

## Semantic Frames of the Urdu Conjoint Verb *lagnā*: A Corpus-based Study

Humaira Jehangir<sup>1</sup>, Musarrat Azher<sup>2</sup>

<sup>1</sup>Postgraduate researcher, LAEL, Lancaster University

[h.jehangir@lancaster.ac.uk](mailto:h.jehangir@lancaster.ac.uk)

<sup>2</sup>Assistant Professor, Department of English, University of Sargodha

[musarratazher@gmail.com](mailto:musarratazher@gmail.com)

### Abstract

In this corpus-based study, we present semantic frames based on principles of FrameNet<sup>1</sup> of Urdu conjoint verb *lagnā*. In research studies on Urdu, while translating from Urdu to English or for grammatical analysis, *lagnā* is typically translated in English as ‘attach’ creating a perception that this is its prototypical meaning. However, Online Urdu Dictionary<sup>2</sup> (OUD) shows at least seven different senses in natural language. The sources for the theoretical framework of the study include frame semantics (Fillmore & Atkins, 1992), and polysemic meaning (Traugott & Dasher, 2002). This situates the study in the cognitive linguistic framework. The analysis examples are extracted from 95.4 million words Urdu monolingual corpus (Jawaid et al., 2014) which can now be accessed using CQPweb<sup>3</sup>. The frequency of occurrence of types of *lagnā* and observation of concordance lines of extracted examples helps in analysing to what extent the different patterns in which the verb *lagnā* generates different meanings. The preliminary analysis showed that there are two overarching patterns NOUN+ *lagnā* and VERB + *lagnā*. Within these two patterns, we observed and analysed twelve semantic frames, of which BEGIN is the most frequent frame followed by TOUCH with the second-highest frequency. This analysis also helps in understanding that certain verbs in Urdu also have polysemous use and their meanings are context-dependent that can be established through a corpus-based analysis. This analysis and its findings are especially beneficial for the Urdu language scholars who have an interest in translation studies or (descriptive) grammatical analysis to look for different interpretations as well as integrate semantic frame and corpus methods rather than relying on restrictive meaning of this verb (as well as other polysemous verbs) for data analysis.

**Keywords:** Cognitive linguistics, corpus-based, frame semantics, lexico-grammatical study, Urdu polysemous verbs

---

<sup>1</sup> <https://framenet.icsi.berkeley.edu/fndrupal/>

<sup>2</sup> <http://202.142.159.36:8081/oud/default.aspx>

<sup>3</sup> <https://cqpweb.lancs.ac.uk/urmono/>

## Semantic Frames of the Urdu Conjoint Verb *lagnā*: A Corpus-based Study

One of the greatest impacts of using computerized data in linguistics studies has been the study of lexicon using corpus-based methods. Corpus linguistics has contributed not only to the development of English lexicographic works (see Biber et al., 2012; Granger & Paquot, 2012; Moon, 2007; Renouf & Sinclair, 1991; Rundell & Fox, 2002) but also in clarifying the relationship between grammar and lexis (see Sardinha, 2019). There has also been considerable corpus-based work on the polysemous nature of various verbs (e.g., Hsiao & Chung, 2018) and modal verbs (e.g., Viebahn & Vetter, 2016).

There has been ongoing work on lexicography in Urdu to digitize sources, especially major contributions are by the Center for Language Engineering<sup>4</sup> (CLE) and Urdu Lughat<sup>5</sup>. Similarly, considerable research on Urdu grammar (e.g., Butt & Geuder, 2011; Davison, 2014; Manetta, 2019), lexical sequence (e.g., Fazalehaq, 2019), lexicon and grammar relation (e.g., Davison, 2011) has been conducted. However, fewer works focus specifically on the polysemous nature of verbs in Urdu (e.g., Ahmed, 2010) and none focusing on the relation between polysemous verbs and varied senses in different contexts. The present research aims to highlight this identified gap and for this purpose, we have studied the Urdu polysemous verb *lagnā*<sup>6</sup> ('to attach') which in translation, especially for descriptive grammar studies of Urdu (e.g., ADD), is rendered in its prototypical meaning (i.e., to attach). Urdu Online Dictionary (OUD) shows *lagnā* 'to attach' can be used in a number of other senses. The purpose of focusing on *lagnā* 'to attach' is to not only examine and report its different senses and functional realization within clauses but also to raise the issue of rendering *lagnā* 'to attach' and other similar verbs in glossing only in their prototypical meaning. This work could serve as a groundwork for future corpus-based studies of polysemous Urdu verbs, and also help researchers in understanding the need to use the relevant meaning of the verb in the syntactic environment of its occurrence rather than a typical meaning that does not do justice to its meaning in context. We address the following questions in our paper:

1. What are the different senses of *lagnā* 'to attach' that data shows?
2. To what extent, the use of *lagnā* 'to attach' in clauses is context dependent?

This paper is structured as follows. In section 2 we briefly inform about verb *lagna* as defined in Urdu grammar and dictionaries. In section 3 we briefly discuss the theoretical underpinnings for this work, that is, frame semantics and lexico-semantics. This is followed by section 4 on research methodology. Then section 5 covers our analysis and findings. In section 6, we discuss and give the conclusion of this work.

### Brief information on verb *lagnā*

*Lagnā* is a conjoint verb and typically functions as a transitive verb. Koul (2008, p. 102) defines the conjoint verb as a verb that occurs in a conjoint verbal construction, that is, construction that consists of a noun, or an adjective, and a verb (typically an auxiliary verb) and a conjoint verb that takes all the inflection as in (1).

- |            |               |             |             |
|------------|---------------|-------------|-------------|
| 1. kām=mēm | lagā          | hūā         | hai         |
| work=in    | busy.PFV.M.SG | be.PFV.M.SG | be.PRS.3.SG |

<sup>4</sup> <https://www.cle.org.pk/>

<sup>5</sup> <http://udb.gov.pk/>

<sup>6</sup> Latinised/Romanised script used in this research follows ISO-15915 developed for Southeast Asian languages including Devanagari and Perso-Arabic script.

[he] is busy in work.

Lagnā can be an intransitive verb where the subject is marked with an oblique case, and the predicate is psychological as in (2).

2.        mujhe            pyās    laggī                    hai  
              1.SG.ACC        thirst   feel.IPFV.F.SG        be.PRS.3.SG  
              I am [feeling] thirsty.

