ISSN: 2666-2795 Vol. 7 No. 1, 2022, Netherland

International Journal of Applied Engineering Research

DESIGN AND AUTOMATIC EXTRACTION OF ARUD RULES FOR URDU POETRY

Asia Zaman [1]*, Zia-Ud-Din[1], Noman Khan[1], Muhammad Daud Khan[1], Javed Iqbal[1]

aasiazaman123@gmail.com,
ziasahib@gmail.com,
nomibaloch8916@gmail.com
daud@gu.edu.pk
javedmarwat7@gmail.com
[1]Institute of Computing and Information Technology, Gomal University, Dera Ismail Khan, KpK, Pakistan.
[1*]Corresponding author: (aasiazaman123@gmail.com).

Abstract:** Arud, also known as "ilam-Arūḍ" (علم عروض), is a type of metrical system that is used to structure verse. It is a product of classical Arabic poetry. In the 8th century, it was developed by an Arab grammarian named Al-Khalil ibn Ahmad al-Farahidi. In order to organize and generate consistent metrical structures, which are referred to as "bahar" (plural: buhur). These guidelines were established to generate rhythmic pattern. This paper delineates the structure of Urdu words, the composition of Urdu poetry using meters and their sub-meters, and gather the rules that are implemented in the Arud system to maintain tone, rhythm, and harmony. This paper investigates the scansion principles that are essential for the conservancy of the rhythm of a poem. The Arud system originated in Arabic and was subsequently adapted and incorporated into Persian and Urdu, Turkish, Punjabi and Asian language poetry as a result of the historical and cultural interactions between these linguistic traditions. In this paper we discuss meters that are the specifically employed in Urdu poetry. The Arud system in Urdu poetry guarantees that the poetry adheres to established metrical patterns, thereby preserving the aesthetic and rhythmic qualities of verse. This integration has allowed Urdu poets to create works that are not only linguistically appealing but also adhere to a traditional and structured format, which enhances their phonetic appeal.

Keywords: ***Arud System, Urdu words structure, Scansion Rules, Natural language processing, Pattern Matching***

1. Introduction

The knowledge of Arud علم العروض is crucial in Urdu poetry because it provides an appropriate process for ensuring that the poet's work complies with weight regulations [1]. Khalil bin Ahmed is the founder of Ilam-e- Aruud [2][4][5]. He designed this knowledge by using the ideologies of the melody. He uses basic knowledge of Arabic poetry; his focus was that every verse ensures some pattern to give proper weight. Weight or wazan is a poetic meter which is the rhythmical arrangement of mutaharik متحرک"Movant" and sakin "quiescent" ساکنletters into lines [3]. To test and measure weight, the ancients devised the Arud system. Researchers in Arabic region laid the groundwork for this knowledge. This knowledge was later applied to Persian and Urdu poetry. It can be viewed as a scale for evaluating the harmony and musicality found in the poems. It has been the subject of numerous books [4][5][6][7][8][9][10], however complete mastery is difficult to achieve. To become proficient in poetry, one does not necessarily need to acquire this expertise. Only what is required to utilize proper language with metric rhythm should be known by an individual. In Urdu poetry, "Arud" is a metric system that is used to determine the meter and rhyme of a poem. "Ilam" is a term used in Urdu poetry to refer to the knowledge and understanding of the rules of prosody. Hence "Ilam Arud" would refer to the knowledge and understanding of the rules of prosody in the context of Urdu poetry using the Arud metric system.

The remaining paper consists of 5 sections. Section 1.2 describe importance of Arud System. Section 2 explores the literature about Arud rules and their implementation in Urdu and Urdu like languages. Section 3 present structure of Urdu poetry. In section 4 scansion, Arud rule formation methodology is depicted with examples. Finally, the last section 5 provides the conclusion of this research.

1.2. Importance of Arud System

The Arud system has many rules to retain Urdu poetry quality. Researchers in [10][9][8] have figure out following benefits to tell the importance of knowledge of Arud.

i. For an understanding of verse wazan/weight and types of verse metrics.
ii. Differentiate between each metric from the other one.
iii. Understanding the pros and corns of each metric.
iv. What modifications and variations can be possible in any given verse?
v. Knowledge of unfamiliar phonetics and their impact on the verse.
vi. Helps in differentiating two Urdu genres Nazam and Nasar.
vii. Differentiating between actual and false metrics.
viii. Identifying all metrics that are authentic in Urdu poetry but not implemented in it due to its complex structure.

It shows that the learning Arud knowledge is as important as learning music Rhythms [11-12]. Without knowledge of rhythms, the singer can't sing a quality song with extra impact, the same reflects in good poetry as well. The sound that comes from only one letter is referred to as a word.

2. Related Work

Sufficient research has been conducted on foreign languages such as English, Chinese, and Arabic. However, there has been significantly less research conducted on Urdu language. Noteworthy contributions in the field include research on authorship attribution [2], sentiment analysis [29], content management [30], Urdu stemmer [28].

