JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

PARADIGM UNIFORMITY VS PARADIGM ECONOMY: A COMPARATIVE STUDY OF NOMINAL AND VERBAL INFLECTIONAL PARADIGMS IN URDU AND ARABIC MORPHOLOGY

Raza-E- Mustafa[*1], Rida Zainab[2]

[*1]*Assistant Professor, Department of English, University of Gujrat*
[2]*Lecturer, Mukabbir University of Science & Technology, Gujrat*

[*1]razaemustafa@uog.edu.pk ,[2]ridazainab@mukabbiruniversity.edu.pk

DOI: http://doi.org/10.5281/zenodo.22108221

Keywords
Paradigm Uniformity, Paradigm Economy, Nominal Morphology, Verbal Morphology, Urdu Language, Arabic Language

Article History

Received: 28 May 2026
Accepted: 15 July 2026
Published: 31 July 2026

Copyright @Author

Corresponding Author: *
Raza-E- Mustafa

Abstract

The paper compares the inflectional paradigm structures of Urdu and Arabic, examining how each language handles the issues of paradigm uniformity and paradigm economy in nominal and verbal morphology. The study draws theoretical support from Word and Paradigm morphology (Matthews 1972; Blevins 2016) and Paradigm Function Morphology (Stump 2001; Stump and Finkel 2013). The analysis shows that Arabic has greater paradigm uniformity, maintaining formally distinct paradigm cells through cumulative exponence, a three-number system, and suffixal case morphology resulting in a higher amount of IC-complexity in its broken plural system. Urdu, on the other hand, achieves economy by distributing case across morphological and syntactic domains through the use of postpositions and auxiliary stems. Both languages have paradigm gaps, including the defective verb laysa in Arabic and the ergative agreement alternation in Urdu. This further highlights how both languages place structural constraints on paradigm cell realization. The findings suggest that uniformity and economy are genuine typological orientations of the languages, each stabilized by language-specific architectural choices that extend beyond the morphological system.

INTRODUCTION

Morphology deals with how languages are organized in terms of the form of words. One of the most important ways of understanding the forms of words is to see how the words are related with each other in certain grammatical contexts, a concept known as the paradigms. Analyzing the morphological structure of a language is not just knowing the terminal formatives of words, but also to know what distinctions are worth making, how the words are related to each other in a paradigm, and what happens when two cells that could be different turn out to be the same. These are some of the most important questions behind the item-and-paradigm approach proposed by Robins (1959), which remains relevant even today. Urdu and Arabic belong to two different language families: the former belongs to the Indo-Aryan branch of Indo-European languages, whereas the latter is a Semitic language. However, both

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

languages share centuries of contact due to shared literary and religious traditions. Urdu gained Arabic loanwords through Persian Masica (1991). Even as Urdu acquired Persio-Arabic vocabulary, its morphological structure remained distinctly Indo-Aryan. It resulted into both languages appear familiar to the speakers of each other, but differ significantly in their morphological structure.

Urdu and Arabic differ in their core morphological architecture. Arabic morphology utilizes the root-and-pattern approach where an underlying consonantal pattern undergoes different vowel patterns to produce noun, verbs, and other classes of words. Roots and patterns in Arabic operate at different phonological levels of representation and cannot be considered as sequentially concatenated morphemes (McCarthy, 1981). The resulting paradigms are complex. Even a single verb in perfective aspect generates as many as thirteen forms. Urdu follows a different path. It uses suffixes and postpositions to mark grammatical functions. Butt and King (2004) have given a detailed account of how case is both a morphological and syntactic property in Urdu nouns. Such complexities do not exist in nominal morphology in Arabic. The Urdu verb agrees in person and gender but the paradigm table for Urdu verbs is much simpler than Arabic verb with fewer cells.

Size, however, is only one dimension of the comparison. Paradigms also differ in how uniform they are, whether each grammatical distinction gets its own distinctive form, or the same form serves multiple grammatical functions. This phenomenon is called syncretism. According to Baerman et al. (2005), syncretism is not random. It generally follows implicational patterns by removing distinctions that are semantically and formally close. Urdu shows considerable syncretism in its nominal system: the feminine oblique form, for instance, is used in different case roles. Arabic has its own syncretism, especially in the dual and broken plural forms, where historical patterns collapse distinctions that the singular paradigm maintains. Ryding (2005) believes that the use of dual is in a decline in spoken varieties of Arabic, but has survived in formal written language.

This study examines these patterns systematically. It tries to analyze how and where the two systems are compatible in terms of paradigm economy i.e. having fewer forms, and paradigm uniformity i.e. having distinct cells for every grammatical function. Carstairs-McCarthy (1994) argues that paradigm economy is not an accidental property of individual languages but reflects a deeper principle determining how inflectional systems are organized. This study also tries to determine what phonological, syntactic, or frequency-based factors might explain the differences between Urdu and Arabic nominal and verbal morphology.

