

The Information Structure of Clause in The Qur'ān: Corpus Based Study of Functional Grammar

1st Muhammad Yunus Anis¹, 2nd Eva Farhah¹, 3rd Afnan Arummi¹
{yunus_678@staff.uns.ac.id¹}

Universitas Sebelas Maret¹

Abstract. This research investigates the information structure of Theme and Rheme in the Qur'an. Theme is defined as the given information and the Rheme is considered as the new information. Arabic clause is divided into two basic information structures (Theme + Rheme) in the (*mubtada* + *khobar*) and (*fi'l* + *fā'il*). The main data in this research was collected from the Quranic Arabic Corpus (QAC). The first problem in this research is to describe the grammatical aspect of trilateral root *sīn hamza lām* (TRSHL) in the clause of Qur'ānic Arabic Corpus. The second problem is to elaborate the trilateral root *sīn hamza lām* in the construction of Theme and Rheme, this trilateral root occurs in the Qur'an 129 times, in six derived forms. This trilateral root was selected to become the model of the other and the next research about the trilateral root at the Qur'ānic Arabic Corpus. Therefore, the annotation of Corpus in Qur'ān had been developed by investigating the information structure based on the Arabic roots. The lack of annotation in the Qur'anic Corpus can be filled by the data mining about the (Theme + Rheme) construction based on the trilateral root at Qur'an.

Keywords: data mining on Qur'an; Theme and Rheme; given information; new information; Quranic Arabic Corpus; Arabic root.

1 Introduction

The study of annotation has the significant role in the process of developing the data mining in Qur'an. One of annotation study in Qur'an can be elaborated from the analysis of Theme and Rheme. Theme is defined as the given information and Rheme is defined as the new information. The Theme and Rheme in Arabic language can be investigated from two types of clause in Arabic language which is constructed from (*mubtada*' + *khobar*) and (*fi'l* + *fā'il*). The following verse is the example:

اهبطوا مصرا فإن لكم ما سألتم

Ihbithū mishran fainna lakum mā sa'altum

Go into (any) settlement and indeed, you will have what you have asked (2:61)

By elaborating the example at Qur'an verse (2:61), we can conclude that there are two components of information structure. First, the information structure found in the first clause is (*fi'l* + *fā'il*), "ihbithūmishran". The word "ihbithū" is considered as the given information. It means that the speaker and the listener have known the information; the new information in this clause is "mishran". Second, the information structure found in the second clause is

(*mubtada' + khabar*) “*fa inna lakum mā sa’altum*”. The theme in the second clause was “*fa inna lakum*” (given information), it was considered that the speaker and the listener have known this information. What is new from the second clause is “*mā sa’altum*” as the Rheme (new information). It was regarded as the new information between the speaker and the listener or the reader. It is obvious that elaborating about Theme and Rheme in Qur’anic Arabic Corpus can annotate the information structure of clause. It is very important to find the information construction (Theme + Rheme) in the Qur’anic verse to gain the information about the given and the new. It can be assumed that the lack of annotation can be supplied by investigating the Theme and Rheme in Quranic Arabic Corpus.

