

Computational Description of Urdu Collocations

Dr. Mahmood Ahmed Awan¹

Dr. Abrar Hussain Qureshi², PAKISTAN

Abstract: *With the rise of the Firthian linguistics, the importance of bringing about a distinct shift from learning and teaching lexical items towards a consideration of lexis in larger units of language organization began to be felt. As meaning in text is cumulative and each lexical item is affected what follows it or precedes it. So the study of lexis does not mean the study of words in isolation. So, collocations play an important role in second language learning and teaching, especially at advance level. There is a slim body of literature on the study of Urdu collocations. The gap is deplorable especially in view of the fact that the use of computer has brought about in the study of language. The EMILLE corpus is a case in point. Employing the EMILLE, the present article aims to explore Urdu collocation vis-à-vis their structure and function. The study shows that collocations play a very significant role in the Urdu language. The finding has pedagogical implications. The teachers and material designers of Urdu must focus on the way collocations create meanings. The article concludes with the hope this article will at a spur to other researchers of Urdu to investigate the manifold aspects of the Urdu collocations.*

Key Words: *Collocation, description, Urdu*

Introduction

There are two main traditions in the study of collocations. The first tradition started with Firth (1951) while the second belonged to Sinclair (1966). The main difference between these two traditions was that Firthian study of collocation was not based on corpus while Sinclair's was based on corpus. In other words, Sinclair developed the Firthian tradition of collocation with its statistical aspects based on corpus. (Carter, R. & M. McCarthy 1988) observe:

“The central concepts of research done by J. R. Firth (1951) were originally "colligation and collocation". Unfortunately, colligation failed to take off and collocation was recognized and accepted only as an insight. Firth, who was considered by many to be the “Father of collocation”, believed in the separation of lexis and semantics because he thought collocation was the

central part of a word's meaning. From his research, he was responsible for bringing the term collocation into prominence during the 60's and the 70's".

The tendency of two words to occur together is called 'collocation' (Firth 1957) or a pair of words is considered a collocation if one of the words significantly prefers a particular lexical realization of the concept, the other represents. The term "collocation" was made more explicit by Halliday (1961; 1966) that encompasses the statistical distribution of lexical items in context. He observes "Collocation is the syntagmatic association of lexical items, quantifiable, textually, as the probability that there will occur, at n removes (a distance of n lexical items) from an item x, the items a, b, c..." (Halliday, 1961:276). Benson (1989) states that "collocations should be defined not just as a 'recurrent word combinations' 'but as 'arbitrary recurrent word combination'. "Arbitrary" as opposed to "regular" means that collocations are not predictable and cannot be translated word by word. According to Sinclair, (1991; 121) collocation is the concept of word co-occurrence, where certain words appear predictably next to or within a certain number of words from each other.

Collocational relations present the greatest problems of replicable identification and that a crucial difficulty remains one of interpreting how loosely or tightly collocating items are strung together (Carter, 1998). Explaining collocation, Lin (1998) observes that even though "baggage and luggage" are synonyms, only "baggage" can be modified by emotional, historical, or psychological considerations. This lack of valid substitution for a synonym is a characteristic of collocation in general (cf. Manning & Schutze 1999) Many have defined these patterns of co-occurrence in the grammatical context that they result primarily from syntactic dependencies or they can be lexical in that, although syntactic relationship is involved. Hoey (2005) defines collocation that it is a psychological association between words (rather than lemmas) up to four words apart and is evidenced by their occurrence together in corpora more often than is explicable in terms of random distribution. This definition is intended to pick up on the fact that collocation is a psycholinguistic phenomenon, the evidence for which can be found statistically in computer corpora.

Nevertheless the definition of the exact nature of a collocation varies from one researcher to the next. It is variously defined as a habitual word combination (Lin 1998) or a recurrent word combination. Of course, if words do regularly collocate in this way, we shall expect to find some semantic relationship among them; but this may be quite complex and indirect. Collocation is a purely lexical relationship; that is, it is an association between one word and another, irrespective of what they mean.

In contrast with collocation, the idea of colligation is that just as a lexical item may co-occur with another lexical item, so also it may occur in or with a particular grammatical function (Hoey, 2005, 43). Fontenelle (1997) defined grammatical collocation that it involved one

element from an open class and an element from a closed class, typically, but not necessarily, a preposition. According to Carter (1998, P.60) colligation is the co-occurrence on nouns, verbs, or adjectives with prepositions or certain features of grammar, for example, a comparative form of an adjective with the word "than", or the verb "deal" with the preposition "with".

Studies of collocation in English, Urdu and many other languages have tended to be within two distinct traditions: one oriented towards specifically grammatical and one towards specifically lexically patterning. The former has tended to result in studies that have been of distinct value to language learners and have to an extent emerged from the demands of particular pedagogic projects. While the tradition of lexical collocation have produced seminal studies which have contributed substantially to our understanding of lexis. According to Carter (1998):

“The processing of collocations involves a number of parameters, the most important of which is the "measure of association", which evaluates whether the co-occurrence is purely by chance or statistically significant. Due to the non-random nature of language, most collocations are classed as significant, and the association scores are simply used to rank the results. Commonly used measures of association include "mutual information, t-score, log-likelihood".

