THE CRITERIA FOR IDENTIFYING URDU COMPOUNDS: A MORPHOLOGICAL PERSPECTIVE

1.   **Abrar Hussain Qureshi**
2.   **Mehmood Ul Hassan**

ABSTRACT:
Compounds constitute a core part in any lexicon of language. Until the mid of 20th century, the compounds were a neglected area in language teaching. But, in the last few decades, these word combinations have attracted the linguists and language teachers across the globe. Their significance has increased manifold after the fact that meaning is present in the longer stretches of utterances rather than in individual words. Their complexity of construction has also been one of the reasons that they have been neglected across the linguistic resources, especially by teachers and lexicographers. The undertaken study focuses on the criteria of identifying Urdu compounds, applying the framework of Plag (2002). It tries to establish the criteria to highlight the Urdu compounds in comparison with phrases, conversion, and derivation. The defining of criteria for Urdu compounds is supposed to be useful especially for language teachers and generally for Urdu bilingual lexicographers.
Key Words: Compounds, Criteria, Identification, Urdu, Pedagogy

1. Introduction

Human beings are creative users of language who borrow words from other languages, rede ne meaning of existing words, create new words and re-form words into new words or into new meanings. Naming newly created or invented products, establishment, ideas and many others, often triggers new word formation. Furthermore, the activities, characteristics or nature of those newly-named produce other new derivatives. For example, one of the most popular social media Twitter (a product) even provides The Twitter Glossary as the words and expressions used in this media may have the same spelling, word class or other grammatical categories with the

The creativity of human language is an established fact. Native speakers of a language keep on coining words from their indigenous language to create new word combinations. This process of coining new combinations has become all the more frequent after the world has shaped into a global village and the need for interactions among nations and group has increased many times.

---

*University of Sahiwal, Pakistan
**Khwaja Fareed UEIT, Rahim yar Khan

There is also a shift of linguistic analysis to word combinations rather individual lexical items. Moreover, a single word for an object in a language may be a compound in the other language. Compounds constitute one of the significant levels of phraseology of any language.

Generally, they are formed by gluing two existing words and as a result, a new compound comes into existence. According to Hatch and Brown, compounds deserve to be written as one word rather as one entry in dictionaries. Sometimes, they are written as two or more orthographic words (Schmitt, 1997). Compounding is a common technique in languages to coin new words and one of the frequently used strategies of word formation. This process of word formation is very rich. They are used in different fashions for example, as multiple words in orthography, sometimes with a hyphen or sometimes without a hyphen as a single word. (Schmitt, 2000). Another feature of compounds is that they consume less number of morphemes as compared with phrases and idioms. Their simplicity is due to the fact that their word order is fixed and unique. On account of their uniqueness, they are the easy choice of the native speakers. According to Bauer (1983) derivation may apply with one or more roots, while two roots are obligatory in compounding.

Compounding is a rich phenomenon in Urdu language and occupies an important place in colloquial speech. The following examples contain one compound verb: بیٹھ جاؤ *beth jao* (sit down), رکھ لینا *rakh lena* (to keep). While the following sentence contains two verbs, not one compound: وہ مجھے دو روپیہ انعام دے گیا *voh mujhe do ropea enaam de geya* (he gave me a reward of two rupees as he went). A distinction can be established between a compound verb and a simple verb. A simple verb only unfolds an action or an event while a compound verb provides contextual information as well. The comparison may be highlighted to the difference between a shaded drawing and an only the outline of drawing (Schmidt 2000).

2. Orthographic criteria

Orthographically, compounds can be written in three different ways. They can be written as one word, or with hyphens and as separate words. For instance in Urdu the compound word سفید خون *sufaid khoon* (white blood), is used. This compound word has two elements, one of which is a modifier and the other is the head. سفید is the modifier and خون is the head.

Another orthographic criterion of compounding is that the parts of the compound words can be written with a hyphen between them. Compounds that are written as one word or hyphenated present no major problem of identification except that they need to be kept separate from derivation. Urdu has no compounds in which hyphen is used while in English, hyphen is frequently used. For instance, the following compounds "*merry-go-round", well-respected and "well-being,*" all have a hyphen in between them.

2.2. Stress

The criterion of stress is used in identifying compound words when they are not written together or hyphenated. It is also used as the main criterion for distinguishing compound words from phrases. According to Katamba (2005:68), phonological compounds and

especially noun compounds have the main stress on the first element and a secondary stress on the second element, whereas in phrases the reverse is true. This criterion has been used for identifying compounds in both the varieties.

For instance, آرمی چیف *Army chief* . In this example the primary stress is on the first element and the second element has the secondary stress, which is why they are included in the list of compound words according to the phonological criteria of the words.