Corpus in the Arabic language can be defined as *mudawwanah* (مدونة) or *majmū’un-nushūsh* (مجموع النصوص). Corpus also can be defined as “textbooks” being analyzed and compared or as sample of sentences or utterances which are analyzed for their linguistic features. In addition, corpus is also defined as “a collection of materials that has been made for particular purpose. To make a fully beneficial corpus needs annotation [1]. There are three types of annotation in corpus, they are: (1) structural mark up, (2) part of speech mark up, and (3) grammatical mark up. This research elaborates the grammatical mark up in the Al-Qur’an. It can be seen that the novelty of this research is the implementation of functional grammar in the data of Arabic Corpus Al-Qur’an. In addition, the main purpose of this research is to develop the model of annotation in Quranic Arabic corpus. The main data was collected from <http://corpus.quran.com>. The trilateral root *sīn hamza lām* (TRSHL) was selected as the main model to gain the data mining on Qur’an annotation. The previous research about “root” in Arabic and Semitic language was elaborated comprehensively by Rubin [2]. But in this case, Rubin did not elaborate the Arabic roots in Qur’an annotation. There is a gap in this research. All verbs and the majority of nouns have a set of root consonants, usually **three**, thus in this case we call **trilateral**, which is used in conjunction with vocalic templates to form words. The roots have lexical meaning, while the templates have grammatical functions. It seems that the roots in Semitic language, especially in Arabic language have significant role in shaping lexical and grammatical meaning. For example, the Arabic root RKB has meaning associated to the act of riding, as seen in the words *rakiba* ‘he rode’, *yarkabu* ‘he rides’, *rākibun* ‘a rider (NOM.)’, *rakkābun* ‘a (professional) horseman (NOM.)’, *markabatun* ‘vehicle; carriage (NOM.)’, and *murakkabun* ‘mounted (NOM.)’. In each of these words, the roots appear in particular set of vowels, along with a prefix and/or a suffix. It is obvious to develop the investigating of root in Quranic Arabic Corpus to develop the analysis of annotation. The verb “annotate” in Arabic language can be defined by two verbs */yu’alliq/* (يعلق) or */yu’tihā mulāchadzāt/* (يعطى ملاحظات) [3]. From this terminology, it can be inferred that the annotation in Qur’an has special task to give the notes, to explain, to supply a work with critical commentary or explanatory notes, to provide interlinear explanations for words or phrases, in this case Arabic grammatical notes, and to develop the information about verse in Qur’an. The previous research about Arabic annotation, especially in translation studies, was elaborated comprehensively by Almann [4]. Related with the Theme and Rheme in Arabic language, Anis elaborated the translation of an object inside the Arabic Theme and Rheme construction [5]. This research describes the variant structure of objects which occurred in Arabic Theme and Rheme construction using the translation perspective. Anis emphasizes that in Arabic texts, the Theme and Rheme are developed in certain patterns (thematic progression patterns)[6]. Anis outlines and compares the constructions of Theme and Rheme in the Arabic and Javanese languages [7]. On the other hand, Sukesti investigates the construction of Theme and Rheme in the Javanese “NgokoBanyumas” dialect [8]. Suparno also investigates Theme

and Rheme constructions, focusing on the local dialect of Malang, East Java, Indonesia [9]. These recent research studies are primarily structurally focused.

Related to functional grammar, it is obvious to refer to M.A.K Halliday in the analysis. Functional grammar refers to the conceptual framework by which has base in functional one rather than formal one. It is functional in three distinct although closely related senses: in its interpretation (1) of texts, (2) of the system, and (3) of the elements of linguistic structures. Bloor emphasized that the term **functional** concerned with **language in use** [10]. The word *function* is often used for the way a word or a group operates in relation to the other words or groups in the clause. Related to language in use, there are three broad categories known as metafunctions, they are: (1) *ideational metafunction* related to the language use to organize, understand and express the human perceptions of the world and of self-consciousness, (2) *interpersonal metafunction* related to the language use to enable the people to participate in communicative acts with other people, to take on roles and to express and understand feelings, attitude and judgments, and (3) *textual metafunction* related to language use to relate what is said (or written) to the rest of the text and to other linguistic events. This involves the use of language to organize the text itself. This research focuses on how textual metafunction works in Quranic Arabic corpus as main annotation in the data mining.