In Urdu poetry Arud system is the key to maintain poetry quality. Arud system was originated in Arabic language by Al-Khalil ibn Ahmad al-Farahidi [12-15]. He introduced 16 meters, having unique rhythmic patterns [14]. Arud system is used in all Arabic derived languages like Persian, Urdu, Punjabi, Ottoman and some Urdu like eastern languages [14]. Arud system using the concept of the melody by ensures some pattern to give proper weight/wazan. Wazan is a poetic meter term which is the rhythmical arrangement of mutaharik and sakin letters into lines [13]. The base knowledge of Arud meter found in chapters 2-4 of the book of Captain Pybus [4] chapters 1–7 of the book Pritchett and Khaliq [5] and knowledge of Urdu poetry metrics and its sub-meters collected from authentic books [6-10].

Consequently, we came across academic literature on Arud system for detecting Arud meter that primarily concentrated on the Arabic [14–23] as well as Punjabi [24,25], and Ottoman languages [26]. Moreover, algorithms have been developed to evaluate Arud meter detection in Arabic language. Howeve, there is no research found on algorithmic meter detection in Urdu poetry. In literature review all Urdu poetry structure and meter detection work is done in published books [1][4-10] in late eighteenth and early nineteenth century. In Two most prominent web systems Aruzz [26] and Rakhta [27] facilitates Urdu poetry learners.

3. Structure of Urdu Poetry

The Urdu poetry contents use structures like words, syllables, and letters. A word is the smallest, independently meaningful unit of a language. Words are "what people generally utter in common usage." Every sentence of the poem is scanned and broken down into a series of syllables. It is used for metrical purposes. A syllable is an essential concept for understanding phonological structure. Urdu language compose of Urdu alphabet from alif 'الف ا' to yaa 'ے'. Appendix I.

Syllables are important for metrical considerations. No line of poetry can be scanned until it is first broken into a series of syllables. The process of breaking words into identified syllables is called scansion. Each word in a line of poetry must be divided into a metrical syllabus according to the below criteria.

1. **Long syllables:** Any syllable consisting of two letters is considered as long syllable
2. **Short Syllabus:** any syllable consisting of one letter is a short one.
3. **Flexible Syllable:** A syllable that consists of only two letters and may be used as long or short, at the poet's pleasure, will be called a flexible syllable.

Every verse of the poem is scanned and broken down into tokens using certain rules. The process of tokenizing verse into letters is called scansions تقطیع. The basic principle of Urdu scansion is based on the distinction between movantمتحرک "mutaharik"بجائےبلند and "sakin" ساکن letters that are termed as بجائےکوتاہ .

Movant letters: The letter accomplished by one of the diacritics, zabar, zer, and pesh is termed as متحرک "mutaharik e.g. دلِ the letter دِ is mutaharik as the zer accomplishes it. Movant letters categoriezed as two types of Harakat i.e. Harkat -e-mutlik, حرکت مطلق harkat-e-basakon حرکت باسکون.

- Harkat Mutlik حرکت مطلق is that Harkat where sakin is not instantly attached to it. e.g. نظر here ن is harkat Mutliq.
- Harkat-e- BaSakon حرکت با سکون is harkat where sakin immediately comes after harakat, such as in گل or in بلبل.

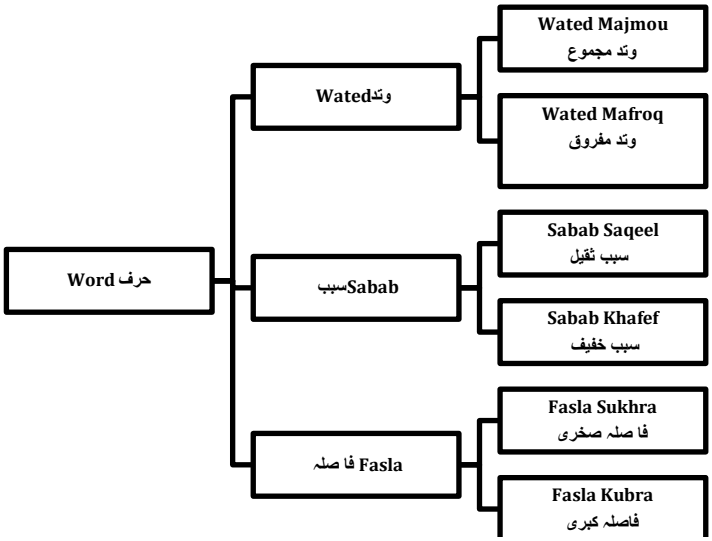

Figure 1: Urdu words structure hierarchy

Quiescent letters: When a consonant is unaccomplished by any diacritics, it is termed "sakin" ساکن. Sakin is defined as a letter of the alphabet that represents the end of a sound. Diacritics such as tashdeed تشدید are meant to represent sounds that are first treated as a sakin and then as mutaharrik [12]. Usually an Urdu word combination of sakin and mutaharik letters. Sakin words are of two types, Sakon-e-Mutlik, سکون مطلق Sakon-bul-harkat سکون بالحرکت.