As Urdu is a gendered language, forms of verbs “agree with their subjects in gender and person” (Koul, 2008, p. 93). In case an auxiliary verb (e.g., *hō* ‘be’) combines with *lagnā*, the auxiliary verb is inflected to show subject and gender marking. Aspect marking is taken by verb element in verbal conjunct constructions (Koul, 2008, p. 95). Urdu, similar to Hindi, has three tense forms: “present, past, future” (Koul, 2008, p. 4), and three grammatical aspects “habitual, progressive, and perfective” (Koul, 2008, p. 105). We present verb forms of *lagnā* in Table 1 and forms of *lagnā*, according to tense and aspect, in Table 2.

**Table 1**

*Verb forms of lagnā*

FORM	EXAMPLE
Stem	Lag ‘attach/feel’
Imperfective	Lagtā(m.sg)/ Lagtī(f.sg )/ Lagtē(m/f.pl) ‘attach/feel’
Infinitive	Lagnā m.sg)/lagnī (f.sg)/lagnē(m/f.pl) ‘to attach/to feel’
Infinitive oblique (with clitic mēm)	Lagne=mēm ‘in attaching/in feeling’
Infinitive oblique (with clitic se)	Lagne=se ‘from attaching/from feeling’
Infinitive oblique (with clitic ko)	Lagne=se ‘for attaching/for feeling’
Perfective	Lagā(m.sg)/Lagī(f.sg)/Lagē(m/f.pl) ‘attached/felt’
Subjunctive	Lagūm(1.m/f.sg)/Lagō(m/f.sg)/Lagēm (m/f.pl) ‘may attach/may feel’

**Table 2**

*Verb forms of lagnā with tense and aspect inflected for gender and numbers (defining details adapted from Koul, (2008, pp. 105-200).*

<b>Form TENSE</b>	<b>Masculine</b>		<b>Feminine</b>	
	<b>Sg</b>	<b>Pl</b>	<b>Sg</b>	<b>Pl</b>
<b>Present</b> (Represents ongoing action, repetitive habit/characteristic/action)	Lagtā hai ‘feels to be’	Lagtē haiṁ ‘feels to be’	Lagtī hai ‘feels to be’	Lagtīm haiṁ ‘feels to be’
<b>Past</b>	Lagtā thā ‘felt to be’	Lagtē thē ‘felt to be’	Lagtī thī ‘felt to be’	Lagtīm thīm ‘felt to be’
<b>Future</b>	Laggē gā ‘will feel to be’	Laggē gēm ‘will feel to be’	Laggē gī ‘will feel to be’	Laggē gīm ‘will feel to be’
<b>GRAMMATICAL ASPECT</b>				
	<b>Masculine</b>		<b>Feminine</b>	
	<b>Sg</b>	<b>Pl</b>	<b>Sg</b>	<b>Pl</b>
<b>Habitual</b>	Lagtā hai ‘feels to be’	Lagtē haiṁ ‘feels to be’	Lagtī hai ‘feels to be’	Lagtīm haiṁ ‘feels to be’
<b>Presumptive habitual</b> (Habitual and presumed actions; not known definitely)	Lagtā hōṁgā ‘would seem/feel to be’	Lagtē hōṁgae ‘would seem/feel to be’	Lagtī hōṁgī ‘would seem/feel to be’	Lagtīm hōṁgīm ‘would seem/feel to be’
<b>Subjunctive habitual</b> (Both habitual and hypothetical/speculative actions; not a direct guarantee to happen)	usē laggē ‘he might feel/think’	unhīm laggē ‘they might feel/think’	Usē laggē ‘she might feel/think’	unhēm laggē ‘they might feel/think’

<b>Progressive</b> (Formed by addition of auxiliary verb immediately after stem form of verb; auxiliary verbs agree with/inflected for person, number gender)	Lag rahā hai(present) ‘feeling’	Lag rahē haīm (present) ‘feeling’	Lag rahī hai (present) ‘feeling’	Lag rahīm haīm (present) ‘feeling’
	Lag rahā thā (past) ‘feeling’	Lag rahē thē (past) ‘feeling’	Lag rahī thī (past) ‘feeling’	Lag rahīm thīm (past) ‘feeling’
<b>Presumptive progressive</b> (Action or state of affairs extended in time and presumed to be occurring)	Lag rahā hogā ‘must be feeling’	Lag rahē hoīngē ‘must be feeling’	Lag rahī hogī ‘must be feeling’	Lag rahīm hoīngīm ‘must be feeling’
<b>Subjunctive progressive</b>	Lag rahā ho ‘possibly/would feel/seem’	Lag rahē hōm ‘possibly/would feel/seem’	Lag rahī hō ‘possibly/would feel/seem’	Lag rahīm hōm ‘possibly/would feel/seem’
<b>Perfective</b> (indicates action or state of affairs has been completed)	Lagā hai (present) ‘has felt/attached’	Lagē haīm ‘have felt/attached’	Lagī hai ‘has felt/attached’	lagīm haīm ‘have felt/attached’
	Laga thā (past) ‘had felt/attached/applied’	Lagae thē (past) ‘had felt/attached/applied’	Lagī thī (past) ‘had felt/attached/applied’	lagīm thīm (past) ‘had felt/attached/applied’
<b>Presumptive perfective</b>	Lagā hōgā (present) ‘would have felt/attached/applied’	Lagē hōīngē ‘would have felt/attached/applied’	Lagī hōgī ‘would have felt/attached/applied’	lagīm hōīngīm ‘would have felt/attached/applied’
<b>Subjunctive perfective</b>	Lagā ho ‘may/might have felt/attached/applied’	Lagē hōm ‘may/might have felt/attached/applied’	Lagī hō ‘may/might have felt/attached/applied’	Lagīm hōm ‘may/might have felt/attached/applied’

## Theoretical framework

One area of focus in corpus-based descriptive grammar research is collocations (i.e., co-occurrence of searched word with other words) and how meaning is created through the interaction of lexicon (vocabulary/semantics) and grammar (structures). Halliday and Webster (2009, p.63) say that speakers and writers make linguistic choices that show “relations between an element and what goes together with it”. Sinclair (1991, p.164) says that in a linguistic analysis grammar and lexicon are not an integration of the two fields, rather it is an interaction, that is, “it is fundamentally grammar with a certain amount of attention to lexical patterns within the grammatical frameworks”. Sinclair (1991, p.83) further points out that we need to start our analysis from an element and look at corpus evidence to observe the patterns that word occurs and resulting meanings and functions of that element within various structures.