Two theories provide the basis for this investigation: Word-and-Paradigm (WP) morphology developed by Matthews (1972) and later expanded by Blevins (2016), and Paradigm Function Morphology (PFM) by Stump (2001). According to Word and Paradigm morphology, the word form is considered to be a primary morphological unit instead of a morpheme. This serves as an important reason for using the WP approach because in broken plurals in Arabic and case markings through postpositions in Urdu, segmentation into separate components cannot be done, hence the use of WP. Blevins (2016) believes that knowing a set of related word forms gives you more information about the other word forms than knowing the separate morphemes composing the words. Thus, Urdu and Arabic languages, due to inherent syncretism present in their systems, cannot be described morpheme-by-morpheme. Another theory that may be useful in our case is PFM, proposed by Stump (2001). In contrast to WP, PFM explicitly defines what combinations of person, number, gender, case, and tense define a paradigm cell and how these feature bundles are mapped onto actual words through realizational rules. Stump and Finkel (2013) later expanded this theory to include the concept of paradigm complexity. This development is especially important since we are going to compare the complexity of paradigms in Urdu and Arabic. Therefore, WP and PFM provide the theoretical

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

base for comparing geometries of paradigms in both languages having the same lexical base.

Implications of this study extend far beyond just the morphological systems of Urdu and Arabic. All languages experience the same issues in their morphological structure: strict grammatical distinctions are required for communicative clarity, but memory and production tend to favor economy. According to Haspelmath (2021), frequency and efficiency determine the structure of inflectional paradigms. Functionally light distinctions are dropped over time and important grammatical distinctions are retained. Languages differ greatly in how they tackle these issues, and this difference is an important aspect of human languages that morphological theory should account for. Urdu and Arabic use two different ways to address the same problem. Placing these languages alongside each other and conducting their morphological analysis through proper theoretical approaches will help us understand the morphological diversity presented by both languages.

1.1 Research Questions

Do Urdu and Arabic paradigms tend toward uniformity (many distinct cells) or economy (syncretism/defectiveness)?

How do root-and-pattern morphology (Arabic) and affixal morphology (Urdu) shape the internal geometry of paradigms?

Where do the two systems exhibit paradigm gaps (defective paradigms) and what conditions license them?

Literature Review

The study of inflectional paradigms has had a long tradition within the field of linguistic theory; however, the issues raised remain largely unresolved. For structuralists, paradigms were seen merely as a tool for presenting related word forms in an organized tabular form. However, the more seriously the linguists studied the paradigms, the more problematic they seemed. For example, Robins (1959) argued that a paradigm represents not just a convenient form of representation, but a true morphological unit whose internal structure cannot be analyzed through individual morphemes. This became the cornerstone of WP theory and served as a point of departure from the other morpheme-based theories.

2.1 The Word and Paradigm Tradition

It was Matthews (1972) who formulated the WP morphology theoretically. Based on Latin verb morphology, he demonstrated that the classical models of morphology based on morphemes, which are typical of American structuralism and which became particularly popular due to the Item-and-Arrangement approach of Hockett (1954), cannot cope with the cumulative and extended exponence that characterizes the inflectional system of most languages. The correct level of analysis, according to Matthews (1972), is the word itself which serves as the realization of one cell within a paradigm. Similarly, Anderson (1992) claimed that morphological operations do not target morphemes but words, and their result is always a complete word, never a concatenation of segments. This theory has important implications for the study of languages such as Arabic and Urdu, which are characterized by similar problems.

The latest account of the WP approach can be found in Blevins (2016). He argues that the structure of paradigms has an implicational nature and certain lexical items imply the existence of other items within a paradigm, and thus the focus of morphological analysis must be this implication. Blevins believes that morpheme-based approaches underestimate the amount of information contained in relations between whole words. This is especially true for Arabic. The famous broken plural paradigm, according to which a singular noun like *kitāb* becomes plural in form through the substitution of vowels, not through affixation of a morpheme, has been a problematic issue for concatenative morphology.

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

2.2 Paradigm Structure in Paradigm Function Morphology

While it was generally understood that paradigm structure was theoretically important, there remained an issue of formalizing paradigm structure. Stump (2001) tried to address that issue through his Paradigm Function Morphology (PFM), which offered a way to define explicitly which bundles of morphosyntactic features make up cells of the paradigm, and how specific ordered realizations map morphosyntactic feature bundles into actual word forms. PFM has been able to do what other WP accounts have traditionally failed to achieve due to their vague nature. PFM has found its uses not only for analyzing inflectional paradigms of European languages but also beyond Europe.

Building on Stump's work in PFM, Stump and Finkel (2013) have devised metrics of quantifying the complexity of paradigms, including IC-complexity, a measure of the unpredictability with which the whole paradigm of a lexeme can be derived on the basis of one of its members. What makes the use of such measurements relevant for a comparative study of Urdu and Arabic inflectional paradigms is the fact that they allow comparing inflectional systems of both languages on a level of formal analysis. A language can have a large paradigm that is highly regular and therefore easy to navigate, or a smaller paradigm filled with complex irregularity. Stump and Finkel's measures try to disentangle these complexities.

2.3 Root-and-Pattern Complexity in Arabic Morphology

Arabic morphology has been of particular interest to theoretical studies for various reasons. The most important among them is the fact that the root and pattern morphology of Arabic poses a challenge to theories of concatenation. McCarthy (1981), using his theory of autosegmental phonology, was able to prove that the root and pattern of Arabic are independent morphological levels and, therefore, Arabic is a non-concatenative language.