Collocations can be in a syntactic relation (such as object-verb in Urdu), lexical relation (such as antonymy), or they can be in no linguistically defined relation. Some collocations are so strong that it seems strange to list them as separate words, to be combined into the collocation each time we need it.

In order to describe collocations, Sinclair states:

“We may use the term node to refer to an item whose collocations we are studying and, we may define a span as the number of lexical items on each side of a node that we consider relevant to that node. Items in the environment set by the span, will be called collocates” (1966, P.415). Consider the following

Urdu Collocations

A recent perspective - corpus based statistical approach - has been adopted here to investigate collocations. As there is no major corpus available of Urdu language like The Bank of English, prepared by University of Birmingham United Kingdom, so, EMILLE, a corpus of Indian language for Bengali, Gujarati, Hindi, Panjabi, Singhalese, Tamil and Urdu, prepared by Lancaster University, United Kingdom, has been used. Though, EMILLE is very small corpus of Indian Languages yet it has been used as a point of departure. “Wordsmith Tools” a software

launched by Oxford University, has been used to analyze concordance and collocational pattern. An example of the collocations of Urdu from the EMILLE corpus of the word سبز is produced below:

N	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	کے	اس	میں	کا	سر	سبز	رنگ	کی	ہے	ہے	اور
2	درے	سال	ہے	کے	ے		اور	کے	کے	کے	ہے
3	ہیں	ہو	کے	اور	اور		درختوں	کا	اور	ڈھاک	ہیں
4		ہو	جاتی	ہے	مہے		و	سے	طرح	میں	کے
5		تھے	ہے	اور	چھ		لکے	میں	سے	اور	کی
6		رنگ	علاقوں	رنگ	ہو		درخت	اور	پے	اے	جو
7		میں	میں	مثلا	گہرا		ہوتا	ہے			
8		وہاں	میں					کو			
9			نے					بہرا			

In the above concordance, the lexical item سبز *sabz* is the node word and its collocates have been shown in a span of five on each side of the node..

Grammatical Collocation

Grammatical collocation is also called syntactic collocation. According to Biskup, 1992 and Bahns, 1993 grammatical collocations are the type in which a dominant word fits together with a grammatical word, for example, in the following concordance the noun is accompanied by a grammatical word □

- 165 موب اور رسمی ہماری ہیں۔ برابر سوہوں نامی # بد میں چھینچ ہے ہی ہے۔ ہماری ہے ک نامی اور
- 166 بڑا کا لوہے اپنا کو ان کے کے تھے نداءں بد کر بہت سے سامان کا آنا جانا دے کہ کر انگشت
- 167 اے سی اے سی ایک دانبتیز اور ہے گنا چھ نسب بد ثابت ٹو ڈانبتیز ہونے کا خطرہ دوسری کمیونٹس کے
- 168 ختم ہی پر وار چوتھے تے سرے # لڑکے ان نصیب بد سکمچی جاتی تھی مگر اے سا بہت کم ہوتا تھا۔ کنگز
- 169 ہوں۔ رہی ہے س مکنی ملا۔۔۔ جواب عورت نصیب بد کر وہ پکارنے لگا۔ اندر کون ہے۔۔؟ اے ایک بوڑھی
- 170 ہوں۔ رہی ہے س مکنی ملا۔۔۔ جواب عورت نصیب بد کر وہ پکارنے لگا۔ اندر کون ہے۔۔؟ اے ایک بوڑھی
- 171 تمام کے اتارنے پر عت بھوت پر لوگوں نصیب بد شے طانی آسے ب قرار دے لے ا جاتا تھا اور اے سے
- 172 اس تھی۔ نتے جہ کا حرکتوں اپنی کو اس نصیبی بد اور وہ تقدیر کا مارا تن تنہا رہ گئے۔ چونکہ عہ
- 173 بی رواج کا دف تک اب ہے ان کہ تھا نصیب # بد ڈھولک کا جواز کہے موجود نہ تھا، اور پنجاب اتنا
- 174 کر خاص اور کو حصے اس کے ملک تھی۔ نظمی بد درس گاہوں کے صدوں تک بہت شہرہ تھا، بڑی
- 175 بھجے # کو اس ہونے مانتے مظہر شعور کو نظمی بد فرق ہوگا؟ جسمانی عا عصوے اتی نظر عہ ذہنی
- 176 جو ہے اقدام کا طرف کی ترے ہی ہے اور نظمی بد دار ہوا میں بکھرنا گورے تنظیم اور ترے ہی سے
- 177 ہے ہوئی شروع سے سال ارب 18 صرف نظمی بد لے لے کہ اس نظر عے کی روسے ہے ترے ہی اور
- 178 منظر ہی جلدی کو اسباب اور مابیت کی نظمی # بد پے دا کر دے۔ عہ خال کے ا جانے لگا کہ دماغی
- 179 کرنا مطالعہ سے طرے # سائنٹیفک کا نظمی بد جن کا مشاہدہ نہ کراے جا سکتا ہو۔ اب اگر ہم ذہنی
- 180 بھی نقائص کے بن فطری غیر اور تصنع، نمائی بد پارکھوں کے نندے اردو ادب میں کہیں کہیں
- 181 دے نا کہہ بات کوئی۔ چاہے دے دے نا کسی نے کی بد اور تصور بھی ہے) اے کہ مقولہ ہے کہ بدی کا
- 182 نہیں لازمی غذائے بدمزہ اور لباس وضع بد نشے نی، تنہائی اور بے کاری نہیں ہے۔ فقر میں
- 183 جس تھے پرانی کی اس سکائے تے کی بصمی بد ڈاکٹروں نے بھی اس کا علاج کے ا تھا۔ اے سٹی اور