*Frame semantics* broadly characterises the integration of lexis and syntactic patterns and resulting lexical meanings in terms of semantic frames, that is, lexical units evoke certain meanings that represent specific situations (Verdaguer et al., 2020, p. 62). Fillmore (1985) explains that

The feature-based approaches using primary categories are not likely to demonstrate the semantic manifestation and fullness of meaning of words because the meanings of words consist of vast information about the words enveloping us which can never be displayed within a few numbers of primary categories (p.383).

This concept is rephrased by Fillmore and Atkins’ (1992) in their notion of frame semantics, who propose that each sense of an element evokes a certain conceptual structure (i.e., semantic frame). Verdaguer et al. (2020, p.62) give the example of *concern* to show that a polysemous element “can evoke other semantic frames”. They say that the verb *concern* evokes the frame “Topic” when it occurs in a sentence “*These questions concern rhetorical issues*” and evokes the frame “Cause\_emotion” when occurs in “*it concerns me that people are not getting enough help*” (Verdaguer et al., 2020, p.62).

Fillmore et al. (1992) in their pioneering work have created FrameNet<sup>7</sup> which is a digitised, corpus-informed system. On FrameNet, various English semantic frames (based on mental/cognitive concepts that are evoked from lexical units occurring in patterns) have been indexed and made available. Each of these semantic frames is “a description of an event, association and its participants which are called elements of the frame” (Safari & Rahmatinejad, 2018, p. 64; see also Boas, 2020). FrameNet is a good source to be used as a reference for corpus-based English examples which have been marked both semantically and syntactically. Each frame has some core elements that are labelled according to the semantic role which the subject or the object has with the main verb (e.g., agent). These core elements are referred to as core frame elements (CFEs).

Fillmore and Atkin’s first lexical resource FrameNet project (1992) on their Frame Semantics theory (1992) has by now diversified from English only to other languages. For instance, Sanacore et al., (2019) research on French derivational relation by application of frame-like structures inspired by Frame Semantics and FrameNet. Another example is of

---

<sup>7</sup> <https://framenet.icsi.berkeley.edu/fndrupal/>

Safari and Rahmatinejad (2018) who developed a frame-based lexicon of Persian words on tourism and touring using FrameNet.

This paper also follows distributional theory in the lexico-semantic approach. Metayard and Vigliocco (2018) say that large data can be used to understand how meanings of words formalize their different relationship with other words in diverse contexts. They say that such relations that words have with each other can be reconstructed through the compilation of distributional information from data. Various methods to compile distributional information is by plotting links between words networks (e.g., Collins & Loftus, 1975), that is, by marking association of words based on how frequently certain words appear near certain other words (e.g., Shaoul & Westbury, 2010). Metayard and Vigliocco (2018, p. 72) say that as corpus methods are used to examine words at phrase or clause level and also to compute the frequency of occurrences of words in different patterns statistically, distributional patterns “formalize what a word means by how often and in what typical patterns it occurs with other words”.

Finally, we take into account the role of context in lexico-semantic approach. Hanks (2006, p.75) says that the meaning of a verb “is determined by the totality of its complementation” (i.e., part of the sentence which completes the meaning of a verb) patterns it occurs in. By application of the lexical semantic approach, “mapping between the semantics of verbs and their associated syntax is discussed in terms of ... feature selection” (Stringer, 2019, p. 180). In other words, the speaker’s choice of verbs involves their inherent meaning as well as the contextual features of the constructions they occur in (Stringer, 2019, p. 188). Therefore, context is an extraneous process that takes place when a word is used by speakers to convey specific semantic information.

In Urdu, there is an increase in corpus-based descriptive studies of grammar. For example, a study of hypotactic and paratactic thematic relations in Urdu clause complex (Yaqub & Shakir, 2019), description of Urdu affixes (Tanveer et al., 2021), a crosslinguistic study of meta-discourse markers in English and Urdu E-newspaper editorials (Shahid et al., 2020), and semantic inventory of collocations of NOUN+VERB in Urdu (Abdullah et al., 2021). However, there still needs to study on the interrelation of lexical items, their semantic meaning, and functions based on syntactic patterns that can be observed through collocational information.

## Methodology

This section presents the data used in this study, the method of extraction of the examples, and the data analysis procedure.

### Data sources

For this study, we use Charles University Urdu Monolingual Corpus 2014 (CUUMC) can be accessed on CQPweb<sup>8</sup>, as our primary source for data retrieval. CUUMC is a monolingual written corpus of approximately 95,000,000 Urdu tokens and open sourced for download<sup>9</sup> by Jawaid et al. (2014). The data has been compiled from Urdu material from

<sup>8</sup> <https://cqpweb.lancs.ac.uk/urmono/index.php>

<sup>9</sup> <https://lindat.mff.cuni.cz/repository/xmlui/handle/11858/00-097C-0000-0023-65A9-5#>