- Sakon Mutlik سکون مطلق that word in which there is no Harakat added on letter, e.g. آس .
- Sakon Bil-Harkat, سکون بالحرکت that word in which there is Harakat added on letter, e.g اب.

Urdu word has three main categories, as shown in figure 1. The sound that comes from a combination of two letters is called Sabab. Sabab is of two types sabab-khafef سبب خفیف (first letter is mutaharik second one is sakin) e.g. abb اب، tuتو and sabab Saqeel سبب ثقیل(both letters mutharik) e.g. Narrhaa نرحا). The sound that comes from three letters is referred to as watadوتد in the Urdu language. When the first two letters are mutaharik and the last third one is sakin such words are referred to as watad majmoue وتد مجموع, e.g. diya , دیا , while the second type of three-letter watad mafrooq وتد مفروقconsist of first and last letter mutahark and the middle one is sakin e.g. پان، بخت. The sound comes from four letters referred to as fasla فا صلہ . when the first three letters are mutaharik and the last fourth one is sakin called fasla-sukhra فا صلہ صغریٰ e.g. نگینا ,Fasla-Sukhra is a combination of sabab-khafeef and sabab Sakeel. While the sound of five letters is referred to as fasla-Kubraفاصلہ کبریٰ. Five letters word is a combination of sabab-saqeelسبب ثقیل and watad-mafroq وتد مفروق.

Urdu poetry is a combination of the above-said word categories. These categories create the basis for the metrics afail افاعیل. Variant arrangement of these words based on certain rules defines Asool Afails اصول افاعیل. Each afail has its numerical value. The combination of different afail in certain patterns define wazen or meter of that verse. Wazan, also known as weight, is a technique of portraying an Urdu word following its afaail, which is actually its phonetic pattern rather than how it is transcribed. The affails split into two categories: a huge number of variants, or muzaahif مضاعف, and a small number of original ones, or salimm [4].

The arrangement of a certain feets/afaaill افاعیل in a specific order is termed as meter بحر in Urdu poetry. Meter is the rhythmical arrangement of mutaharik متحرک and sakin letters into lines. Arabic Arud lies in the system of sakin and Mutharik letters. "Movent" Harkat حرکت and "quiescent" Sakin. If any Urdu letter accompanied by any one of three zabr, zeer, pesh, is referred as movent harkat letter. While Urdu consonants letters unaccompanied by any harkat are referred as quiescent or Sakin. For example بلبل (Nightingale bird)، (Mosque) مسجد these two Urdu words are same in quantification/syllable arrangement. س and د are sakin in masjid word, while م and ج are mutaharik because of Zaber on م , and zeer on ج . similarly word بلبل have two sakin and two mutaharik words. With same sequence of Mutaharik, Sakin, Mutaharik, Sakin. All meters in Urdu are composed of one or more of eight standard feet/Afail. These eight feets divided into two types one is known as five-lettered (khamasiفعولن ، فاعلن)خماسی، and second type is seven letters (Saabaii سباعی)، مفاعیلن، فاعلاتن، مستفعلن ، مفعولات، مفاعلتن متفاعلن، . these Afails can be used as its catalexis زحاف . These Standard feet/ Afail with their Zehaf زحاف are describe in table 1. Zehaf word means mould from the original one. Meters comprise numerous sub-meters, each with a unique arrangement of "afail" and its "zehaf" or catalexis. Catalexis is the term used to describe any atypical alteration in the meter. These modifications can be implemented by incorporating or eliminating "sakin" or "mutaharik" words from the afail. The zehaf is subject to specific regulations. The principles that are most frequently implemented, or their combinations, are as follows:

- Converting mutaharik words to sakin.
- Eliminating sakin or mutaharik letters.

- Increasing the number of letters in the afail.

The zehaf of each foot is used to define the number of meters and their sub-meters.
Table 1 provides a comprehensive list of all potential feet, along with their catalexis.The rhythmical arrangement of feets mentioned in table 1 generates a pattern that is called meter, Beher, weight, or wazan, which are distinguished by unique rhythmic patterns. Each pattern consists of a combination of afails افاعیل. Arrangement of afail in certain sequence creates specific meter. This pattern creates a melody when uttering the verse following a certain melodic aspect that is audible when read aloud.
Nineteen meters in Table 2 are most commonly used in Urdu poetry. Recently, in the nineteenth century one more meter Jameel جمیل is added in Arud system [13]. Urdu poetry meter has its roots in the Arabic poetic tradition. This system enhances the aesthetic coherence and auditory appeal of poetic creations. By using table 2 and table 1, it was found that there can be number of basic meter's sub- meters. We observed that from basic 20 Arud meters there are over 290 drived submeters used in Urdu poetry. The list of 290 meters can be found in Annexure I. List describe numbers of identified sub-meters for each basic meter shown in Table 3.
Every appealing poetry can be written by following any one of the given rhythm in table 2. Poetry in Urdu language can be implemented according to the defined meters by applying scansion rules on a verse. A couplet can be created by applying above meters. These meters are categories into two main types i.e. simple meter, compound meter. The compound meter is a combination of two or more feet. Feet in verse are divided into the following parts, the first foot is termed as Sadar صدر, the second, and third foot termed Hashuu حشو the fourth is Aruzz عروض, the fifth foot is abtada ابتدا, the sixth & seventh foot is called Hashuu حشو and the eighth foot called Zerb ضرب .A meter of 8 feet (4 per hemistich) is called Musamanمثمن.A meter of 6 feet (3 per hemistich) is called Musadas مسدس . A meter of 4 feet (2 per hemistich) is called Murabaa مربع. The couplet feet compose of the combination of any meters.