2 Methodology

This research used distributional method to analyze the trilateral root *sīn hamza lām* (TRSHL). There are three basic steps in this research, namely: (1) collecting the data by using the observation method. The observation began with the elaboration of TRSHL in Qur'anic Arabic Corpus. It can be classified by the variants of clause in the position of TRSHL in the Quranic verse. The TRSHL (<http://corpus.quran.com/qurandictionary.jsp?q=sAl>) is in the scope of Qur'an Dictionary. The trilateral root *sīn Hamza lām* (س آل ل) occurs 129 times in the Quran, in six derived forms:

1. 106 times as the form verb I *sa-ala* (سَأَلَ)
2. nine times as the form verb VI *yatasā'alu* (يَتَسَاءَلُونَ)
3. once as the noun *suāl* (سُؤَال)
4. once as the noun *su'l* (سُؤْل)
5. seven times as the active participle *sā'il* (سَائِل)
6. five times as the passive participle *masū'l* (مَسْئُول)

The TRSHL was randomly selected to be the main data by some reasons: first, because it occurred in the first position at the letter /sīn/. This research becomes the introductory research for the prototype to the other roots at the Quranic Arabic Corpus (QAC) as the modification of annotations; and second, because this research elaborates the information structure in QAC. As Halliday proposed, in order to gain the main theme in a clause is by using the question “what does the clause tell about?” and “What do I tell about?”. By elaborating the root TRSHL, the information structure in QAC can be analyzed by looking at the TRSHL. The Theme and Rheme in QAC can be divided by TRSHL, for example:

يَسْأَلُونَكَ عَنِ الْأَهْلِ قُلْ هِيَ مَوَاقِيْتُ لِلنَّاسِ وَالْحَجِّ
yas'alūnaka 'an = THEME = given information
al-'ahillah = RHEME = new information
qul = THEME = given information
hiyamawāqītulinn-āsiwal-chajji

This research is classified as qualitative research because it adapts to the research problems. The aims of this research are (1) elaborating the grammatical aspects of TRSHL and (2) investigating the information structure of TRSHL in QAC. Miles elaborated that qualitative analysis data in research can be divided into three basic steps, they are: (1) reduction of data, (2) display of data, and (3) conclusion or verification of data [11]. In this research, the reduction of data describes that the clause is reduced and selected randomly from *jumlah ismiyyah* and *jumlah fi'liyyah* which have the TRSHL inside the clauses. The data reduction refers to the process of selecting, focusing, simplifying, abstracting, and transforming the data that appear in written-up field notes or transcriptions. The data reduction in this research focuses on the variant models of Theme and Rheme in the Qur'an and the models of TRSHL inside the clauses. The display of data is divided into two basic problems, they are: (1) grammatical aspect of TRSHL and (2) the TRSHL inside the construction of Theme and Rheme based on SFL theory (systemic functional linguistics). Data reduction is not something separate from analysis. It is part of analysis. Thus the data reduction had been followed by the classification of the conclusion data.

3 Result and Discussion

There are two basic findings in this research. The first finding and discussion in this research is related to the grammatical aspect of the trilateral root *sīn hamza lām* (TRSHL) in the Qur'anic Arabic Corpus. The second problem is to elaborate the trilateral root *sīn hamza lām* (TRSHL) in the construction of Theme and Rheme.

3.1 Grammatical Aspect of TRSHL

This case shows the grammatical aspect of TRSHL in the clause at QAC. Elaborating root in language analysis has significant role, especially in data mining in Quranic Arabic Corpus' annotation. Haywood asserts that Arabic language is rich in derived verb forms which extend or modify the meaning of the root form of the verb, giving many exact shades of meaning [12]. TRSHL in QAC was found dominantly in the verb form. TRSHL was found in QAC just two nouns: seven forms of active participle and five forms of passive participle in some verses. For example in the verse (93:10:2).

وَأَمَّا السَّائِلَ فَلَا تَنْهَرُ

waammas-sāila = THEME (data 1)

falātanhar = RHEME

وَفِي أَمْوَالِهِمْ حَقٌّ لِّلسَّائِلِ وَالْمَخْرُومِ

Wafīamwālihim = THEME (data 2)

Chaqunlis-sā'iliwal-machrūmi = RHEME

By looking both of the data, it can be concluded that the TRSHL in QAC is in the position of Theme and Rheme. The TRSHL analysis in this part focuses on transitivity concept. This concept hopefully becomes the annotation in QAC. Transitivity is defined as the set of options relating to cognitive content, the linguistic representation of extralinguistic experience, either the phenomena of the external world or of feelings, thoughts, or perceptions. There are three main components of the process of transitivity: (1) the process,