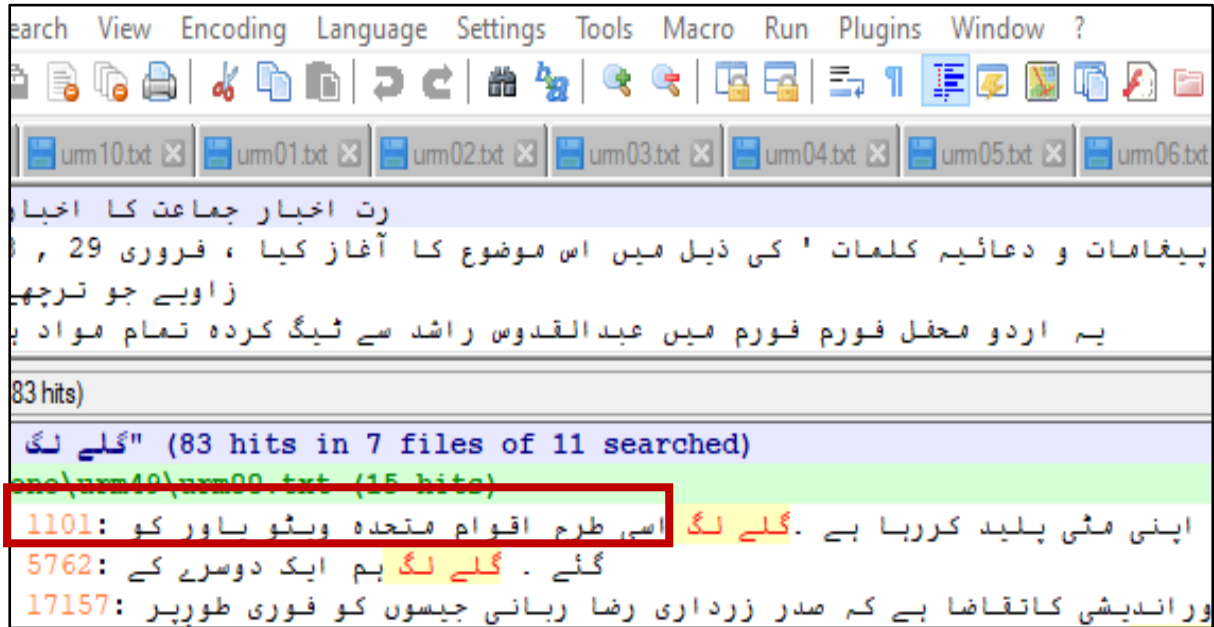
various websites<sup>10</sup>. The text of data consists of both formal and informal writings online (e.g., news, blogs, literature). It has been POS-tagged (annotation with parts of speech) by Sajjad and Schmidt's (2009) Urdu POS tagging scheme.

In addition, we have consulted an online dictionary by the Center for Language Engineering (CLE) which has provided meanings for words with examples from corpus. The online Urdu dictionary has helped us in establishing the most common senses of *lagnā* in use. These examples of *lagnā* in the Online Urdu dictionary were later used for cross referencing for the meanings we identified while we observed concordance lines of the data used in this study.

All the examples given in the analysis section are from CUUMC. As the data has been gathered from online sources, therefore, the corpus builders of CUUMC data have scrambled and then split the sentences at clause level before making it available for the public to avoid any ethical or copyright issues. Thus, we split the whole data into 49 files of approximately equal size of 14 MBs. Then for each example from the CQPweb, we searched the files on Notepad++ (see Figure 1) and then added the reference number to the example. So, each example has a reference number as follows (00\_123456) where 00 is the file number and 123456 is the line number as displayed in Figure 1. Figure 1 shows that once a phrase from example (20) is searched in multiple files (in image11files), the result shows (see in the left bottom highlighted part) that the searched example occurs on 00 file out of 00-10 files, in line number 5762.

**Figure 1**

*Search for example in multiple files of CUUMC in Notepad++*



<sup>10</sup> Web sources: <http://www.bbc.co.uk/urdu/>; <http://www.urdulibrary.org/>; <http://www.urduweb.org/planet/>; <http://shahfaisal.wordpress.com/>; <http://awaz-e-dost.blogspot.cz/>; <http://www.minhajbooks.com/urdu/control/txtformat/>







2012, p. 14) and to account for any example giving different sense in their correct frame. Also, within each category we assessed, for the sake of description, whether the patterns evoke the meaning in a literal sense or metaphorical one as in Figure 5, word *āg* ‘fire’ can be seen to occur in both types. Though for this study we have not separated the senses according to their literal and metaphorical meaning. We rather classed them according to the meaning these phrases evoke in the context they occur in.

Once frames are assigned based on syntactic and semantic properties of *lagnā* patterns, the coordination between meanings and semantic environments (based on lexicogrammatical environments, Tucker, 2001) of *lagnā*, the association of distribution (defined as how often words appear with other words by Shaoul & Westbury, 2010) and frequencies of each sense was analysed.

### Figure 5

*Concordance result of āg ‘fire’ cooccurring with lagnā in CUUMC*

Your query "[word="آگ"] [word=". \*"] [word=" | لکھی گئی | لکھا گیا ]" returned 125 matches in 1 text (in 95,038,214 words [1 text]; frequency: 1.315 instances per million words) [0.453 seconds]

|<  < >> >|

Show Page: 1

Line view

Show in random order

Choose action... ▼

No	Text	Solution 1 to 50	Page 1 / 3
1	<a href="#">only</a>	ہوئی تھی ۔   بہت دیر ہو گئی ، عمار ، وارث ، فہیم ، لوڈ	<a href="#">آگ نہیں لگی</a>   اس وقت بلوچستان میں نپلم کے لیے   مقام صحرات آباد تو اتنی آیت تکمیل دیں
2	<a href="#">only</a>	گا۔   بڑی دیر کر دی مہریاں نے اُترے آئے   ہماری طرف سے بہت سا پیار کجھے	<a href="#">آگ خود لگے</a>   چائنگ نے کہا کہ وہ بھی اپنے اسلاف سے بدل سکتا ہے یہ لکھ لیں
3	<a href="#">only</a>	ہوئی   ویسے آج کل سارے نیٹ ورکش خود کار کوئفگریشن فراہم کرتے ہیں ، فون	<a href="#">آگ پراپر لگی</a>   ہر ایک کے میں ٹھوڑا چیوا تھا تو ہے ۔   سو غز قبول کر کے مشکور ہوں

## Analysis

## Core semantic meanings of Urdu verb lagnā

Corpus evidence shows that there are twelve core semantic meanings of the Urdu verb *lagnā*. According to FrameNet<sup>12</sup> the twelve frames that use of verb *lagnā* evokes, are presented with their description and core frame element (CFE) in this section.

The verb *lagnā* evokes the frame of ATTACH<sup>13</sup> as ATTACH\_physical and ATTACH\_emotional. Frame ATTACH always occurs in Noun + lagnā (with various tense forms). It is ATTACH\_physical when CFE Agent<sup>14</sup> physically connects one thing to another, or an entity is physically connected to another thing as in (3). ATTACH\_physical also includes

<sup>12</sup> <https://framenet.icsi.berkeley.edu/fndrupal/>

<sup>13</sup> All frames are written in capital in this paper

<sup>14</sup> All Core FEs are written with a capital first-letter

the sense of an entity joining some organisation for job as in (4). Some common elements in the frame are *phūl* ‘flower’, *istēhār* ‘advertisement’, *tālā* ‘lock’, *darwāza* ‘gate’, *niśāna* ‘target’, *jagha* ‘place’.