Musadas Example:

ٹھیک ہے خود کو ہم بدلتے ہیں
شکریہ مشورت کا چلتے ہیں
Romen: (Teek hy khud ko hum badalty hain
Shukriya Mashwarat ka chalty hain)
Translation: Well, we change ourselves, Thanks for the advice.
Feet: فاعلاتن مفاعلن فعلن
Meter: خفیف مسدس مخبون محذوف مقطوع

Musaman Example:

یقیں عمل محکم پیہم،محبت فاتح عالم
جہاد زندگانی میں ہیں یہ مردوں کی شمشیریں
Romen
(Yaqeen Mohkam, Amal Peham, Mohabbat Faateh-E-lam
Jahad-E-Zindagani Mein Hain Ye Mardon Ki amsheerain)

Translation:
Firm certainty, eternal action, the love that conquers the world— These are the swords of men in the holy war of life.
Feet: مفاعیلن مفاعیلن مفا عیلن مفاعیلن
Meter: ہزج مثمن سالم

4. Scansion تقطیع

The scansion word comes from the Arabic word, which means to cut or divide into small parts. This is one of the basic parts of illm-Arud علم العروض. In Urdu poetry, "scansion" refers to the process of analyzing and determining the meter and rhyme of a poem. The rules of scansion in Urdu poetry are based on the "Arud" metric system.
The rules of scansion in Urdu poetry are based on the "Arud" metric system, which is used to determine the meter and rhyme of a poem.
The Arud system is based on the number of syllables in a line of poetry, and the specific pattern of long and short syllables. The basic unit of meter in Aruuz is the "qafia" قافیہ, which is a group of two or three syllables. The number of qafias in a line of poetry determines the meter of the line.
The rhyme in Urdu poetry is determined by the final syllable of the line, which must match the final syllable of at least one other line in the poem. The rhyme scheme of the poem is determined by the pattern of rhymes used throughout the poem.

It's important to note that many poets use the rules of Arud more flexibly or creatively and deviate from the strict rules of the system. Additionally, scansion can be difficult to do by machine, as it requires a thorough understanding of the rules of Arud, as well as a deep understanding of the cultural and historical context of the poem.

4.1. Rules of Scansion

Overall, scansion in Urdu poetry can be complex and challenging due to many rules and variations in the Arud metric system, as well as the complexity of the Urdu script and diacritics.

Rule #1: Dropping letters

The first rule is that only those letters which are pronounced are taken into account and those which are not pronounced are rejected. For example, بالکل contains five letters, however it is pronounced without the sound of ' ' ا as if it is written like بلکل. During scansion only articulate letters are taken into account. Some most important rules [9] implemented in scansion procedures are discussed below:

The following are the letters most commonly dropped in scansion.

a. ا aliph in Arabic words such فی ا لحا ل having 7 seven letters which are scanned as if split into فیلحال six letters word.
b. ھائےمخلوطه Do Chashmi h ھ, is aspirated consonant, although it affects the pronunciation of the syllable in which it occurs, is metrically invisible. e.g. کھانا is 5 letter word scanned as کانا 4 four letters word.
c. یائےمخلوطه in کیا kya meaning "what" will be scanned as کا. Having 3 letters instead of 4
d. واو معدوله letter و used in words like خوش، خود is scanned as if split as خش , خد respectively.
e. نون غنه ں , the nasal N if preceded by any حروف علت i.e. ا ، ی ، و . e.g. word جھونکا having 6 letters will be scanned as جو کا have 4 letters.

Rule #2: Doubling letters

The second rule is related to characters which are counted twice.

a. مشدّد letters are of course counted as two e.g.فعولن = تر د د = ترّد, and واللہ= و ل لا ہ
b. الف ممد ود ہ Aliph counts as two, the first being mutaharik the second sakin as آنا to انا ا the above word was 3 letters scanned as 4 letters.

Table 1 Basic feet with their Catalexis

c. Hamzated واو count as two as in the wordرؤف۔ which is scanned asرووف.

The persian izafat is counted in scansion, e.g. عروسی چمن = عروسِ چمن.