2.3 Difference between Compound words and Phrases

Katamba (2000) claims that "in English, compounds may be distinguished from phrases on phonological, syntactic and semantic grounds". As in English, words are also recognized by a single primary stress in Urdu so, compounds are often recognized by the stress pattern and lack of juncture. For example the phrase کالا رنگ' black color) has stress on each word and a juncture between the two words, whereas مزیدار) delicious) has one primary stress and no juncture. Word order also distinguishes compounds from phrases, as some compounds use what may be considered ungrammatical and unusual word order in Urdu. For instance, the structure of Adjective + Noun is not a usual phrase pattern in Urdu but is used in compounds such as لمبا لڑ کا (tall boy). Compounds are also characterized semantically as they acquire specialized meanings like idioms. In very few cases the meaning of the compound is derived from its constituents in the literal sense. In most cases the meaning of one constituent is obscure. For example کالا رنگ black color may be green or blue.

2.4 Compounding and Derivation

There is a further distinction between derivation and compounding. According to Adams (2001), compounds are the combinations of two free morphemes as they can stand alone independently, for instance پتلئ لڑ کی (slim girl), موٹا بچه) fat child). The elements of compounds have the identifying characteristics of single words as their order is fixed and their elements cannot be separated by other forms. In a derived word the affix is a bound form and characteristically gives the word a more literal meaning than one would expect a grammatical element to have. Thus compounding is different from derivation in respect of the absence and presence of bound forms and whether both roots are lexical or not. According to Katamba (2000), the word class of the item can be changed by derivational affixes and in this way new words are formed that belong to the various word classes. "They are inner with respect to inflection, so that if derivations and inflections co-occur, derivations are inner, closer to the stem, and inflections are outer, furthest from the stem". For instance in Urdu کھیل-و کھلاڑ ی کھیلنا (to play, player, play (imperative), the base form is کھیلنا , and when و is added it will become an inflection and by adding ئ it will become derivation.

2.5 Compounding and conversion

One problem arising while identifying compound words is conversion. In conversion a word belonging to one class is transferred to another word class without any change of form. There are many words in these articles that were handled as cases of conversion. For instance, in Urdu the compound word دیکھا دیکھی (in imitation) and دلچسپ (interesting) are used. In these words, although دیکھا (to see) and چسپ (to stick) are originally verbs, they are used as nouns after conversion. Conversions from noun to verb and from verb to noun are the most productive word formation processes in Urdu . According to Adams (2001), the productivity of conversion can be measured only in terms of change in word class. This type of word formation is highly productive especially when it involves the open word classes. In this regard most nouns may be

used as verbs and many verbs can be used as nouns; adjectives can also be turned into nouns and verbs.

3. The Notion of Head

The notion of headedness is a significant issue in the discussion of compounds, especially in the work of generative morphologists such as Katamba (2005). The notion of head of a phrase in syntax has played a crucial role in generative grammar, and the notion of head is also used in morphology. For example, since most Urdu compounds are endocentric, they have heads. Platts (2002) argues that in English compounds the right-hand member is the head because it determines the category of the complex word. In Urdu morphology, we define the head of a morphologically complex word to be the left hand member of that word that may be called the left hand Head Rule.

The lexical class membership of the English and Urdu compounds   is correctly explained by the RHR and LHR.

 English compounds (Katamba 1993: 311)
[bird N watch N] N [over P react V] V [sugar N daddy N] N
[blue A book N] N [blue A black A] A [wind N screen N] N
Urdu compounds LHR
[نیلا آسمان] N   (blue sky)   [لمبا درخت] **N**   (tall tree)
[دو نالی]   A   (double barreled)   [خوبصورت] A   (beautiful)
[اچھا ہونا] V   (to get well)   [سو جانا]   V   (to go to sleep)

So far, we have investigated the notion of headedness from the viewpoint of the determination of the lexical category of the entire compound. However, there are also other kinds of arguments on which to base headedness in compounds, or, alternatively, the individual members of a compound may be regarded as 'heads' in other senses than the lexico-morphological one that we have discussed so far. Thus, Adams (2001) gives three different kinds of arguments to base headedness on.

Similarly, to phrases, words also have heads. In the case of compounds, syntactically the head is the dominant constituent of the construction, which means that the inflectional properties of the compound are inherited from the head element. Semantically, the head of a compound specifies the class of entities to which the compound belongs (Katamba 1993). It is a general assumption that the majority of Urdu compounds follow the Left-Hand Head rule. There is also idiosyncratic nature of endocentric Urdu compounds. The vast majority of Urdu compounds are interpreted in such a way that the right-hand member somehow modifies the left-hand member. Thus, پتلی لڑکی (slim girl) is a kind of girl), موٹا بچہ (fat child) is a type of child and so on.