3. Tālāb=mēm                      **filter**      **laggē**                      hūē  
     pond=in                      **filter**      **attach**.IPFV.M.PL.OBL                      be.PFV.M.PL  
     haim  
     be.PRS.3.PL  
     “The pond has filters” (34\_79133).
4. Behrqēf                      ūskī                      **nōkarī**                      **lag**                      gāī  
     Anyway                      his                      **job**                      **attach**                      go.PFV.F.SG  
     “Anyway he got a job” (26\_70079).

ATTACH\_emotional frame is evoked when CFE Experiencer emotionally attaches themselves to people, places, or things. The elements in the frame are *jī/dil/mann* ‘heart’, *caskā* ‘addiction’.

5. Iqrā=kā                      **dil**                      **lag**                      gēā  
     Iqra=GEN.M.SG                      **heart**                      **attach**                      go.PFV.M.SG  
     pūrī                      tarah                      yahān  
     completely                      way                      here  
     “Iqra has adjusted here very well” (44\_23507).

Three common patterns of ATTACH\_emotional frame are *lag gēā* ‘got attached’, *dil lag līā* ‘(consciously made effort) to get adjusted’, and *lag hī jāē gā* ‘will eventually get attached’, as in example (5).

The verb *lagnā* is used to evoke the frame of BEGIN where CFEs of BEGIN including all those action verbs that are used when a process, event, or activity is started at a particular time and place as in (6).

6. cāī                      **pīnē**                      **laggā**                      hūm  
     Tea                      **drink**.INF.OBL                      **begin**.IPFV.M.SG                      be.PRS.1.SG  
     “[I] am about to drink tea” (19\_7974).

Frame BEGIN is evoked when any kind of action such as *qahqahā lagayā* ‘started to laugh’, *jūstajū mēm lag gēā* ‘begin to investigate’, *parhanē lagī hī thī* ‘was about to begin reading’.

The verb *lagnā* is used to evoke the frame of CATCH\_FIRE. It has a very specific element in its frame, that is, fire. It is evoked when something catches fire. This phrase is used both literally and metaphorically. For instance, in example (7) it is a literal fire where a physical thing is put to fire deliberately or accidentally. Metaphorically, in Urdu, it is usually used in the sense of motivation, love, hate, jealousy, inflation, or lawlessness. Sometimes, these terms are mentioned in the sentence along with fire as in (8) but often these feelings can be perceived from the context as in (9).

7. Bārūd=ko                      **āg**                      **lag**                      gāī  
     Dynamite=DAT                      **fire**                      **catch**                      go.PFV.F.SG  
     “Dynamite caught fire” (00\_11784).
8. Vō                      dil=mēm                      **āg**                      **lagānē=kī**                      bāt

- They heart=in fire catch.INF.OBL=GEN.F.SG talk  
kartē haim  
do.IPFV.M.PL be.PRS.3.PL  
“He talks about kindling the fire in the heart” (34\_95781)
9. Mērī zātī sōc jō ke main  
1.SG.DIR personal thought.PL REL COMP 1.SG  
bēyān kar dūm tō buhat sorh=kō  
state do give.SBJV.1.SG CONJ many of=ACC  
āg lag jātī  
fire catch go.IPFV.F.SG  
“My personal thoughts which if I express then it would offend/anger many”  
(01\_2666).

The verb *lagnā* evokes the frame of ENGROSS with Experiencer or Agent as CFEs. The CFE is Experiencer, the subject in an NP of a sentence, who has a stimulus that brings about a particular experience/emotion causing Experiencer to focus on or that the Experiencer is engaged in some activity as in (10). Whereas the subject of a sentence is an Agent attempts to attain an explicitly mentioned goal as in (11). In case the sentence is passive, it may not explicitly mention the Agent but he can be inferred from the context as in (12). Common elements are *kām* ‘work/job’, *bātēm* ‘talk/gossips’, *dēnē* ‘to give’, *banaānē* ‘to make’. A common pattern is *NOUN+ lagā hūā* VERB ‘NOUN am/is/are/was/were busy/engaged in VERB, *NOUN+ lagā hai/haim/thā/thē/thī/thīm* ‘NOUN busy/engage AUX’

10. Sab auratēm apnē kām=mēm laggī  
All woman.F.PL they.OBL do=in busy.PFV.F.PL  
thēm  
be.PST.F.PL  
“All [Experiencer: women] were busy in their work” (00\_65914).
11. Main is=kē bajāē is kośī=mēm lug  
1.SG DEM=GEN.M.PL instead DEM effort=in busy  
gāi ke kisī tarah diprēsān=sē  
PFV.F.SG COMP some way depression=from  
chuṭkāṛā hāsil karūm  
rid.of get do.SBJV.3SG  
“Instead of that [Agent: I] started trying to get rid of depression” (01\_52579).
12. Kuch arjant=kē ārdar thē un=kō fāinal  
Some urgent=GEN.M.PL order.PL be.PST.3.PL DEM=DAT final  
karne=mēm laggā huā thā  
do.INF.OBL=in apply.IPFV.M.SG be.PP.M.SG be.PST.3.SG  
“‘There were some urgent orders, [Agent: I] was busy in completing them”  
(03\_100651).

*Lagnā* evokes the frame of EMOTION\_of\_mental\_activity where CFE Experiencer undergoes some kind of emotion or emotional state as induced by some form of Stimulus as in (13). The stimulus can be another person, circumstance/event, or a state of affair that evokes a certain emotional reaction in the Experiencer. These emotions are psychological feelings but may be reflected in words or expressions of the Experiencer. Common elements used to express EMOTION\_of\_mental\_activity are nouns such as *gussa* ‘anger’, *khuśī* ‘happiness’, *burrā* ‘bad’, *achā* ‘good’, *afsos* ‘sadness’, *dukh* ‘sadness’ *dar* ‘fear’ which when occurring with *lagā* read as psychological experience.