The grammatical collocation may also be analysed with the approximate collocates of a grammatical item in the form of a span on each side of the node. Most of the immediate collocates of Urdu grammatical words are nouns as in the following concordance:

- 120 تک آخر سے شروع جوان تو وہ ہے نہیں علاج لا لوگوں سے عہ کہوں گا کہ سطرع کے رےض
- 121 شفا سے علاج مخصوص کے آپ مرےض علاج # لا ترے بومے وہے تھ مے نہوتا ہے۔ سے نگرؤں
- 122 ہے ماری اس بشرطے کہ رہی نہیں ہے ماری علاج لا جہاں کار سے نوجن استعمال ہونگا ہو۔ کے نسر اب
- 123 چھوٹ سے ہے ماری والی جانے سمجھے علاج لا دے تا ہے۔ جس سے بفضلہ مرےض کی جان اسی
- 124 گا۔ کہنے پلے زمان تو ہے اگر؟ ہے مرض علاج لا رہی ہے۔ اس نے پوچھا تھا کہ کے ا پولے و اے ک
- 125 تک آخر سے شروع جوان تو وہ ہے نہیں علاج لا لوگوں سے عہ کہوں گا کہ سطرع کے رےض
- 126 کو ہے مارے ذہنی سے مدد کی جن ہے علم لا پڑتا ہے کہ قاری نہ صرف ان اصولوں سے قطعی
- 127 غیر مستقبل کہ ہے سمجھتے ہم اور ہے علم لا مستقبل کے متعلق (177) 19
- 128 ہے، مصروف میں اعمال ذہنی ہے جے دہ، علم لا ہے۔ دونوں # حصے اے ک دوسرے سے قطعی
- 129 ان ہے رہا کر کے ا ہاتھ دابند کا اس کہ گا رہے علم لا انداز میں جاری رکھے گا اور اس بات سے قطعی
- 130 عام کہ ہوا عہ مطلب کا اس، کرتی ظاہر علمی لا کے جاتے تو وہ اس کے فے ت سے # قطعی
- 131 ہے بات تو کی ہی کل کہ بتاے اور کی ظاہر علمی لا بورن بتاے اور دوکان کے بارے # میں قطعی
- 132 15 (33) تک ہے بلکہ کرتی۔ ظاہر # علمی لا پوچھے جاتے پر وہ اس تمام واقعے سے قطعی
- 133 ہے۔ کرتا ہے دا مسائل لے ہمارے اطلاق عمومی لا چہ مسلم معاشرے پر اسلامی صلی مطالعاتی نمونے

Lexical collocation may also be called semantic collocation. To attain a clearer understanding of lexical collocations, it is helpful to try to distinguish them from free combinations and idioms. Actually, it is a matter of frequency that constitutes lexical collocations, for example, سبز

“*sabaz*” (green) may collocate with لباس “*libaas* (clothes), جھنڈا *jhanda* (flag), میدان *midaan* (ground), etc. while free combinations do not occur in language frequently while on the other hand, idioms are relatively frozen expressions whose meanings do not reflect the meanings of their component parts. Lexical collocations do not contain preposition, infinitives, or clauses, but consist of various combinations of nouns, adjectives, verbs, and adverbs (Bahns, 1987). It normally consists of combinations of two basically “equal” words such as noun+verb, noun+noun, adjective+noun etc. in which both words contribute to the meaning.

N	Word	With	elation	Total	Left	Right	L1	Centre	R1
1	سبز	سبز	0.000	99	1	1	1	97	1
2	رنگ	سبز	0.000	20	3	17	3	0	17
3	اور	سبز	0.000	13	6	7	6	0	7
4	سر	سبز	0.000	12	12	0	12	0	0
5	ے	سبز	0.000	8	6	2	6	0	2
6	درختوں	سبز	0.000	7	0	7	0	0	7
7	و	سبز	0.000	6	1	5	1	0	5
8	مہے نے	سبز	0.000	4	4	0	4	0	0
9	درخت	سبز	0.000	4	0	4	0	0	4
10	لکے ر	سبز	0.000	4	0	4	0	0	4
11	ے	سبز	0.000	3	3	0	3	0	0
12	گہرا	سبز	0.000	3	3	0	3	0	0
13	ہوتا	سبز	0.000	3	0	3	0	0	3

The above concordance shows that the node word سبز “*sabaz*” (green) is frequently accompanied by nouns.