The study of the broken plural phenomenon in Arabic has yielded interesting results. According to Ratcliffe (1998), broken plural of Arabic should be viewed within the framework of the paradigmatic hierarchies rather than feature specification since there exists a systematic interaction of roots and patterns in this morphological phenomenon of Arabic morphology. Holes (2004) demonstrates that the classical paradigm involving three cases has been mostly maintained in Modern Standard Arabic, but has completely degenerated in most spoken dialects of Arabic, persisting only in some set phrases or in formal speech. These observations have been corroborated by Ryding (2005) who believes that the dual form is retained in written standard Arabic as a result of grammatical necessity rather than communicative needs of speakers.

2.4 Case, Agreement, and Postpositional Complexity in Urdu

Unlike Arabic, which has received sufficient theoretical attention, Urdu morphology has remained largely underexplored. Some of the more detailed grammatical accounts include Platts (1874) and Schmidt (1999), although neither would be considered primarily theoretical in nature. Theoretically oriented studies of the inflectional morphology of Urdu have mainly emerged from within the framework of LFG. In particular, Butt (1995) provided an analysis of the Urdu complex predicates that demonstrated an interaction between agreement in person and number based on the gender of the subject and the agreement between aspectual auxiliaries. Butt & King (2004) further applied these principles to case and showed that Urdu makes use of case markers in such a way that the morphosyntactic boundaries are blurred. While the nominative/oblique case is marked on the noun, the instrumental, dative, locative, and other cases are realized by postpositions that are affixed to the oblique form of the word.

Earlier, Mohanan (1994) had studied the structure of arguments in South Asian languages, such as

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

Urdu, and in her study, she highlighted the complex relations between ergativity of the language and its verbal agreement system, a phenomenon that does not have any parallel in Arabic. In Urdu, ergativity is exhibited by the verb in the perfective aspect, as the verb agrees with the object instead of the subject; thus, the paradigms for such verbs contain grammatical forms whose interpretations are clause-bound and not dependent only on morphological features of nouns. Such phenomena are precisely the sorts of issues that the WP analysis fails to address, and that is why the combination of WP and PFM is helpful, because PFM is capable of representing morphosyntactic properties of words established in syntax.

2.5 Comparative and Typological Perspectives

Comparative research addressing the similarities or differences in paradigm structure of Urdu and Arabic is rare. Studies related to inflectional paradigms in the crosslinguistic context have mainly considered inflectional morphological features of European languages, whereas the languages of other families, such as Indo-Aryan and Semitic, are mentioned only as marginal cases. An insightful typological analysis of inflectional morphology of Arabic and Urdu from the comparative point of view is offered by Haspelmath and Sims (2010). The three-way number distinction in Arabic and special types of verbal agreement triggered by broken plurals are listed among the more salient typological facts (Corbett, 2000; Corbett, 2006). However, a comparative theoretical account of inflectional paradigms in the two languages is yet to come.

As such, these findings form the basis upon which a comparative study between Urdu and Arabic paradigm structure is conducted. However, what it lacks is a comparative analysis that uses a common framework to compare the characteristics of the two paradigms by measuring their sizes, their syncretisms, defective cells, as well as the influence of economy and uniformity in both languages. This is what this study intends to undertake.

2.6 Theoretical Framework

Any comparison of inflectional paradigms requires the selection of a set of analytical tools. Framework selection cannot be neutral, as different approaches to morphology bring about different results. A morpheme-based framework emphasizes segmentation and composition, whereas, a paradigmatic approach emphasizes the relationships between the elements within the paradigm and the structure of the cells they occupy. Since this study aims to explore the organization of paradigm space in Urdu and Arabic languages, as well as to determine the amount of paradigm space utilized, the points at which paradigms collapse and where gaps occur, the most adequate analytical framework is the one that takes a paradigm as its main unit of analysis. There are two such frameworks: Word and Paradigm Morphology (WP), formulated by Matthews (1972) and further developed by Blevins (2016), and Paradigm Function Morphology (PFM), introduced by Stump (2001).

2.6.1 The Limits of Morpheme-Based Analysis

Before looking at what WP and PFM approaches offer, it is important to clarify why morpheme-based analyses are inadequate in this case. The major competitor to the WP approach is the Item and Arrangement (IA) theory introduced by Hockett (1954). IA holds that words consist of morphemes, which are concatenated during morphological analysis. Such a system is fairly productive for languages with simple agglutinative morphology in which each affix conveys one morpheme and affixes are attached in predictable manners. However, Arabic and Urdu belong to neither language family.

More specifically, the biggest difficulty can be observed in the case of the former. Consider a word such as *kataba* (he wrote). Here, perfective aspect, third person, masculinity, and singularity are all present in an intertwined form, where one morpheme cannot be separated from another in principle. McCarthy (1981) demonstrated that this situation is best analyzed within non-concatenative morphology, where the consonantal

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

skeleton of the root and its vocalic melody form two different morphological levels. However, no prefix or suffix are attached in the transition of *kitab* (book) to *kutub* (books); the transformation takes place internally within the word as a consequence of a conditioned vocalic alteration. The application of the IA model to this phenomenon thus involves postulating abstract morphemic entities without any phonological content or seeing it as a string of phonological transformations induced by morphological triggers – an unsatisfactory position in both cases.