13. Pehlē                      mujhē                      tumhārī                      bātōm=par                      ḡussa  
 First.OBL                      1.SG.ACC                      2.SG.DIR                      talk.PL=on                      anger  
 ātā                      thā                      phir                      yūm                      hūā                      ke  
 come.IPFV.M.SG                      be.PST.M.SG                      then                      so                      be.PFV.M.SG                      COMP  
**hansī**                      ānē                      **laggī**                      phir                      afsos  
**laughter**                      come.INF.OBL                      **feel**.PFV.F.SG                      then                      sorry  
 hōnē                      **laggā**                      phir                      khāmōśī=kō  
 be.INF.OBL                      **feel**.PFV.F.SG                      then                      quiet=ACC  
 fōqīyat                      dī  
 prefer                      give.PFV.F.SG  
 “At first I used to feel anger at your words, then it so happened that I felt your talk  
 laughable, then I started to feel sorry [for you], then I preferred to keep quiet”  
 (29\_40311).

The verb *lagnā* can evoke two types of PERCEPTION frames: PERCEPTION\_active and PERCEPTION\_evaluation. *Lagnā* evokes the frame of PERCEPTION\_active where CFE Perceiver or Agentive performs some sort of activity to have a perceptual experience as in (14). Typical elements that occur in these frames are *dēkhnā* ‘to watch/look’, *sūngnā* ‘to smell’, *cakhnā* ‘to taste’, *sūnanā* ‘to listen’. Typically, *lagnā* occurs in an oblique form of the verb (i.e., *lagnē*) followed by auxiliary inflected for the gender and numbers, that is, *lagī* (F.SG), *lagā* (M.SG), *lagē* (M.PL), and *lagīm* (F.PL)

14. Ham                      ūs=par                      kharē                      hō                      kar  
 1.PL                      DEM=on                      stand.PFV.M.PL                      be.PRS.1.PL                      do  
 tamāśa                      **dēkhnē**                      **laggē**  
 performance                      **watch**.INF.OBL                      **start**.PFV.1.PL  
 “We stood on it to watch the performance” (24\_71844).

*Lagnā* evokes the frame of PERCEPTION\_evaluation in which CFE Perceiver typically undergoes an emotional state that is accounted as an internal experiential state. Most commonly *aēsā* ‘such’ with *lagnā* to give the meaning of ‘it seems that’ or *lagnā* is used in imperfective form *lagtā* followed by *hai/thā* ‘be’ and also sometimes *ke* (that-complementiser) meaning ‘seems/seemed or seems/seemed that’. Typically, constructions *lagtā/ lagī hai/thā/thī* ‘it seems/seemed’ and *lagtā/ lagī hai/thā/thī ke* ‘it seems/seemed that’ give an epistemic possibility reading that is, these constructions also convey a speaker’s judgement about the truth value of the statement or drawing an inference from the given circumstances as in (15) and (16).

15. Fahīm                      tō                      **lagtā**                      **hai**                      rāt=kā  
 Faheem                      EMPH                      **seem**.IPFV.M.SG                      **be**.PRS.3.SG                      night=GEN.M.SG  
 idhar                      hī                      hai  
 here                      EXC                      be.PRS.3.SG  
 “It seems Faheem has been here since night” (00\_1727).
16. Kum-sē-kum                      dō                      dīn=kā                      khēl                      mukamal  
 Less-from-less                      two                      day=GEN.M.SG                      game                      complete  
 hōnē=kē                      bād                      tō                      aēsā                      hī                      **lagtā**  
 be.INF.OBL= GEN.M.PL                      after                      EMPH                      such.that                      EXC                      **seem**.IPFV.M.SG  
 hai ke                      Garēm                      Ismith=kā                      Emrān                      Ṭāhir=kō  
 be.PRS.3.SG COMP                      Graham                      Smith=GEN.M.SG                      Imran                      Tahir=ACC  
 na                      khilānē=kā                      faēsā                      thīk                      nahīm                      thā.  
 NEG                      play.OBL=GEN.M.SG                      decision                      right                      NEG                      be.PST.3.SG

“At least after completion of the two-day game, it seems that Graham Smith’s decision to not let Imran Tahir play was not a good decision” (01\_34200).

The verb *lagnā* is used to evoke the frame of PHYSICAL which can also be divided into two parts: PHYSICAL\_sensation and PHYSICAL\_action. PHYSICAL\_sensation includes those bodily or psychological sensations that are experienced by Perceiver. These include *bhūkh* ‘hunger’, *pēās* ‘thirst’, *dard* ‘pain’, *thand/thandā* ‘cold’, *garam* ‘hot’, *garmī* ‘heat’.

17. **Bhōkh**                      **laggī**                      hō    tō                      ghussā                      buhat  
**Hunger**                      **feel.IPFV.F.SG**                      be    CONJ                      anger                      much  
āṭā    hai  
come   be.PRS.3.SG  
“If you feel hungry then you get very angry” (03\_110782).

PHYSICAL\_action includes those physical actions that occur when CFE Impactor makes a sudden and usually forceful physical contact with an Impactee. This also includes bodily harm such as *cōṭ* injury that are experienced by the sentient. The elements that cooccur with *lagnā* include *thapaṛ* ‘slap’, *capēr* ‘slap’, *mukkā* ‘punch’, *jūtē* ‘shoes’, *dakkā* ‘shove’.

18. farāns=mēm    ēk    aurat=ne    pīcce    bhāg    kar  
France=in    one    woman=ERG    after    run    do  
mard=kō    **thapaṛ**    **laggāyā**  
man=ACC    **slap**    **hit.PFV.M.SG**  
“In France, a woman ran back and **slapped** a man” (43\_83055).

It also includes metaphorical use of *dil kō dakkā/ghūnsa lagā* ‘it pained/shocked my mind’ and *(dil) kō lag gēā/gēī* ‘it affected/impacted my thoughts’.

The verb *lagnā* is used to evoke the frame of QUEUE. Semantic frame QUEUE is not part of any existing frames on FrameNet. However, as this is a recurring frame in this Urdu data, we have added it. The CFEs of QUEUE in Urdu are *lāin* ‘line’ or *qaṭār* ‘queue’ that sentient entity makes by standing in an ordered sequence for an event. Some common core elements are tickets, bill, fee, customer. In Urdu, this frame can be evoked when *lagnā* is used after core elements occur in a noun phrase (NP) and *lag* follows a postpositional *mēm* ‘in’ such as in (19).