Rule #3: Positioning of Two Sakin Letters

If two sakin letters occur together except at the end of the line (where both are counted Sakin) , otherwise the first is referred to as sakin and the second mutaharik.

Rule #4: Positioning of Three Sakin Letters

If three sakin letters occur together in a word but not at the end of line , the first remains sakin, and the rest of the two is governed by poetic license. If the second and third sakin letters are preceded by ا، و، ی the second sakin is not counted and the third becomes mutaharik i.e. دوست will scan as دست.

Rule #5: Poetic Desertion

According to rule #5 there is flexibility for the poet that he can switch sakin letter into mutahark letter and vice verse to keep the rhythmic situation consistent for metrical purpose in Urdu poetry. This rule helps poets to get out of the bed rhythmic situation. The rule of poet desertion complicates the implementation of Arud principles.

For example د و ستدارئ, کا شت کا ر ی، will be scanned as دوسدارئ، کاشکاری۔ by converting mutaharik letter 'ت' as silent letter .

Table 1 List of Afail & its Zehaf

	Afaiil/feet افاعیل	Zehaf زحافات
1	Fauolan فعولن	فعول، فعلن، فعول، فعل، فعولان، فعل، فع , فعلان
2	Faeelan فاعلن	فعلن، فعلن، فع، فاعلان،فاعلاتن، فعلاتن، فعل، فعلان، فعلن
3	Mafaeelan مفاعیلن	مفاعلن، مفاعیل، مفعولن، مفاعیل، فعولن، فعل، مفاعیلان، مفعول،فاعلن، فاع، فع، فعلان، فعلن، ,مفعولان, فعول
4	Faeelaatan فاعلاتن	فعلاتن، فاعلات، فاعلات، فاعلن، مفعولن، فاعلیان، فعل، فاع، فع، فعلات، فعلن، فعلان، فعلان، فعلن، فعلیان، فعلان، مفعولان
5	مستفعلن Mustafeelan	مفاعلن، مفتعلن، فاعلن، مفعولن، فعلن، مستفعلان، مفعولان، فعلان، فعولن، فعلتن، فاع، فع، ،مفاعلان، مفتعلان، فاعلاتن، فعلتان، مفاعلاتن، مفتعلاتن
6	Mafeolat مفعولات	فعولات، فاعلات، مفعول، مفعولان، مفعولن، فعلن،، فاع، فع، فعلات، فعولان، فاعلان، فعلان، فعولن، فاعلن، فعلن
7	Mutfaeelan متفاعلن	مستفعلن، فعلاتن، فعلن، متفاعلان، متفاعلاتن، مفاعلن، مفتعلن ، مفعولن، فعلن، مستفعلان، مستفعلاتن، مفاعلاتن، مفاعلان، مفتعلاتن، مفتتعلان
8	Mafaeelatan مفاعلتن	مفاعیلن، مفتعلن، مفاعلن، مفاعیل، فعولن، مفعولن، فاعلن، مفعول

S.No	Metric Name	Meter Feet/ Afaiill بحر افاعیل
1	Mutakarib متقارب	فعو لن، فعو لن،فعو لن، فعو لن
2	Rajiz رجز	مستفعلن، مستفعلن مستفعل مستفعلن
3	Ramel رمل	فاعلاتن، فاعلاتن فاعلاتن، فاعلاتن
4	Kamel کا مل	متفاعلن، متفاعلن متفاعلن، متفاعلن
5	Wafer وا فر	مفاعلتن، مفاعلتن مفاعلتن، مفاعلتن
6	Hazej ہز ج	مفاعیلن، مفاعیلن مفاعیلن، مفاعیلن
7	Mutadarik متدارک	فا علن، فا علن فا علن، فا علن
8	Maded مدید	فاعلاتن، فاعلن فاعلاتن، فاعلن
9	Taweelطو یل	فعو لن، مفاعیلن فعو لن، مفاعیلن
10	Baseet بسیط	مستفعلن، فاعلن، مستفعلن، فاعلن
11	Jaded جدید	فاعلاتن، فاعلاتن، مستفعلن
12	Khfeef خفیف	فاعلاتن، مستفع لن، فاعلاتن
13	Sariee سریع	مستفعلن، مستفعلن، فاعلات
14	Qareeb قریب	مفاعیلن، مفاعیلن، فاعلاتن
15	Mujtass مجتث	مستفعلن، فاعلا تن مس تفع لن، فاعلا تن
16	Muzare مضارع	مفاعیلن، فا ع لا تن مفاعیلن، فا ع لا تن
17	Muktazeb مقتضب	مفعولات، مستفعلن مفعولات، مستفعلن
18	Munsareeح منسر	مستفعلن، مفعولات مستفعلن، مفعولات
19	Mushakeelمشاکل	فاعلاتن، مفاعیلن، مفاعیلن
20	Jameel جمیل	مفاعلاتن مفاعلاتن مفاعلاتن مفاعلاتن