A quite different challenge arises in the analysis of the grammar of Urdu language. Unlike Arabic, Urdu nominal morphology is suffixal and transparent: an oblique suffix is used in all nouns. However, unlike in Arabic where the case of a noun is expressed morphologically, in Urdu it is done using postpositions, which appear after the noun, is attach to the oblique form, and contain the semantic content of what would be a case suffix in Arabic. According to Butt and King (2004), case in Urdu is represented in both morphological and syntactic structure simulataneously. In such circumstances, an IA analysis limited only to the boundaries of the noun falls short. However, the paradigm should take into account the postpositional layer within the case system. The WP approach is able to do this better compared to the IA approach since it does not favor the morpheme over the word or the clause.

2.6.2 Word and Paradigm Morphology

A basic assumption of Word & Paradigm Morphology is that the word form and not the morpheme is the basic unit of morphological analysis. In his seminal work on Latin verbs in 1972, Matthews showed how the old-fashioned morpheme-based morphological theories were unable to accommodate phenomena like cumulative exponence, extended morphs, and the paradigmatic contrasts that characterized Latin verb inflections without any serious theoretical modification. In Matthews' theory, the word form embodies a set of grammatical properties like person, number, tense, and mood, which together make up a cell in the paradigm.

This idea had already been put forward by Robins (1959) in opposition to the then prevalent morpheme-based theories of mid-twentieth century American linguistics; however, only Matthews (1972) succeeded in providing the WP framework with enough theoretical substance to challenge the generativist morpheme-based theories. Anderson (1992), in his proposal for A-Morphous Morphology, extended this logic of WP even further by arguing that morphological rules map sets of properties onto sequences of sounds resulting in whole words rather than combinations of morphemes. This means that there is no room for morphemes in the lexicon and that all one has is words and morphological rules that connect these words to their paradigmatic neighbors.

Unlike previous WP approaches, Blevins (2016) treats implications drawn from the nature of paradigms seriously. His main point is that paradigms contain predictive relationships among cells. By knowing one forms of a particular lexeme, a speaker can derive other forms, which becomes the very basis of morphological knowledge. Blevins coins the term the Paradigm Cell Filling Problem, which, in his view, cannot be addressed by relying on morpheme inventories because it implies focusing on whole-word patterns. The application of Blevins' approach to the study at hand will have two important aspects. First, it enables the researcher to conceptualize the structure of paradigms in terms of predictability. Second, it provides a basis for accepting the Arabic broken plural as well as the Urdu postpositional case system as paradigmatically real despite their resistance to segmentation, because the WP criterion for paradigmatic reality rests in implication rather than morpheme identity.

In the present comparative analysis, WP morphology provides the framework. Cells of a paradigm are defined by morphosyntactic feature bundles which are instantiated by words. A paradigm for a lexeme is thus the complete pairing among word forms and feature bundles. It is this

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

structure of paradigm, its size, its regularities, its gaps, and its syncretic collapses that is compared here.

2.6.3 Paradigm Function Morphology

WP morphology is more theoretical than formal. It talks about how paradigms are built but does not give a computational account of how words can be derived from properties. In response to this problem, Stump (2001) provides a model in Paradigm Function Morphology (PFM). According to PFM, Paradigm Function maps from a lexeme and morphosyntactic property set defining a paradigm cell to a phonological form, i.e., a word form that realizes that cell.

The organization of rules in PFM allows for the encoding of the layered process in the derivation of verbal/nominal inflectional paradigms in Arabic. This can be illustrated by the case of Arabic perfect tense paradigm that consists of thirteen cells that can be generated by rules encoding the person/number/gender combination, with weak-root forms being dealt with by override rules that override the standard consonantal pattern.

When it comes to Urdu, the handling of verbal agreement is relatively straightforward, in terms of rules specifying person, number, and gender properties that can then be assigned the corresponding suffixes through the realization rules. More challenging, however, is the nominal domain because according to Butt and King (2004), the property set of the case-marked nominal must include not only the morphological case features of the suffixes but also postpositional specifications of the case. It is also possible to incorporate such an observation into PFM by including post-positional forms within the paradigm function. This, of course, would be a theoretical decision rather than something that comes out naturally, but it is defensible considering the data.

By further expanding the PFM, Stump and Finkel (2013) introduced quantification tools for analyzing paradigm complexity. Among the concepts introduced by them, the measure of IC-complexity (which assesses the amount of uncertainty about the remaining paradigm after knowing just one cell in a paradigm) is especially pertinent to the topic at hand. An IC-complex language is a language whose paradigms are characterized by mutually unpredictable cells, while the internal redundancy of paradigm cells is characteristic of a language with lower IC-complexity. It can safely be assumed that a language like Arabic with its broken plurals and weak root alternations should display greater IC-complexity than Urdu with its regular suffixes. Yet, this too is something that the current analysis would aim to prove.

2.6.4 Integrating the Two Frameworks

The compatibility of WP and PFM lies in their shared assumptions: first, word-form identity is they key; second, the paradigm is the basic structure; third, morphological competence is competence in relationships between paradigm cells. The point of divergence comes in the degree of formalization of the two frameworks: WP, especially Blevins' implementation of it, is best suited for the description of implicational structures within paradigms and proving the presence of information within whole-word relationships that morpheme-based frameworks fail to account for. PFM is more effective in calculating actual forms based on feature specification and quantifying the complexity of the system.