19. Maim=nē                      Sājīd                      bhāī=kō                      buzurgh  
1.SG=ERG                      Sajid                      bhai=ACC                      elderly  
samajhtē                      hūē                      ūnhēm                      ēk                      taraf  
consider.IPFV.M.PL                      be.PFV.M.SG                      him.ACC                      one                      side  
bēthanē=kā                      kahā                      aur                      khud                      jā  
sit.INF.OBL=GEN.M.SG                      tell.PFV.M.SG                      and                      self                      go.PFV.M.SG  
kar                      **qaṭār=mēm**                      **lag**                      gēā.  
do                      **line=in**                      **queue**                      PFV.M.SG  
“I told Sajid bhai, considering him an elderly, to sit on a side and went to stand in the queue myself” (03\_6188).

The verb *lagnā* evokes the sense of TOUCH where Impactor comes in contact physically with Impactor. The role of Impactor and Impactee (CFEs) is realised in a noun phrase (NP). Sentient entities such as pronouns (e.g., *mēm* ‘I’, *tum/āp* ‘you’), *aurat* ‘woman’, *sahib* mister, are common. Common patterns are of perfective form, such as *galē lag gē*

‘hugged’, *galē lagā lā* ‘embraced’, imperative *galē lagā lēm/lījē* ‘do hug’, or subjunctive *galē lagā lēm gāē* ‘would hug’.

20. Hum	aik	dusre=kē	<i>galē</i>	<i>lag</i>	gāē
We	one	other.INF.OBL=GEN.M.PL	<i>embrace</i>	<i>touch</i>	go.PFV.M.PL
“We <b>hugged</b> each other” (00_5762).					

## Results and discussion

Our findings show that all twelve frames of *lagnā* are used in Urdu but not necessarily in a uniform manner. This can be seen in the distribution of *lagnā* across the twelve frames as given in Table 3 which is followed by the graphic visualisation of the values in Figure 6.

**Table 3**

*Distribution of lagnā ‘attach/feel’ across various frames in CUUMC*

S.NO.	FRAME	NUMBERS	PERCENTAGE
1.	ATTACH_Emoional	2241	1.2%
2.	ATTACH_Physical	21852	11.6%
3.	BEGIN	82040	43.5%
4.	CATCH_FIRE	2862	1.5%
5.	EFFORT/ENGROSS	6284	3.3%
6.	EMOTION_of_mental_activity	5551	2.9%
7.	PERCEPTION_active	2577	1.4%
8.	PERCEPTION_Evaluation	7023	3.7%
9.	PHYSICAL_Action	2562	1.4%
10.	PHYSICAL_Sensation	2017	1.1%
11.	QUEUE	487	0.3%
12.	TOUCH	53191	28.2%
	<b>TOTAL</b>	<b>188687</b>	<b>100.0%</b>

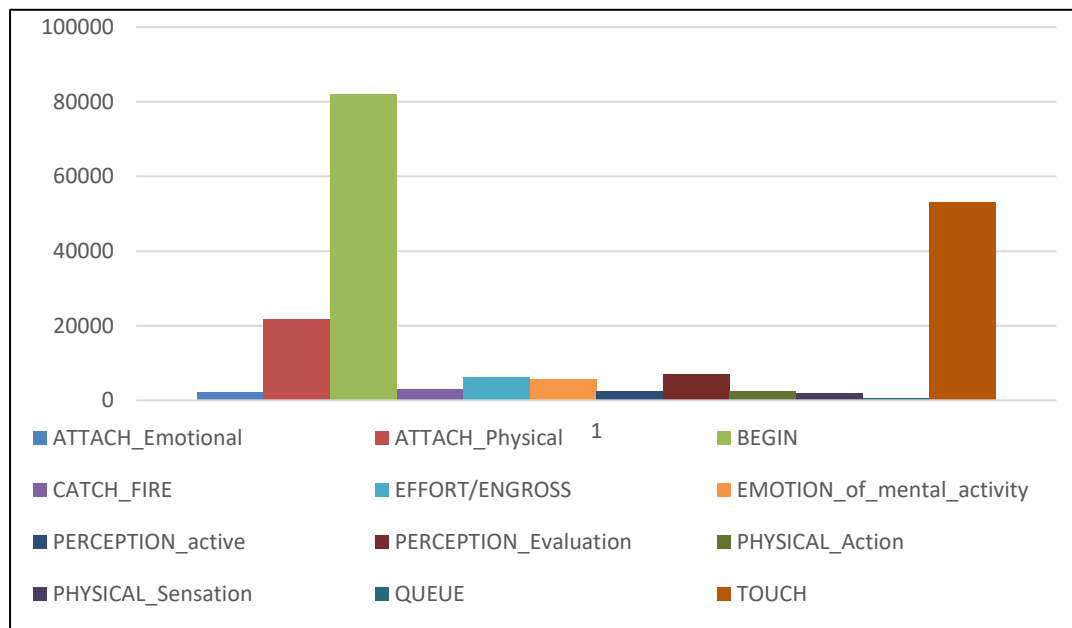
The most common frame is BEGIN (43.5%), followed by TOUCH (28.2%) and ATTACH\_Physical (11.6%). The rest of the frames are not as frequent as the first three, but they are regularly used to evoke a particular meaning in a given context. All these frames point out that in Urdu there are certain verbs such as *lagnā* that when used in certain construction may evoke a different meaning. Two broader constructions are of NOUN+ *lagnā* and VERB+ *lagnā*. These then get further divided into all the twelve categories. All three major frames are of VERB+ *lagnā* constructions. However, the use of the presumed prototypical meaning of verb *lagnā* as ‘attach’ cannot be justified. Therefore, it is the responsibility of the researchers to use a particular translation of the word that is closest to the sense *lagnā* conveys in the constructions it is occurring in and the lexical items it is co-occurring with.

Our study conforms with Hank’s (1996) viewpoint that the meaning of a verb such as *lagnā* can be understood in its totality in the patterns it occurs in. Moreover, this study also proves that the lexical items *lagnā* cooccurs with and the tense used in the construction, also influences the diverse contexts in which the meaning of *lagnā* is construed (Metayard & Vigliocco, 2018).