(2) the participants, and (3) the circumstances [4]. These three main components of transitivity can be used as annotation of the roots in Qur'an. The participants are divided into six basic processes, they are: (1) actor/ goal (doing process), (2) sensor/ phenomenon (sensing process), (3) carrier/attribute vs. identified/identifier (relational process), (4) behavior (behaving process), (5) sayer/ verbiage/ receiver (saying process), (6) existent (existing process). For example, in data 1 and data 2, the TRSHL has role as the participants. In data 1, the word /as-sāil/ "one who asks" can be classified as participant (GOAL) and in data 2, the word /as-sāil/ can also be classified as participant (GOAL).

Circumstance is divided into five basic circumstances, namely: (1) circumstance of extent: spatial, (2) circumstance of location: temporal, (3) circumstance of manner: quality, (4) circumstance of cause: reason, (5) circumstance of cause: purpose. The TRSHL in this case can be found as circumstance of manner. It can be found from TRSHL as the passive participle in the data 3 below.

لَهُمْ فِيهَا مَا يَشَاءُونَ خَالِدِينَ كَانَ عَلَى رَبِّكَ وَعْدًا مَسْنُورًا

Kāna 'alaRabbika = THEME (data 3)

Wa'danmasūlan = RHEME

The word *masūlan* "questioned" can be classified as circumstance of manner.

Finally, process is divided into six basic processes, namely: (1) material process (process of doing), (2) mental process (process of sensing), (3) behavioural process (process of behaving), (4) verbal process (process of saying), (5) relational process (process of being), and (6) existential process (process of existing). By elaborating the TRSHL in QAC, it can be seen that TRSHL can be divided into two basic form of verb, they are: (1) verb I "to ask" and (2) verb form VI "to question, to ask, and to inquire". By looking at data 4 below, it can be seen that the TRSHL can act as relational process.

وَأَقْبَلَ بَعْضُهُمْ عَلَى بَعْضٍ يَتَسَاءَلُونَ

waaqbalaba'dhuhum 'alaba'dhin = THEME (data 4)

yatasāalūn = RHEME

The word *yatasāalūn* "questioning one another" can be classified as relational process. There are 9 words *yatasāalūn* in QAC.

Finally, the TRSHL is actually related to the mental process. Some processes do not involve material action but involve phenomena best described as states of mind or psychological events. Hence, they are classified and labeled as *mental processes*. Mental processes tend to be realized through the use of verbs like *think, know, feel, smell, hear, see, want, like, hate, please, disgust, admire, enjoy, fear, frighten* [10]. They can be found in data 5 below (79:42:1) *an-nāziāt*.

يَسْأَلُونَكَ عَنِ السَّاعَةِ أَيَّانَ مُرْسَاهَا

Yas'alūnaka 'an 'they ask you' = THEME (data 5)

As-sā'ahayyānamursāhā = RHEME

3.2 TRSHL in the Construction of Theme and Rheme

This part investigates comprehensively TRSHL in the construction of Theme and Rheme. Before doing the analysis, the definition of Theme and Rheme is elaborated first. Theme is the units of language which serve as the starting-point for the message in the text: it is what the clause is going to be about. Theme typically contains familiar, or given, information between the listener and the reader. In this case, the information has already been mentioned somewhere in the text or is familiar from the context [13].

Rheme is part of the clause in which the Theme is developed. Rheme typically contains unfamiliar or new information. The identification criteria for the Rheme is simple; every unit of language in the text that is not Theme is Rheme [13]. The TRSHL can be used as Theme and it indicates Rheme. It can be seen in data 6 and 7 below.

سَأَلُ سَائِلٌ بِعَذَابٍ وَاقِعٍ
Sa'alasā'ilun = THEME (data 6)
Bi'adzābinwāqi' = RHEME

By looking at data 6, the TRSHL derived as the verb /sa'ala/ 'to ask' is classified as the Theme of the clause.