Cowie and Howarth (1995) suggest that such lexical phrases can be placed on a 4-level scale of complexity. At Level 1, idioms are multiword lexemes that have frozen collocation. If any variation is inserted into the idiom, it ceases to exist as a unit. In the idiom تارے گننا *taaray ginnaa* (not to sleep all the night) the meaning would be lost if even one element was changed. Thus idioms are the least complex because they allow no variation. Moving to Level 2, the collocation is still fixed, but the meaning is not idiosyncratic any longer; rather it is more transparent because it is composed from the meanings of all of the component words. This level is somewhat more complex because the meaning has to be composed from several lexemes, rather than coming from a list of words e.g. سر تا پیر *sir ta paer* (from head to foot). Level 3 has a slot that can be filled with a limited list of words, most of which are similar in meaning and would be placed at one point of collocation. This choice introduces both variation and complexity e.g. مثال قائم کرنا *mesaal qaim kerna* (to set example) and مثال دینا *mesaal deina* (to give example) etc. Level 4 has two slots instead of one, adding increased variation and complexity and new words may be placed at the both ends of collocation e.g.:

1. خوشگوار موسم *khushgawar*----- *mausum*
(pleasant) (weather)
2. ٹھنڈی رات *Thandi* -----*raat*
(cold) (night)

In the following cluster, the node word **پانی *paani* (water)** is placed with a variety of words on either side.

24	پانی خشک ہو جائے	14	3
25	گے لن پانی میں	14	3
26	میں ملا کر	14	3
27	کے لئے پانی	14	3
28	اور پانی کے	13	3
29	ارک گے لن پانی	13	3
30	کو پانی سے	13	3
31	کا پانی خشک	13	3
32	بے اور پانی	13	3
33	اس میں پانی	13	3
34	اس پانی میں	12	3
35	پانی کا ارک	12	3
36	اور پانی میں	12	3
37	سمندر کے پانی	12	3
38	کے پانی کا	12	3
39	اور تھوڑا پانی	12	3
40	پانی کے چھٹے	12	3
41	ٹھنڈے پانی میں	12	3
42	تالاب کے پانی	12	3
43	کی سطح پر	12	3
44	کی وجہ سے	12	3
45	سمندر کا پانی	11	3
46	بے کہ پانی	11	3
47	پانی دے جاتا	11	3
48	مقدار میں پانی	11	3
49	پانی میں جوش	11	3
50	پانی کی طرح	11	3

According to Leo Wanner (2002) “to summarise, the idiosyncratic interpretation of the term "collocation" covers all non-free combinations also covered by its distributional interpretation, but, in addition, it covers rare idiosyncratic combinations and rejects all recurrent free combinations. Therefore, it is the interpretation of the term "collocation" as an idiosyncratic combination that underlies the present study. It is to be noted that we make use of the following properties implied by this interpretation

1. a collocation is a binary combination of lexical items,
- 2 . a collocation possesses a coherent syntactic structure, i.e the base and the collocate always possess the same grammatical function with respect to each other,
- 3 . a collocation is a lexically restricted word combination, i.e. it cannot be constructed using universal (semantic) selectional restriction rules; rather, the base predetermines the set of lexical items (collocates) it may appear with in a combination based on idiosyncratic, collocation type-specific grounds

In what follows, some collocations from the EMILLE corpus have been analysed in order to see how far they measure up to the above interpretation.

Theoretically, it is possible for any lexical item of Urdu to co-occur or keep company with any other lexical item. However, for any particular lexical item A there are other items which have a high probability of being found near A. For instance, we might expect *نیا* *naea* (new) to have a high probability of co-occurrence with *گھر* *gher* (house), *کتاب* *kitab* (book), *شہر* *sheher* (city), etc. but a low probability of co-occurrence with *روٹی* *roti* (loaf), *ہوا* *hawa* (air) *بجلی* *bijly* (electricity), etc. The fact that the patterns are probabilistic raises some methodological problems. For example a fundamental difficulty in measuring collocational relations is that of deciding the maximum distance between items that can be said to be collocating. The solution generally adopted has been a relatively arbitrary of restricting the collocating items to a span of a fixed number of words on either side of the specified focal word whose patterning is being investigated. It is also difficult, however, to appropriately demarcate what might be the upper limit of a lexical unit and to decide whether a larger unit could itself be the limit of the collocational relationship. In Sinclair (1966) and Sinclair et al (1970) the aim was to study large quantities of text in order to focus in a statistically significant way on the company kept by particular words and for the strength and weakness of partnership to be expressed in terms of percentile frequencies of co-occurrence. But not enough data could be processed for

either a wide range of lexical collocational probabilities or interesting lexical sets to be described. The corpus used in the present study is too small to solve this thorny issue. A corpus larger in fairly size may do so.