Table 2: Arud system basic meters with Afails

Table 3: Number of Basic meters's Sub-meters detected in Urdu poetry

S.No	Meter Name	No of its sub-Meters
1	Mutakarib متقارب	24
2	Rajiz رجز	24
3	Ramel رمل	12
4	Kamel کا مل	19
5	Wafer وا فر	5
6	Hazej ہز ج	48
7	Mutadarik متدارک	28
8	Maded مدید	17
9	Taweelطو یل	3
10	Baseet بسیط	10
11	Jaded جدید	5
12	Khfeef خفیف	14
13	Sariee سریع	6
14	Qareeb قریب	8
15	Mujtass مجتث	16
16	Muzare مضارع	15
17	Muktazeb مقتضب	7
18	Munsareeح منسر	7
19	Mushakeelمشاکل	8
20	Jameel جمیل	5

Rule #6: Assigning weight

Some **common rules** for assigning weight to words are based on the position of vowels and consonants. The Urdu language has 36 alphabets, in which some letters like ا، آ،ی،ے،و are treated as vowels while the remaining letters are treated as consonants.

a. The consonant will have the value of '1'. As 'j' in 'aaj', 'r' in 'pyar' 'ba' in 'bahaar'.
b. Two consonants, if spoken together, will have a value of '2'as in 'hum', 'ab' 'chal' 'yad' in 'shaayad'
c. Vowel 'e ی ', 'o و' , 'u ے ', 'i ا' have a value of '1' and with a consonant will have the value '1' e.g. 'mein', 'ae', 'ku' in 'kuch', 'lo' in 'lori'.
d. If two consonants joined by an above vowel and spoken together have the value of 2 as 'tum' تم , 'fir' پھر, and 'dil' دل all have a value 2.
e. Aa' ا will alone or with consonants will always have the value of 2 as 'kaa' کا in 'kaam' کام , 'yaa' یا in 'qayaamat' قیامت 'haa' ہا in 'bahaar' ب.
f. Vowel 'ee', 'oo', 'ao' will have a value of 2 as 'bhi' in 'kabhi', 'hi' in 'nahin', 'ho' in 'kaho'.e, o, u, alone or with a consonant, have a value of 1 or 2 provided by the scratch of a syllable, e.g. 'mere میرے' may have 1-1, 2-1, 2-2, or 1-2 anyone, 'tumhare، تمھارے' maybe 2-2-1 or 2-2-2.

The Arud system is challenging since poet diacritics exist, which allow the poet to drop any sakin word to mutahaik or treat any mutaharik letter as sakin in order to maintain proper meter. Such letters and words are classified as flexible syllables. It is abbreviated as X. X flexible syllables can be quantified as either one or two.

By applying these rules, we can verify any verse in the defined meters mentioned in table 3. Taking examples of different verses can describe the implementation of Arud rules.

4.2. Describing Arud Rules with Examples

In order to understand the impact of Arud rules on the Urdu verse we need to go through some of the examples. If we consider Mirza Ghalib's famous verses to apply arud rules on it. This verse tokens will be quantified by applying Arud rules and then identifying its syllable.

Let take the verse 1 is as an example of implementing Arud rule in Urdu poetry.

نکلنا خلد سے آدم کا سنتے آئے ہیں لیکن

All the words in the verse 1 are parsed as below in figure 2.

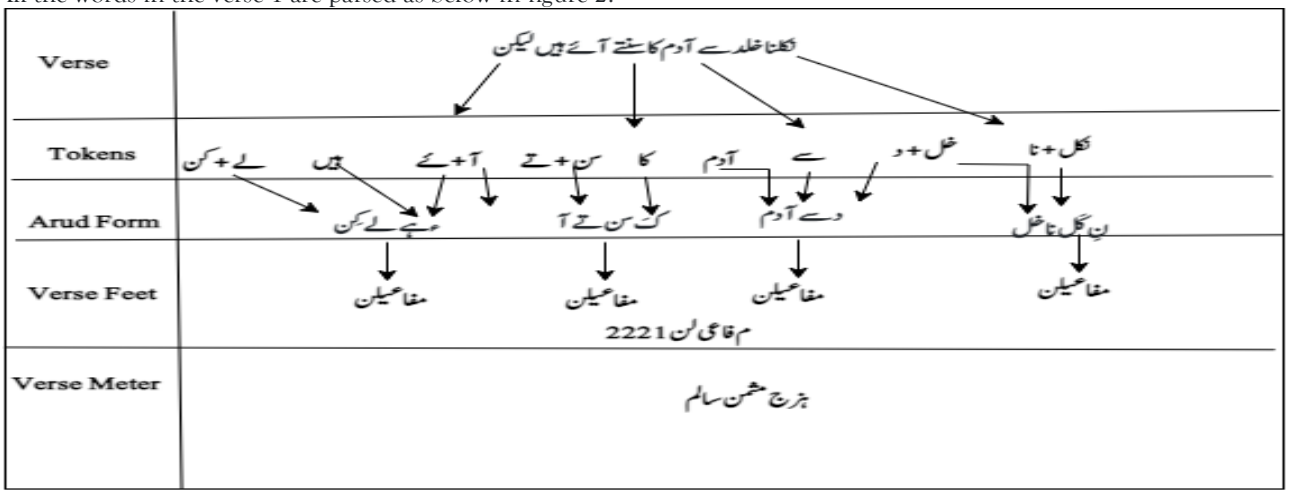

Figure 2: Verse 1 example

To get Arud form of the verse following rules mentioned in section 4.1. Implemented on the verse 1.
Scansion Rule #1.a implemented on word "کا" of the verse where "ا" Alif is omitted from the word کا "kaa" as کَ.