While WP is applied here as a conceptual framework for interpretation, it helps understand the essence of paradigms and explains the theoretical relevance of paradigm comparison across languages. PFM is employed as an analytical tool which helps calculate the forms, identify syncretisms via the merging of rules of realization and paradigm gaps in the absence of relevant rules, and quantify the complexity of the realization system.

The combined application of both approaches makes it possible to analyze the essential questions of the present research in a scientific manner. Is the Arabic verbal paradigm larger, in terms of cells,

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

compared to the Urdu paradigm? In which situations do similarities exist, and does such syncretism fit into the patterns observed in all languages according to the Baerman et al. (2005) analysis? Is the Urdu verb smaller because this language shifts morphological complexity towards syntactical complexity, dispersing its categories not only within the word but also using postpositional forms? And, finally, what can the indicators of IC-complexity tell us about the ease of learning?

Data Analysis

In this section, the theoretical approach developed earlier is applied to paradigm data in two natural languages – Urdu and Arabic. There are four comparative dimensions under examination, namely paradigm size and cell density, syncretism, paradigmatic gaps and defectiveness, and complexity of IC as an indicator of paradigm predictability. Each of these dimensions will be explained using the data from each language in comparison to each other.

3.1 Paradigm Size and Cell Density

The first and foremost architectural distinction between Urdu and Arabic inflectional paradigms lies in the sheer number of cells within these paradigms. The Arabic paradigms, especially the verbal paradigms, are considerably larger than those seen in Urdu. It does not mean that Arabic is a more complex language because it contains a larger paradigm.

3.1.1 Arabic Verbal Paradigm: Perfect Tense

The Arabic perfect tense verb *kataba* (to write) generates thirteen distinct conjugated forms in the singular and plural alone, not counting the dual. Table 1 below displays the full paradigm.

Table 1: Paradigm of Arabic verb *kataba*

Person	Gender	Number	Form	Gloss
3rd	Masculine	Singular	*kataba*	he wrote
3rd	Feminine	Singular	*katabat*	she wrote
2nd	Masculine	Singular	*katabta*	you (m.) wrote
2nd	Feminine	Singular	*katabti*	you (f.) wrote
1st	–	Singular	*katabtu*	I wrote
3rd	Masculine	Dual	*katabā*	they (m. dual) wrote
3rd	Feminine	Dual	*katabatā*	they (f. dual) wrote
2nd	–	Dual	*katabtumā*	you (dual) wrote
3rd	Masculine	Plural	*katabū*	they (m.) wrote
3rd	Feminine	Plural	*katabna*	they (f.) wrote
2nd	Masculine	Plural	*katabtum*	you (m. pl.) wrote
2nd	Feminine	Plural	*katabtunna*	you (f. pl.) wrote

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

1st	–	Plural	*katabnā*	we wrote

There are thirteen forms for a single verb in a single tense, each being phonologically distinct. However, the distinction of gender in this paradigm is made in second and third persons for all three numbers but not in the first person where the difference of gender is neutralized. Two additional cells are added by dual. In PFM framework, realization rules need to refer to person, gender, and number all at once, since none of them is enough on its own for determining the form.

3.1.2 Urdu Verbal Paradigm: Perfective Aspect

The perfective system for *likhna* (to write) works according to a different principle. It consists of the verb stem with a participle agreeing with the subject in gender and number and a perfective particle expressing the tense. This is the paradigm of the simple past from table 2:

Table 2: Paradigm of Urdu verb *likhna*

Gender	Number	Verb Form	Auxiliary	Full Form	Gloss
Masculine	Singular	*likhā*	*hai*	*likhā hai*	he has written
Feminine	Singular	*likhī*	*hai*	*likhī hai*	she has written
Masculine	Plural	*likhe*	*haiN*	*likhe haiN*	they (m.) have written
Feminine	Plural	*likhī*	*haiN*	*likhī haiN*	they (f.) have written

Urdu has four forms for the present perfect as compared to thirteen for the Arabic perfect. There is no person marking on the participle suffix in Urdu. Unlike Arabic, the person marking is done via auxiliary verbs and is less richly morphological than Arabic's suffixation system. Gender is marked in the singular form but not in contrast between feminine singular and feminine plural, representing a syncretism that will be discussed in detail in the next section. What is most important is that Urdu does not mark the dual. As far as both noun and verb paradigms go, this feature is simply not a part of Urdu grammar.

From a PFM perspective, the Urdu paradigms are simpler because the suffix block needs only gender and number, but not person. However, the additional grammatical meaning carried out by Arabic within its verbal suffix is found in the auxiliary in Urdu. Although the Urdu paradigm is smaller, yet it does not mean absence of grammatical functions. The functions are only relocated syntactically.

3.2 Syncretism Patterns

Both languages have syncretism, the use of one form in different grammatical cells; however, in both languages, syncretism is based on different structural principles. In the WP model, syncretism provides information about the internal structure of a paradigm. If two adjacent cells share the same form, it is assumed that those cells must have some sort of grammatical closeness to each other.