In Urdu and many other languages of the world some words are relatively freer than other in forming pairs, e.g. اچھا *accha* (good) can have freer combination with many nouns اچھا گھر *accha ghar* (good House), اچھا لڑکا *accha larka* (good boy), etc. Some words occur only in company of some specific class of words. For instance, the adjective باسی *baasi* (not fresh) makes a collocation like باسی روٹی *baasi roti* (stale bread) باسی سبزی *baasi sabzi* (stale vegetable) but does not make a collocation with some uncooked edibles like گھیوں *gehu* (ghee) as باسی گھیوں *baasi gehu*. In a semantic field the relationship between the two words is paradigmatic forming a system of vertical substitution. In a collocation the relationship between the two words is syntagmatic.

A full description of the collocational behaviour of even a small group of lexical items requires the analysis of a great number of texts from different registers. The practical problems, such studies used to encounter, can now be partially overcome with the development of computational linguistics and the use of computers in linguistic analysis, and particularly in frequency counts. The analysis of larger samples will possibly give more valuable information about how language works at the level of lexis. In the first place collocational descriptions will indicate the different syntagmatic relationships; a certain lexical item enters into in different contexts of situation. The occurrences of a certain item on the syntagmatic axis are part of the meaning.

TYPES OF COLLOCATIONS IN URDU

Following are the main types of collocations in Urdu

- i. Unrestricted collocation: This describes the capacity of particular lexical items to be open to partnership with a wide range of item. Most core words fall into such category. For example پانی *paani* (water) collocates relatively unrestrictedly with a wide range of animate and inanimate, concrete and abstract entities: e.g. پانی لگانا *panni lagaana* (to water), پانی پینا *paani peena* (to drink water), پانی ڈالنا *panni daalnaa* (to throw water), پانی پلانا *paani peelana* (to make some one drink), پانی ملانا *paani meelana* (to mix water), etc. There are,

however, 'transitional areas' (Cruse, 1986 p.41) between free combinations/collocations and collocations/idioms. The fact may be proved with the help of corpus that the node word پانی *paani* (water) accompany with a big slot of words.

ii.

N Word	With elation	Total	Left	al Right	L1	Centre	R1	
1	پانی پانی	0.000	2,427	2	2	2	2,423	2
2	میں پانی	0.000	607	147	460	147	0	460
3	کے پانی	0.000	392	159	233	159	0	233
4	کا پانی	0.000	245	157	88	157	0	88
5	کی پانی	0.000	238	12	226	12	0	226
6	سے پانی	0.000	213	66	147	66	0	147
7	اور پانی	0.000	187	139	48	139	0	48
8	کو پانی	0.000	147	62	85	62	0	85
9	گرم پانی	0.000	93	85	8	85	0	8
10	پر پانی	0.000	66	51	15	51	0	15
11	دے پانی	0.000	65	0	65	0	0	65
12	ہے پانی	0.000	60	48	12	48	0	12
13	کہ پانی	0.000	43	43	0	43	0	0
14	اس پانی	0.000	42	32	10	32	0	10
15	ٹھنڈے پانی	0.000	39	38	1	38	0	1
16	اے پانی	0.000	37	21	16	21	0	16
17	جو پانی	0.000	31	26	5	26	0	5
18	خشک پانی	0.000	31	0	31	0	0	31
19	کر پانی	0.000	29	29	0	29	0	0
20	تھوڑا پانی	0.000	27	26	1	26	0	1

ii. Semi-restricted collocation

This category embraces lexical patterns in which the number of items, which can be substituted in different syntactic slots, is more determined. For example, the collocations of آسمانی *aasmaani* (heavenly) are comparatively restricted e.g. آسمانی کتب *aasmaani katub* (heavenly books), آسمانی رحمتیں *aasmaani rhematae* (heavenly blessings), آسمانی آفتیں *aasmaani aaftaen* (heavenly sufferings)

N Concordance

- 24 جن اور ہےں ہوچکی ضائع جو تھےں کتابےں آسمانی ہےں کہ وےدوں کے ظہور سے پہلے بھی کچھ
- 25 تو ہےں ڈالئے نظر پر اس ہےں کتابےں آسمانی سرےانی زبانوں میں جن میں قرآن کے قبل کے
- 26 سے سب ہمارے چاند چونکہ میں کروں آسمانی اور 000.00.50.3 میل دور ہےں۔
- 27 ثابت نے اس میں کتاب اس، حرکت کی کروں آسمانی، میں اےک کتاب شائع کی، جس کا عنوان تھا
- 28 موروثی جاگےر تھا۔ کرگے ا اختے ار شکل کی آسمانی دار پکارتے تھے۔ عہد محمد شاہی میں اجارہ بلانے
- 29 اولمپس کوہ نام کا جس تھا ضرور مسکن آسمانی انسانوں کی رسائی نہ ہوسکے۔ حالانکہ ان کا اےک
- 30 میں زمینے پرانے ہے۔ مقصد کوئی کا مظاہر آسمانی منعی پوشدہ ہےں اور کے ا ان کے اور دوسرے
- 31 صحیح پر طور مجموعی نقشہ جو کو نظام آسمانی کا موضوع ہےں کوئی ترتےب قائم نہیں کرسکتے۔
- 32 پر ہم جو ہے میں رازوں ان ابتدایی کی نظام آسمانی ہوتا ہے، برقرار رکھتا ہے۔ ے تمام عظیم الشان
- 33۔ ہے حصہ اےک کا قانون مضبوط کے نظم آسمانی کی حرکت ہےں ان کرتی ہے کہ ے، مظہر قدرت

iii. Restricted Collocation: These restricted are based neither on semantic range nor or meaning but confined to one or two words only e.g. چکنی *chilchilati dhoop* (scorching sun), چو پڑی با تیں *chikni chupri baate* (flattering words), گورا رنگ *gora rung* (white complexion). In other words, these collocations can appear only in the company of words of a specific class and not with others.