Scansion Rule # 1.c یائےمخلوطہ applied on word آئے where "ی" of the ئے omitted , similarly Rule 1.e applied on word " ہیں" where ں is omitted as it doesn't contain any weight in meter.

Verse 2: ستم کو ہم کرم سمجھے جفا کو ہم وفا

By applying Arud rules where necessary verse will be converted into Arud-form as in figure 3.

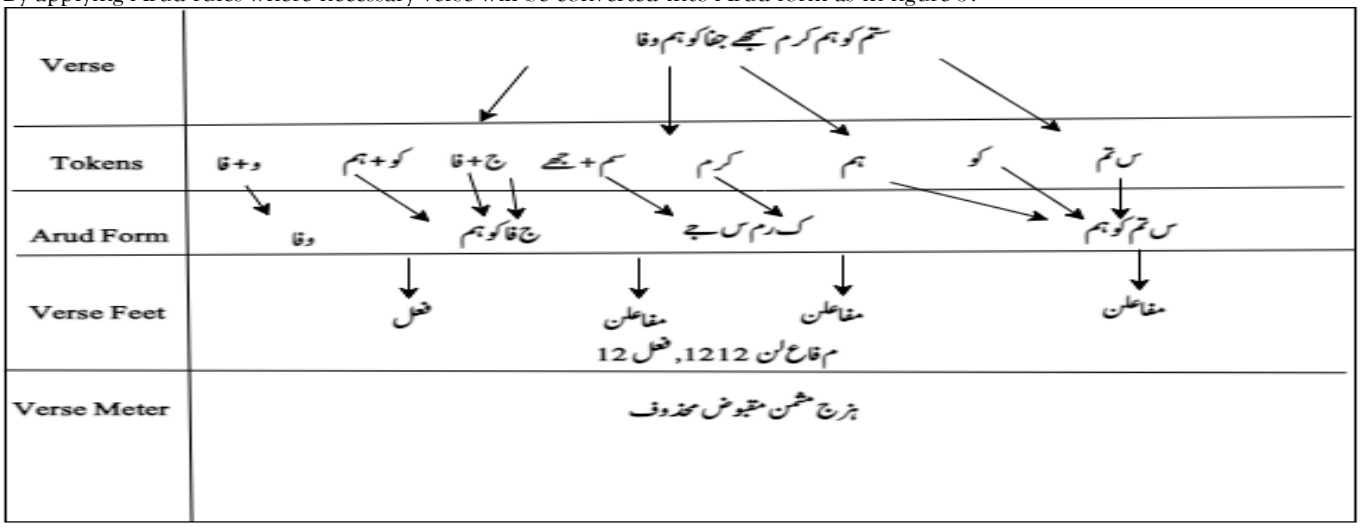

Figure 3 Verse 2 Example

Here poet uses diacritics by applying rule #6 to drop word سمجھے as ج - س.

Arud form ensures that these verse patterns help poets to maintain a rhythmic structure that is both pleasing to the ear and consistent with traditional forms. These patterns ensure that the verse flows melodically and adheres to the classical standards of Urdu poetry. Arud –form pattern clearly differentiates word's syllables and can be easily quantify.

5. Conclusion

The paper thoroughly examines the Arud system, illustrating its significance in the evolution of Urdu poetry. It delves into the structure of Urdu poetry and the precise techniques utilized to convert Urdu poems into Arud form by removing words that do not contribute to the verse's phonetics. Arud's principles do not apply to all written words; only phonetically relevant terms are quantified and considered. In this document, all Arud system structures, such as feet, catalexis, and scansion rules, are collected and placed in one location. In light of the paradigm transition from print to digital media, there is a huge increase in need for automated systems that recognize Arud systems. These principles make it easy for new Urdu poets and readers to identify the meter and rhythm of any verse. According to the literature review, Arud system research in Urdu poetry has been ignored due to limited resources, and all work on the Arud system dates back to the early 19th century. There is a need to investigate the Arud method in simple words so that new learners can understand it and apply it to Urdu poetry composition. As arud system counts only that words which are phonetically utterd so there is need to have a system that recognize speech to convert voiced verse into arud form.