3.2.1 Syncretism in Urdu Nominal Paradigms

The Urdu nominal morphology involves two cases – direct and oblique. All other case relations such as dative, instrumental, locative, genitive, and ablative are expressed by means of postpositions applied to the oblique case. Thus, one oblique form may be used as the base for several grammatical functions. This is a classic example of

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

syncretism in Urdu. Consider the noun *laRkī* (girl):

Table 3: Case form of Urdu noun *laRkī*

Case	Form	Construction	Gloss
Direct Singular	*laRkī*	bare noun	the girl (subject)
Oblique Singular	*laRkī*	bare oblique	(base for postpositions)
Dative	*laRkī ko*	oblique + *ko*	to the girl
Instrumental	*laRkī se*	oblique + *se*	by/from the girl
Locative	*laRkī meN*	oblique + *meN*	in/at the girl
Genitive	*laRkī kā/kī/ke*	oblique + *kā*	of the girl

The direct and oblique forms of the *laRkī* are phonologically similar in the singular form without any difference in suffixes. This is an instance of syncretism between direct and oblique cases for feminine nouns in the singular form, rather than an accidental similarity in the pronunciation of the case form. According to Schmidt (1999), such a case occurs with all feminine nouns ending with *–ī*.

Contrast the masculine noun *laRkā* (boy):

Table 4: Case Forms of Urdu noun *laRkā*

Case	Number	Form	Gloss
Direct	Singular	*laRkā*	the boy
Oblique	Singular	*laRke*	(oblique base)
Direct	Plural	*laRke*	the boys
Oblique	Plural	*laRkoN*	(oblique base)

In this case, a different type of syncretism arises: the form of the oblique singular masculine *laRke* is phonologically identical to that of the direct plural masculine *laRke*. A listener cannot decide whether *laRke* is an oblique singular or a direct plural based on only the noun itself. The form can be identified on the basis of whether there is a postposition or not, or from the subject-verb agreement. Baerman et al. (2005) calls this accidental syncretism, because the two forms oblique singular and direct plural are not adjacent in terms of thier grammatical features. Yet this syncretism keeps occurring all through the class of masculine *-ā* nouns.

3.2.2 Syncretism in Arabic Nominal Paradigms

Arabic nominal syncretism operates in a different way. Whereas the traditional system of three grammatical cases i.e. nominative, accusative, and genitive, continues to be maintained in modern standard Arabic, but shows systematic collapse in some morphological contexts. Broken plurals is one of the most studied cases of such syncretism.

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

In many forms, the accusative and genitive cases merge, losing their singular distinctions:

Consider *kitāb* (book) and its broken plural *kutub*:

Table 5: Case forms of Arabic noun *kitāb*

Case	Singular	Plural (broken)
Nominative	*kitābun*	*kutubun*
Accusative	*kitāban*	*kutuban*
Genitive	*kitābin*	*kutubin*

All three categories are formally distinct in the singular form via the use of nunation suffixes *-un*, *-an*, and *-in*. In the broken plural, the same three distinctions are also formally marked via identical nunation. However, in the case of a diptote category, there are several patterns within the broken plurals and loanwords:

Table 6: Case forms of Arabic noun *masājidu*

Case	Diptote form	Gloss
Nominative	*masājidu*	mosques
Accusative	*masājida*	mosques
Genitive	*masājida*	mosques

In the above examples, the genitive and accusative cases are syncretic because they share the *-a* ending while nominative retains the *-u*. Ryding (2005) has described this phenomenon extensively, noting that it occurs in a large number of broken plural forms. In PFM terminology, this means that the realization rule for accusative in diptotes is identical to that of genitive and can therefore be expressed by one underspecified rule applied to both property sets.

In Urdu, syncretism leads to the reduction of the direct/oblique distinction for particular noun categories and thus reduces morphology towards the unification of the cases. In Arabic, on the contrary, syncretism in the case system leads to the reduction of the accusative/genitive distinction while retaining nominative. This is consistent with findings of Baerman et al. (2005) who believe that nominative case is most resistant to syncretism.

3.3 Paradigm Gaps and Defective Paradigms

The phenomenon of paradigm gap occurs when a word form is absent from a cell where one would expect it according to the paradigm rules of the language. Both Urdu and Arabic are languages with defective paradigms, but the triggering conditions for paradigm defectiveness in these two languages differ.

3.3.1 Defectiveness in Arabic

There are some well-known instances of paradigm defectiveness in the Arabic verbal system. One of these examples concerns the negative copula *laysa* (not to be). This verb does not have an imperfect form. The perfective form, however, is completely inflected for person, gender, and number. Therefore, the paradigmatic cells for the imperfect form of this verb are empty, while the cells for its perfect paradigm are filled with the forms of *yakūnu* (the imperfect of *kāna*). In PFM, the

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

paradigm function for *laysa* is defined only for the perfect property sets; the imperfect cells offer no output.

A more subtle form of defectiveness can be found in the representation of the dual form in spoken Arabic dialects. According to Holes (2004), the dual form that should be used in the standard language with respect to nouns, adjectives, and verbs is not represented in colloquial Arabic dialects. The speakers of Egyptian Arabic, for example, always use plural forms when the standard language asks for the dual agreement. The empty cells exist in the standard language but are simply not realized in the spoken language.

3.3.2 Defectiveness in Urdu

Urdu shows a different type of defectiveness, especially in terms of ergativity in the verbal paradigm. When transitive verbs occur in perfective aspect, they show ergativity and agree with the object and not the subject; meanwhile, the subject is marked by the ergative postposition *ne*. As a result, certain cells are not available in particular syntactic environments. The verb *likhā* in the sentence *us ne xat likhā* (he wrote a letter) agrees with the masculine singular object *xat* rather than with the masculine singular subject *us*.