Katamba (1993) disagrees with Selkirk (1982) when he tries to suggest new rules for semantic analysis for exocentric compounds. Rather, he suggests simple indexing of meaning of these compounds. Katamba asserts that exocentric compounds and idiomatic expressions are "listemes" semantically, although, they are complex sometimes and not subject to compositionality. The researcher strongly believes that it is due to the fact exocentric Urdu compounds have high frequency than the endocentric compounds that have comparatively less frequency in Urdu language. However, the semantic dimension of endocentric compounds is

complex then there is as such no compulsion in using them. In this regard, it is more reasonable if Katamba's theory of compounding is followed for such type of analyses as is required for semanticlly endocentric compounds.

It can be safely concluded that the inner structure of an Urdu compound originally highlights a head and a modifier. The term "head" denotes generally to the core unit of a complex linguistic structure. In all the above Urdu compounds, it the central unit of a compound, namely "Head" that is modified by the other part of the compound.

 With reference to structure of the head, Urdu compounds generally have a very organized and systematic property. Heads in Urdu compounds most generally occur on the left side (commonly known as left-hand rule for compounds). Urdu compounds inherit their syntactic and semantic information from their heads. Thus if the head of an Urdu compound is verb, the consequent compound will be a verb e.g. سو جا نا (to go to sleep)**,** And if the head of an Urdu compound is a countable noun, the consequent compound will be a countable noun (e.g. مو ٹا بچہ *(fat child)*. Similarly, if the head of an Urdu compound has a feminine gender, the consequent Urdu compound will have a feminine gender (e.g. پتلئ لڑ کی (slim girl). The head of an Urdu compound has another property that if there is a need to pluralize the head of an Urdu compound then plural marking will affect the head and not on the rest of the parts of the Urdu compound. Thus, لڑ ک پتلیا ں is not the plural of پتلئ لڑ کی only پتلئ لڑ کیا ں can be the plural form.

3.1 Compositionality

Compositionality is a language phenomenon where the meaning of an expression can be expressed in terms of the meaning of its constituents e.g. خلا ئی د و ر *khalaaee daur* (space age),خلا ئی جہا ز *khalaaee jahaaz* (spacecraft), خلا ئی تجر بہ گا ہ *khalaee tajerbah gah* (space station), some are non-compositional e.g سفید خون *sufaid khoon* (insincere attitude), and some in between e.g. فلک بو س *faluk boce* (extremely high) . Compositionality is lost in idioms. The meaning of آ نکھیں د کھا نا *ankhain dekhana* (to threat)   has nothing to do with either آ نکھیں (eyes) or د کھا نا (show) . Idiomaticity is found among words as well as phrases. There is a gliding scale from fully compositional (or motivated) to fully idiomatic (or conventional). Somewhat easier to parse example of nominal compound is: جنگ نا مہ *jang nama* (the book of battle).

Summing up, the principle of compositionality states that semantic interpretation of words, to some extent, reflects their derivational history and internal make-up. It is useful as a guiding principle for semantic analysis (of words and phrases), in the sense that a compositional analysis is usually to be preferred over a non-compositional one.

4. Concluding Remarks

 Multi-word items are the result of lexical (and semantic) processes of fossilisation and word-formation" Multi-word items in Urdu cover a wider range. They may be pair or group of words associated paradigmatically or a pair or group of words associated syntagmatically, to form an idiom. Best examples of significant multi-word items in Urdu are compounds.

Compounds cover a core part in the vocabulary of the Urdu language. Compounding means two or more free forms are combined to form a single lexical item that has a particular meaning. From a semantic point of view, the new lexical items formed in this way are of two types: (a) those possessing the total some of meaning of their constituents (i.e. the free forms) e.g. خلا ئی د ر*khalaaee daur* (space age) (b) possessing a meaning that is distinct from the meaning conveyed by the constituents e.g. شہر یا ر *shehriyar* (king) . Summing up the arguments, there has been made a deliberate attempt to determine the criteria for the definition and description of

Urdu compounds. As Urdu is a very complex language that developed under the influence of many regional languages, it has grown all the more a difficult task to arrive on a comprehensive criterion to identify Urdu compounds. In the end, all the possible dimensions to identify Urdu compounds, may not have explored. However, the research motivates lexicologists to further investigate the criteria to identify compounds of one of the most complex languages of the world.

REFERENCES

Adams, V. (2001). *Complex Words in English.* London: Pearson Longman.

Bauer, L. (1983). *English Word-formation, Cambridge*: Cambridge University Press

Hook, P. E. (1974). *The Compound Verb in Hindi*. The University of Michigan.

Katamba, F. (1993). *Morphology. Modern linguistics series*. New York: St.Martin's
               Press

Katamba, F. (2005). *English words*. Routledge Taylor and Francis group.
                London and New York.

Platts, J. T. (2002). *A Grammar of Hindustani or Urdu Language*, Sang-e-
                Meel Publications.

Schmmit, N. (2000). *Vocabulary in Language Teaching*. Cambridge University Press.

Schmitt, N.   (1997). *Vocabulary Description, Acquisition*. Cambridge University Press.