N Concordance

- 75 اپنی نے جنہوں ہوئی ثابت رحمت میں زحمت دھوپ چلچلاتی تو لئے کے سے احوں کچھ کے ملکوں
- 76 کبھی # جاتا. چلا ہوئے لگے ئ چھتری میں دھوپ چلچلاتی کی گرمی پوش سفے د کوئی کبھی گزرتے
- 77 کبھی # جاتا. چلا ہوئے لگے ئ چھتری میں دھوپ چلچلاتی کی گرمی پوش سفے د کوئی کبھی گزرتے
- 78 اثرات اچھے پر جسم اور ذہن انسانی دن والے دھوپ چمکے لی خوشگوار بھی دارومدار کا صحت انسانی

iv. Inherent collocation: Some words inherently require a human participant and some inhuman, and some either of them, similarly, some words are used in the context of animate participants and some of inanimate, and some of both. For instance, رونا *rona*: ‘to weep; represents a universal human activity

N Concordance

- 13 کر مار من نے اس کار آخر چلتا؟ نک کب **دھونا رونا** رکھا۔ مگر واہ رے زندگی کی کڑوی سچائی۔۔ ے
- 14 بھی خوشی بلکہ ہے نہیں ہی یہ صرف **دھونا رونا** کرنی چاہیئے اور دوسرا یہ کہ انسان کی قسمت میں
- 15 پہونچکر وہاں چلے۔ گھر اب چلو، ہوچکا **دھونا رونا** نے منشی کے کندھے پر اپنا ہاتھ رکھ کر کہا اب بہت
- 16 نمودار سلاہو اےک وہاں وقت اسی تھی ربی **رو رونا** زارو قطار رونے لگی۔ جس وقت وہ اپنی تقدیر کا
- 17 وگ و لے رہنے میں محلے کے اس رہتا۔ **رونا رونا** کی نے ند حرام ہوجاتی۔ وہ اکثر اپنی بدنصیبی کا
- 18 حقوق گھنٹے چوبے س نانی تمہاری ہے۔ **روتے رونا** دم گھنٹا ہے۔ تمہارے باپ ہے وقت اپنی بے ماری کا
- 19 مجھ کردے۔ # اعلاج میرا آپ بھرنا۔ **سسکے ان**، **رونا** میں چل پھر نہیں سکتا۔ کچھ کر نہیں سکتا اور
- 20 دے دلاسم اسے آپ ہے ظاہر بھرنا۔ **سسکے ان**، **رونا** ہوگے (52) 17 پھر رونا۔ اور
- 21 باہر تاکہ بونا سے زور زور بھرنا # **سسکے ان** **رونا** کر آئے تھا۔ پر اب تو میں چل بھی نہیں سکتا۔
- 22 ہے تھا وہ جب باہر حالانکہ ہے۔ دے تا کر **شروع رونا** ہے اور اب جب بھی آتا ہے تو # آنے کے ساتھ ہی

PATTERNS OF LEXICAL COLLOCATIONS IN URDU

1. Noun+Noun Collocation

N Concordance

- 88 جگر تھا۔ کے نسر کا جگر # کو اس تھا۔ پر **مرگ بستر** میں اےک مریض کو دے کھنے گئے اتھ۔ وہ
- 89 کہنا کا اشخاص موجود # کو فریب کے **مرگ بستر** واقع 30 اکتوبر 1883ء کو پے ش آئے۔ ان کے
- 90 باپ شفےق # انتہای اےک ہوئے بڑے پر **مرگ بستر** کی سمجھ سے بھی بالائے نہی۔ دے اٹند کے چچا
- 91 جگر تھا۔ کے نسر کا جگر # کو اس تھا۔ پر **مرگ بستر** میں اےک مریض کو دے کھنے گئے اتھ۔ وہ
- 92 کہاں جانے نہ بخود خود کے وچادر تھے ہ **مع بستر** صوفہ کا بٹن دبا ے جس میں سے اےک بہت بڑا
- 93 از کم کا گھر میں موسم سرد رہیں۔ گرم بھی **میں بستر** یا سر کا سکارف استعمال کیجیئے۔ رات کے وقت

2. Noun+Verb Collocation

According to Allerton (1984) verb-noun constructions in which the verb carries little semantic content and is used for the sake of its structural properties only (1984) as in