This results in a paradigm in which the cells corresponding to subject agreement are not filled in the ergative marking. They are blocked by the syntactic constraints which force agreement with the object. This is not the case of lexical gap, when a certain lexical item lacks forms because the verb does not have such forms at its disposal. This is a paradigmatic gap that arises because of the syntactic context, and, as such, can be described with respect to agreement mismatch. The PFM account of such gap would require the paradigm function to encode not only lexeme and properties of the construction but also the syntactic environment as well.

3.4 IC-Complexity and Paradigm Predictability

Based on the discussion of Stump and Finkel (2013), the degree of complexity of inflectional paradigms in Urdu and Arabic can be measured by how uncertain one would be of the entire inflectional paradigm of a word when presented with just one of its forms. High predictability implies low IC-complexity while unpredictability implies high IC-complexity.

3.4.1 Arabic Nominal IC-Complexity

The Arabic nominal system carries certain unpredictable elements. The most unpredictable one is the formation of the broken plural. It is difficult, sometimes even impossible, to predict the type of broken plural by looking at the singular form. Consider for example the following pairings:

Table 7: Arabic broken plurals

Singular	Broken Plural	Pattern
kitāb (book)	*kutub*	CuCuC
bayt (house)	*buyūt*	CuCūC
walad (boy)	*awlād*	aCCāC
madīna (city)	*mudun*	CuCuC
rajul (man)	*rijāl*	CiCāC

Of the five different singular forms, there are five different patterns of broken plurals. The consonant root remains unchanged but the vocalic pattern changes according to the template, but

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

which template can be applied to a particular root cannot be determined from the singular form. Knowing the singular form *kitāb* does not enable you to predict the plural form kutub; they need to be memorized independently of each other. Such is precisely the situation in which the IC-complexity becomes very high.

3.4.2 Urdu Nominal IC-Complexity

IC-complexity in the nominal paradigms of Urdu is relatively less. Given the direct singular form of a noun, it is easier to predict all other forms i.e. the oblique singular, the direct plural, and the oblique plural based on the gender and terminal phoneme of the noun. The pattern of morphological ending of *-ā* is regular.

Table 8: Case forms of the noun *laRkā*

Cell	Form	Rule
Direct Singular	*laRkā*	stem
Oblique Singular	*laRke*	stem + *-e*
Direct Plural	*laRke*	stem + *-e*
Oblique Plural	*laRkoN*	stem + *-oN*

With knowledge of the word *laRkā*, one can the whole paradigm following a limited number of rules. In this respect, the syncretism between the oblique singular and the direct plural forms works in favor of linguistic economy. Rather than learning four forms, the learner would have to learn only three forms of each noun. For the feminine nouns in *-ī*, the paradigm remains predictable, with direct-oblique syncretism in the singular reducing the number of forms even further. Thus, Urdu nominal morphology manages to do the grammatical job with a relatively small number of generally applicable and exceptionless rules.

The comparison with Arabic is obvious and theoretically instructive. The paradigms in Arabic depend heavily on lexeme-specific knowledge, with each noun requiring knowledge of the appropriate pattern for its broken plural. Urdu paradigms, being more predictable, are much easier to learn because of their lower IC-complexity.

Conclusion

The study aimed to compare the paradigm structures of Urdu and Arabic along two axes: the degree of paradigm uniformity, and paradigm economy by means of syncretism and grammatical information redistribution. The analysis shows that both these dimensions reflect systematic architectural differences between the two languages and are theoretically important.

The results suggest that Arabic has greater paradigm uniformity in terms of maintaining formally distinct cells in its paradigms and preserving gender and number distinctions within the verbs. Moreover, the nouns in Arabic preserve a three-way distinction within the category of case, number and gender. Its nominal and verbal systems encode greater information inside the inflectional categories instead of syntactically redistributing the grammatical information across the clause. A higher paradigm uniformity results in higher IC-complexity, especially in the nominal morphology, which features the plural forms that cannot be derived directly from the corresponding singular. It requires a more complex realization system, more lexeme-based learning, and more complex mappings between the properties that make up morphosyntactic structures and phonological realizations.

JOURNAL FOR SOCIAL SCIENCE STUDIES
An International Peer-Reviewed Open Access Journal
ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

Urdu, on the other hand, is a more economical language. Its nominal paradigms are smaller; verbal forms have less agreement and uses redistribution of the grammatical functions. Arabic, for instance, marks case with nominal suffixes, while Urdu uses postpositions for that purpose. Similarly, Arabic encodes tense and aspect in a single synthetic verb form, whereas Urdu distributes these features between participle stem and auxiliary. This results in a system which has a reduced IC-complexity but relies more on the syntactic context for complete understanding of the grammatical information.