فِي جَنَاتٍ يَتَسَاءَلُونَ
Fījannātin = THEME (data 7)
Yatasāalūn = RHEME

In data 7, it can be seen that TRSHL derived as the verb /yatasālūn/ 'asking each other' is the Rheme of the clause. It can be concluded that the concept of information structure has significant role to develop the annotation of the Qur'anic Arabic Corpus.

4 Media Accuracy of The Hadith Studies

The conclusion in this research is related to the two findings and discussion, they are: (1) grammatical aspect of TRSHL in the clause, and (2) TRSHL in the construction of Theme and Rheme. By looking the grammatical aspect, it can be concluded that TRSHL is constructed dominantly by the verb, two nouns, seven forms of active participle and five forms of passive participle. The roots in Arabic language is related to the concept of transitivity. This concept can be main annotation of Quranic Arabic Corpus (QAC). In the TRSHL, the active participle can be identified as participant (goal). On the other hand, the passive participle can be identified as the circumstance of manner. TRSHL in QAC can be divided into two basic forms of verb, for example, the word yatasāalūn as derivative form of TRSHL can be classified as the relational process. Finally, the three main components of the process of transitivity: (1) the process, (2) the participants, (3) the circumstance, can be main annotation in QAC and they can be the model of roots analysis in QAC. It can also be concluded that The TRSHL can be Theme and Rheme.

Acknowledgment. This research was supported by LPPM UNS. The Institution of Research and Servitude toward the Society, under the Fundamental research grant (Hibah Penelitian Fundamental PNBP UNS 2018).

References

- [1] Meyer F C 2004 *English Corpus Linguistics An Introduction* (Cambridge: Cambridge University Press)
- [2] Rubin A D A B 2010 *Introduction to the Semitic Languages* (USA: Gorgias Press LLC)
- [3] Collin P H, Kassis N and Angel T 1990 *Harrap's English Dictionary for Speakers of Arabic* (Canada: Kernerman Publishing, Inc)

- [4] Almanna A 2016 *The Routledge Course in Translation Annotation: Arabic-English-Arabic* (New York: Routledge)
- [5] Anis M Y 2016 The Translation of An Object in Arabic Theme and Rheme Constructions *International Conference on Linguistics & Translation Studies* pp 255–266
- [6] Anis M Y, Arifuddin A and Farhah E 2015 Pengembangan Tema dalam Buku Al-Qira'ah Ar-Rasyidah untuk Pelatihan Menulis Kreatif Bahasa Arab *Arab. J. Pendidik. Bhs. Arab dan Kebahasaaraban* **2** 144–53
- [7] Yunus Anis M, Farhah E and Arummi A 2018 Arabic-Javanese Translation: A Case Study on Textual Equivalence of Islamic Moral Ethic Book *Proceedings of the International Conference on Culture and Language in Southeast Asia (ICCLAS 2017)* (Paris, France: Atlantis Press) pp 161–5
- [8] Sukesti R 2011 Tema – Rema dalam Bahasa Jawa Ngoko Dialek Banyumas: Kajian Penataan Organisasi Informasi *J. Hum.* **23** 219–28
- [9] Suparno 1993 *Konstruksi Tema Rema dalam Bahasa Indonesia Lisan Tidak Resmi Masyarakat Kotamadya Malang* (Jakarta: Pusat Pembinaan dan Pengembangan Bahasa)
- [10] Bloor T and Bloor M 2004 *The Functional Analysis of English* (London: Arnold)
- [11] Miles M B and Huberman A M 1994 *An Expanded Sourcebook Qualitative Data Analysis* (London: Sage Publications)
- [12] Haywood J . and Nahmad H M 1979 *A New Arabic Grammar of The Written Language* (London: Lund Humphries)
- [13] Eggins S 2004 *An Introduction to Systemic Functional Linguistics* (New York: Continuum)