- 443 0 ابتدائی ہوا۔ معلوم اچھا بہت مجھے لکھوایے۔ حظ مری انسانی ن سب بچوں سے جبہ ہر مجھے
- 444 کو کونسل کنسیومرس گیس ریجنل اپنی یا لکھنے خط آر ڈی کو 41 مل بینک، لندن ایس ڈبلیو
- 445 کو آپ۔ کریں درج شکایت اپنی میں جس لکھیں خط ہے کہ آپ اس لیفلٹ میں شامل فارم کو پُر کریں، یا
- 446 ایس کاپیاں کی رپورٹ۔ بھیجیں میل ای یا لکھیں خط۔ بصورتِ دیگر اوپر دیئے ہوئے پتہ پر اپنا کیولا کو
- 447۔ کرسکین مدد کی آپ ہم کہ ہے سکتا ہو۔ لکھیں خط کو مکمل کریں، یا ہمیں اپنی شکایت کے بارے میں
- 448 ملی پسندیدہ کا آپ لیے کے آپ ہم تاکہ، لکھیں خط ہیں یا پاکستانی، جلتزنگ آپ کا اپنا پروگرام ہے ہمیں
- 449 پوری فرمائش کی آپ ضرور ہم تو گئے لکھیں خط سنانے کا بہت شکر یہ، بھئی اگر آپ محنت سے ہمیں
- 450 پروگرام حسین ممتاز، ہے یہ پتہ ہمارا، لکھیں خط بھیجیں یا خود email، کے خطوں کا انتظار رہتا ہے
- 451 د کافی، گی لکھوں ضرور جی انشالہ، گی لکھیں خط اور ہمیں امید ہے کہ آپ پروگرام کے بارے میں ہمیں
- 452 اور ہیں چاہتے ہوا منٹ دس کر بج گیارہ، لکھیں خط کرتے ہیں، اگر آپ نے کوئی خاص نغمہ سننا ہے تو
- 453 ہے یہ پتہ ہمارا لیے کے اس سامعین اور لکھیں خط اگر آپ نے کوئی خاص نغمہ سننا ہے تو آپ ہمیں
- 454 سی بی بی جلتزنگ پروگرام حسین ممتاز لکھیں خط ہے، کیا آپ کسی کو سلام و آداب کہنا چاہتے ہیں
- 455 چٹھی کی بچے کے آپ اگر۔ گی لکھے خط کو اور سالانہ نظر ثانی میں ماخوذ پیشہ ور ماہرین کو
- 456 تاریخ کی اجلاس کے نظر ثانی میں جس گا لکھے خط کی حامی بھرے گی۔ آپ کے بچے کا سکول آپ کو
- 457 پتہ ہمارا تو چاہیں لکھنا خط ہمیں آپ اگر، لکھے خط گئے ہمارا ساتھ دیا، ہمیں فون کیے، خوبصورت
- 458 گئی دی وجوہات کی فیصلہ میں جس گی لکھے خط ضروری نہیں تو وہ آپ کو اور آپ کے سکول کو
- 459 قلم میں خدمت کی آپ پھر آج، ہونے لکھے خط سلام، آپ کہتے ہیں کافی عرصہ بیت گیا آپ کو
- 460 بھی ے کو کسی سے # مے ہم اور گئے لکھیں خط نہیں کر سکتے تھے کہ ڈاکٹر محمود اس قسم کا
- 461 ہیں درج تفصیلات میں آخر کے لیفلٹ س۔ لکھیں خط کی انکوائری لائن کو فون کریں یا اس کے آفس
- 462 نقص بصرہ گورنر اب۔ پہونچ ساتھ اے کر لے خط مسجد کی شہادتیں لے کر، اور قاصد خلعہ کا

3. Adjective + Noun collocations.

Noun has often typical adjectives which go with them.

- 1,080 بے ظاہر ہیں۔ جاتے گزر ہوئے دے کہتے سے نظر سرسری اے کو ان ہم مگر ہیں گزرتی چیزے
- 1,081 اخبارات صرف وہ کہ ہوگا معلوم بھی سے ڈالنے نظر سرسری اے کو پر تاریخ کی صحافت میں
- 1,082 ثابت مند سود خبیروں صرف اور ہے ڈالنے نظر سرسری پر اخبارات قاری ہے۔ چکی بڑھ بہت
- 1,083 کی اشتہارات اور خبیروں صرف اور ہے ڈالنے نظر سرسری پر اخبارات قاری ہے۔ چکی بڑھ بہت
- 1,084 کے تربیت بغیر چنانچہ ہے دے کہتے سے نظر سرسری نہایت اشے کی گردوپے شے بچ
- 1,085 اخبارات صرف وہ کہ ہوگا معلوم بھی سے ڈالنے نظر سرسری اے کو پر تاریخ کی صحافت میں
- 1,086 وہ کہ ہے ہوتا پے دا ے سوال اب تھے۔ آتے نظر سرشار سے احساس ہی اسے کے مسرت پر
- 1,087 زند تشاکل غر کا وقت کہ ہے ہوتی سے کرنے نظر سرف سے حقے قت اس غلطی ہے۔ جاتا کہا Force
- 1,088 سے شکار اور کرنے جمع 0 غزا گے۔ ڈالے نظر سرف # سر پر اداروں معاشی اور معیشت کی شکار
- 1,089 اصول روحانی روسے کی نظر ے اس ہے۔ ڈالی نظر سرفی سر بہت پر جبلت میں نظر ے اتے نفسے اتی
- 1,090 کچھ دے۔ کر زدہ حیرت ہمیں نے جس آئی نظر سطح نے لی نے لی اے کو پر دور تو اٹھی اور پر سے
- 1,091 اس ظاہر یہ جو گی ملے مثالیں کئی سے نظر سطحی ہے۔ ناممکن پر بنا کی دوم قانون ے لے کن
- 1,092 بتدریج وہ کہ ہے ہوتا معلوم سے دکھنے سے نظر سطحی، ہے تعلق تک موٹف منڈل تک جہاں 1
- 1,093 نے برےخت میں ڈرامے پر۔ ہے لگتا آتے نظر سفر والا ہونے ختم نہ کیھی اور طوےل کا اس سے
- 1,094 نم لگو متفگطے س قرےن کے گس ہے نے گینے نظر سفوف کے لوہے گور گنہک نہی گن کو نم کے گ