References:

1. Nadim Balghi. Tafheem ul Arud, Behar Afseet press, Patna; 1985.
2. Nida Tariq "Identification of Urdu Ghazal Poets using SVM" Mehran University Research Journal of Engineering & Technology, Vol. 38, No. 4, 935-944 October 2019
3. Waqas Anwar, Xuan Wang, Xiao-Long wang, "A survey of automatic Urdu language processing", Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, 2006
4. Pybus GD. A Text-book of Urdu Prosody and Rhetoric. Baptist Mission Press; 1924.
5. Pritchett FW, Khaliq AK. Urdu Meter: A Practical Handbook. FW Pritchett and KA Khaliq; 1987.
6. Mirza OajLakhnawii. Mikyass-ul – Ashaar. Prosody book, 1875.
7. Sarwar Alam Raz Sarwer, "Kitab-r-Arooz", 2008.
8. Book "Tafheem ul Arooz" by Jamal ud-din Jamal
9. Book "Mikyass-ul – Ashaar" by Mirza OajLakhnawii"Aruzz Asbaq" by SarwerAlam Raz Sarwar.
10. Mohamed Y. Dahab, Ablah AlAmri, Bayan Bagasi "Automatic Identifying Rhythm of Arabic Poem" International Journal of Computer Applications (0975 – 8887) November 2016
11. Qureshi R. Tarannum: The chanting of Urdu poetry. Ethnomusicology. 1969 Sep 1:425-68.
12. Molwi Najmal-ul- Ghani " Bahr-ul-Fasahet" Tayaba Noor printers, Lahore, 2018.
13. Alnagdawi MA, Rashideh H, Aburumman F. Finding Arabic poem meter using context free grammar. Journal of Communication and Computer Engineering. 2013;3(1):52-9.
14. Shalabi R, Kana'an G, AL-Jarah A. Computing System for Analyzing Arabic Poems Meter. Yarmouk Research, Yarmouk University. 2003.
15. Alsharif O, Alshamaa D, Ghneim N. Emotion classification in Arabic poetry using machine learning. International Journal of Computer Applications. 2013 Jan 1;65(16).
16. Dahab MY, AlAmri A, Bagasi B, AlMalki E, AlBeshri O. Automatic Identifying Rhythm of Arabic Poem. International Journal of Computer Applications. 2016;975:8887.
17. Baïna K, Moutassaref H. An efficient lightweight algorithm for automatic meters identification and error management in Arabic poetry. InProceedings of the 13th International Conference on Intelligent Systems: Theories and Applications 2020 (pp. 1-6).
18. Abuata B, Al-Omari A. A rule-based algorithm for the detection of arud meter in CLASSICAL Arabic poetry. Update. 2018 Jul 1;26:06.
19. Deo A, Kiparsky P. Poetries in contact: Arabic, persian, and urdu. Frontiers of comparative metrics. 2011:147-73.
20. Yousef WA, Ibrahime OM, Madbouly TM, Mahmoud MA. Learning meters of Arabic and English poems with Recurrent Neural Networks: a step forward for language understanding and synthesis. arXiv preprint arXiv:1905.05700. 2019 May 7.
21. Zeyada S, Eladawy M, Ismail M, Keshk H. A Proposed System for the Identification of Modem Arabic Poetry Meters (IMAP). In2020 15th International Conference on Computer Engineering and Systems (ICCES) 2020 Dec 15 (pp. 1-5). IEEE.
22. Al-Talabani AK. Automatic recognition of arabic poetry meter from speech signal using long short-term memory and support vector machine. ARO-The Scientific Journal of Koya University. 2020 Apr 14;8(1):50-4.
23. Kaushal A, Dutta K. Analysis of Performance Metrics for Classification of Punjabi Poetry using Machine Learning Techniques. In 2022 International Conference on Artificial Intelligence and Smart Communication (AISC) 2022 Jan 27 (pp. 680-684). IEEE.
24. Abbas MR, Asif KH. Computing prosody to detect the Arud meter in Punjabi Ghazal. Sādhanā. 2020 Dec;45:1-20.
25. Kurt A, Kara M. An algorithm for the detection and analysis of arud meter in Diwan poetry. Turkish Journal of Electrical Engineering and Computer Sciences. 2012;20(6):948-63.

26. www.aruzz.com access on 3rd August 2021
27. https://www.rekhta.org/ accessed on 3rd August 2021.
28. Jabbar A, ul Islam S, Hussain S, Akhunzada A, Ilahi M. A comparative review of Urdu stemmers: Approaches and challenges. Computer Science Review. 2019 Nov 1;34:100195.
29. Khan IU, Khan A, Khan W, Su'ud MM, Alam MM, Subhan F, Asghar MZ. A review of Urdu sentiment analysis with multilingual perspective: A case of Urdu and roman Urdu language. Computers. 2021 Dec 27;11(1):3.
30. Bashir MF, Javed AR, Arshad MU, Gadekallu TR, Shahzad W, Beg MO. Context-aware emotion detection from low-resource urdu language using deep neural network. ACM Transactions on Asian and Low-Resource Language Information Processing. 2022 May 8;22(5):1-30.