The difference between Urdu and Arabic is not merely that of morphological complexity between a synthetic and an analytic language; it is the different ways in which they manage the competing pressures that every inflectional morphological system faces. On the one hand, there is the pressure for clear formal distinctions: morphologically rich languages include clear grammatical distinctions in their paradigms without having to recover information through contextual inference by the hearer. On the other hand, there is the pressure for economy with fewer forms to produce and retrieve. In Arabic morphology, former pressure seems to outweigh the latter. Urdu, on the other hand, prefers economy over formal distinctions. Moreover, each linguistic system uses a means of making this choice consistent. For example, in Arabic, this is achieved through templatic morphology using roots and patterns. However, Urdu has postpositions and auxiliaries, which enable the realization of the distinctions in a compositional way, reducing the need to have a large number of inflectional forms without reducing the distinctions themselves.

The implicational version of Word and Paradigm morphology by Blevins (2016), and Paradigm Function Morphology by Stump and Finkel (2013) were ideally suited to reveal such intricate differences between the morphological architecture of both languages. The morpheme-based approach would not have worked well for Arabic broken plurals, Urdu postpositional case system, or for the ergative agreement of transitive perfective verbs in Urdu. By treating word forms as the primary units of analysis and analyzing paradigmatic structures in terms of their implicational hierarchies freed the researcher from imposing an artificial structure on the data that would not fit. Moreover, the use of PFM helped to describe realization rules, detect syncretism, and calculate IC-complexity, thus making the cross-linguistic comparison feasible. Neither framework alone would have been sufficient; together, however, they have made this comparison viable.

A few specifics are worth carrying over into future research. Firstly, there are different syncretism tendencies in both languages. While Arabic tends to retain the nominative but lose the distinction between the accusative and genitive in diptotes, this conforms with the syncretism hierarchy described by Baerman et al. (2005). On the contrary, Urdu tends to lose the distinction between direct and oblique cases in some noun classes and retain it in others, which can be explained by lower functional load of case morphology in a language which uses postpositions as means of expressing case relations. Secondly, whether these syncretic tendencies are universal or developed historically as a result of language use. This can be addressed through a large-scale typological study involving other Semitic and Indo-Aryan languages.

The second interesting point is related to what can be called the compensation principle. The reduced paradigms due to economy should not lose the grammar information contained in them. Instead, the information can be found in another part of the sentence or clause. Such compensation occurs in Urdu due to its postpositional system and its auxiliary chain. In effect, paradigm economy appears to be a characteristic not of the morphological system in isolation, but of the morphological system in relation to the syntax and pragmatics. A language can economize on its paradigm forms if it has other ways of encoding the distinctions left unencoded by those forms.

One limitation of the present study is that the data used here is drawn from Modern Standard Arabic.

ISSN (Online): 3006-2071 | ISSN (Print): 3006-2063
Volume 4 • Issue 3 • July 2026

The data from spoken dialects of Arabic, especially the prominent colloquial ones, featuring distinct paradigm organization, a heavily reduced dual system was neither available, nor used. The analysis of the two languages in terms of paradigm uniformity and economy through the addition of informal Arabic to the equation along with formal and colloquial Urdu and formal Arabic would have provided more insight into the matter. However, this will be considered for future research. In this way, the present research has provided a point of reference from which further comparisons and analyses may be carried out.

REFERENCES

Anderson, S. R. (1992). *A-Morphous Morphology*. Cambridge University Press.

Baerman, M., Brown, D., & Corbett, G. G. (2005). *The Syntax-Morphology Interface: A Study of Syncretism*. Cambridge University Press.

Blevins, J. P. (2016). *Word and Paradigm Morphology*. Oxford University Press.

Butt, M. (1995). *The Structure of Complex Predicates in Urdu*. CSLI Publications.

Butt, M., & King, T. H. (2004). The status of case. In V. Dayal & A. Mahajan (Eds.), *Clause Structure in South Asian Languages* (pp. 153–198). Springer.

Carstairs-McCarthy, A. (1994). Inflection classes, gender, and the principle of contrast. *Language*, 70(4), 737–788.

Corbett, G. G. (2000). *Number*. Cambridge University Press.

Corbett, G. G. (2006). *Agreement*. Cambridge University Press.

Haspelmath, M. (2021). Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. *Journal of Linguistics*, 57(3), 605–633.

Haspelmath, M., & Sims, A. D. (2010). *Understanding Morphology* (2nd ed.). Hodder Education.

Hockett, C. F. (1954). Two models of grammatical description. *Word*, 10(2–3), 210–234.

Holes, C. (2004). *Modern Arabic: Structures, Functions, and Varieties*. Georgetown University Press.

Masica, C. P. (1991). *The Indo-Aryan Languages*. Cambridge University Press.

Matthews, P. H. (1972). *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge University Press.

McCarthy, J. J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry*, 12(3), 373–418.

Mohanan, T. (1994). *Argument Structure in Hindi*. CSLI Publications.

Platts, J. T. (1874). *A Grammar of the Hindústání or Urdú Language*. W. H. Allen.

Ratcliffe, R. R. (1998). *The "Broken" Plural Problem in Arabic and Comparative Semitic*. John Benjamins.

Robins, R. H. (1959). In defence of WP. *Transactions of the Philological Society*, 58(1), 116–144.

Ryding, K. C. (2005). *A Reference Grammar of Modern Standard Arabic*. Cambridge University Press.

Schmidt, R. L. (1999). *Urdu: An Essential Grammar*. Routledge.

Stump, G. T. (2001). *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge University Press.

Stump, G. T., & Finkel, R. A. (2013). *Morphological Typology: From Word to Paradigm*. Cambridge University Press.