1. Adverb + adjective collocations: Some adverbs modify particular types of adjectives.

N Concordance

- 2 سے نظروں بھری محبت طرح کی باپ شفیع # انتہای دے اند کے چچا بسٹر مرگ پر پڑے ہوئے اے ک
- 3 بو۔ رہا ہو فے ل اور ہو۔ ہوچکا لاغر انتہای بہت زے ادہ ہوتا ہے۔ بالخصوص اس وقت جب دل
- 4 بو۔ رہا ہو فے ل اور ہو۔ ہوچکا لاغر انتہای بہت زے ادہ ہوتا ہے۔ بالخصوص اس وقت جب دل

Concluding Remarks

Although it is widely acknowledged that collocation are both indispensable and problematic for language learners and that they therefore should play an important role in teaching and learning second language, especially at an advanced level. Study of the Urdu language at the level of Collocation is a new phenomenon. But owing to its significance, It has great importance. The present study tried to investigate collocations in Urdu language in detail yet, the study is far from being comprehensive in view of the fact that the size of the corpus used in this study was small. A larger corpus would have made this study more comprehensive in which more dimensions might have been investigated in the end. Investigation of collocations of the Urdu language is a vast and interesting field. The authors of this article hope that this research will motivate linguists to explore more dimensions of Urdu collocations in more detail.

References

- Allerton, D.J. (1984) Three or Four Levels of Co-occurrence Relations. *Lingua* 63:17-40.
- Benson, M. (1989) *The Structure of the Collocational Dictionary*.
International Journal of Lexicography. 2 (1): 1-13.
- Biskup, D. (1992). *L1 influences on learners, renderings of English collocation: A Polish/German empirical study*. In P.J.L.Arnaud and Bejoint H. (Eds.),
Vocabulary and Applied Linguistics (pp.85-93). London: Macmillan.
- Bahns, J. (1993). *Lexical Collocations: A Contrastive View*. *ELT Journal* 47.
- Carter, R.(1998). *Vocabulary. Applied Linguistic Perspective* (2nd Edition)
London and New York: Routledge.
- Carter, R. & McCarthy, M. (1988) *Vocabulary and Language Teaching*,
(London, Longman) pp. 56-75

Cowie, A.P. (1978) “ The place of illustrative material and collocations in the design of a learner’s dictionary. Firth (1951) ??????????????????????

Fontenelle, T. (1997) *Turning A Bilingual Dictionary Into Lexical-Semantics Database*. Published in cooperation with the dictionary Society of North America (DSNA) and the European Assciation for lexicography (EURALEX)

Halliday, M. A. K. (1961) Categories of the Theory of Grammar. *Word* 17: 241–292.

Halliday, M. A. K. (1966) Lexis as a Linguistic Level. In: C. E. Bazell et al., editors, *In Memory of J. R. Firth*. Longman, London.

Hoey, M. (2005) *Lexical Priming, A New Theory Of Words And Language*. Routledge, Taylor And Francis Group, London And New York.

Lin, D. (1998) Extracting collocations from text corpora. In *First Workshop on Computational Terminology*, Montreal, Canada.

Manning, C. D. and Schutze, H. (1999) *Foundations of Statistical Natural Language Processing*. The MIT Press, Cambridge, Massachussts.

Sinclair, J. (1966) Beginning the study of Lexis. In the memory of J.R.Firth., Charles E.Bazell, John C. Catford, Halliday, and Robert H. Robins (eds.), 148-162. London: Longman.

Sinclair, J. (1991) *Corpus, Concordance, Collocation*. Oxford: Oxford University press.

Wanner, L. (2002) *Towards automatic fine-grained semantic classification of verb-noun collocations*. Natural Language Engineering. Cambridge University Press

First Author: Dr.Mahmood Ahmed Awan. Assosicate Professor Govt.College Burewala. Pakistan

Second Author: Dr. Abrar Hussain Qureshi , Assistant Professor of English Govt. College Chichawatni Pakistan