**Figure 6**

Percentage frequency of *lagnā* ‘attach/feel’ frames in CUUMC



Interestingly, the analysis shows that *lagnā* occurs in constructions to give literal, idiomatic as well as the metaphorical meaning of the sense created through those patterns. For instance, for construction NOUN [body part] + *lagnā* the construction may be literal if it is *pāōm lagā* ‘foot touched’, but idiomatic if it is *ānkh lagī* ‘lit. eye touched’ which means ‘to shut eye’ or ‘sleep’. Similarly, an extended context needs to be considered in some cases. Such as if the phrase is *par lag gaē* ‘got the wings’ it can have a metaphorical reference to high speed with which a piece of news or hearsay is spread (e.g., *is khabar kō par lag gaē haim* ‘the news spread very fast’), sudden rise in prices (e.g., *qīmatōm kō par lag gaē* ‘there is a meteoric rise in the prices’). The separation of phrases into literal and idiomatic is beyond the scope of the present study. We have accounted for the idiomatic phrase according to the frame it evokes.

### Conclusion

In this study, we have primarily followed Fillmore and his fellow researchers and used their FrameNet English frames in the field of frame semantics. We have analysed Urdu polysemous conjunct verb *lagnā* which in glossing is translated in English as ‘attach’. All in all, it is our finding that the meaning of *lagnā* as ‘attach’ is one of the predominant meanings, however, it is not the most frequent one. Therefore, it does not seem appropriate to consider it as a representative of all types of meanings that *lagnā* evokes when it occurs in various constructions. For each pattern that the verb *lagnā* occurs in, it expresses specific semantic information. This information is important especially for the researchers who use corpus-based Urdu examples and then gloss and translate them into English for their readers that correct sense of the verb *lagnā* is used to avoid any ambiguity for the reader who does not know the language.

This study answers the two questions proposed in the introductory section. In the analysis section, we have briefly described each of the twelve frames the verb *lagnā* creates. Each description includes CFEs, typical patterns, and at least one example. All these

descriptions are strictly corpus-based evaluations. The descriptions have also explicitly clarified that the use of the verb *lagnā* and the senses it evokes is context dependent. On its own, the conjunct verb *lagnā* in any of its forms may not be meaningful.

As pointed out in the discussion section, our study proves Stringer's (2019) analysis that speakers while using polysemous verbs do mentally process feature selection that results in specific verbs in particular contexts. This may be of significance in future research because the data analysis in this study shows that Urdu speakers are well aware of the contextual meaning verb *lagnā* evokes because it is regularly being used in both literal and metaphorical meanings.

It is mentioned in section 3, about the use of FrameNet (Fillmore et al, 1992; Fillmore & Baker, 2015) by researchers for research on languages other than English by aligning the semantic frames in their respective languages with original FrameNet frames in English. Similarly, for this study we have not used any translations, rather we glossed and translated original Urdu examples from natural language data into English to successfully apply the semantic frames provided on FrameNet without distorting the meaning in the source text (Urdu) or target text (English). This has significance for future studies on Urdu and other regional languages because the other researchers can also utilise FrameNet as a resource which has approximately 1200 semantic frames and 13,600 lexical units. Similar to the method we followed, other researchers can also create alignment of semantic frames evoked by their researched natural language examples with those provided on FrameNet. It can also help in identifying any additional semantic frames not mentioned on FrameNet or a frame that is specific only to Urdu (or any language researched). Semantic frame QUEUE is one such example that is not specifically mentioned in any of the semantic frames on FrameNet. There are examples on FrameNet of *standing in sequence*, *waiting in sequence* but *queuing* has not been added to those frames till the writing of this paper. Semantic frame QUEUE establishes the usefulness of corpus-based analysis because natural language data not only validates the use of language as authentic but also helps establishing frames specific to a language that may have gone unnoticed if researchers rely only on intuition or follow the already existing frames.

However, this study is very broad in its nature. This can be further refined in future research by taking one particular frame or some particular set of frames to analyse, for instance how apparently similar NOUN+ *lagnā* constructions may evoke different senses, both literal and idiomatic. Another topic for future research can be on verb *lagnā* used for the modal meaning of epistemic possibility (see examples 15 and 16 in this paper). Our aim is to introduce this strand of research in the corpus-based descriptive analysis of Urdu and other regional languages.

### Abbreviations

=	clitic boundary
1	first person
2	second person
3	third person
ACC	accusative
COMP	complementiser
CONJ	conjunction
DAT	dative
DEM	demonstrative
DIR	direct
EMPH	emphatic
ERG	ergative
EXC	exclusive particle <i>hī</i>
F	female
FUT	future
GEN	genitive
INF	infinitive
INS	instrumental
IPFV	imperfective
M	masculine
NEG	negation, negative
OBL	oblique
PFV	perfective
PL	plural
PRS	present
PROG	progressive
PST	past
SBJV	subjunctive
SG	singular

## References

- Abdullah, F., Ahmed, T., & Anjum, U. (2021). Urdu Conjunct Predicates (N+ V) Inventory from Urdu Universal Dependency Corpus. *Corporum: Journal of Corpus Linguistics*, 4(1), 15-29.  
<https://journals.au.edu.pk/ojsrcr/index.php/crc/issue/view/7/15>
- Ahmed, T. (2010, November). The interaction of light verbs and verb classes of Urdu. In *Proceedings of Interdisciplinary Workshop on Verbs - The Identification and Representation of Verb Features*, Pisa 2010.  
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.727.635&rep=rep1&type=pdf>
- Boas, H. (2020). A roadmap towards determining the universal status of semantic frames. In R. Enghels, B. Defrancq, & M. Jansegers (Eds.), *New approaches to contrastive linguistics: Empirical and methodological challenges* (pp. 21-52). De Gruyter Mouton. <https://doi.org/10.1515/9783110682588-002>
- Butt, M., & Geuder, W. (2011). Light verbs in Urdu and grammaticalization. In R. Eckardt, K. Heusinger & C. Schwarze (Eds.), *Words in Time: Diachronic Semantics from Different Points of View* (pp. 295-350). De Gruyter Mouton.  
<https://doi.org/10.1515/9783110899979.295>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407. <https://doi.org/10.1037/0033-295X.82.6.407>
- Davison, A. (2014). Non-finite complements and modality in de-na ‘allow’ in Hindi-Urdu. *Natural Language & Linguistic Theory*, 32(1), 137-164.  
<https://doi.org/10.1007/s11049-013-9224-6>
- Fazalehaq, H. M. (2019). *Identification of formulaic sequences in Urdu language and their pedagogical implication for SLA (ESL/USL)* (Doctoral dissertation). The University of New Mexico. [https://digitalrepository.unm.edu/educ\\_llss\\_etds/117](https://digitalrepository.unm.edu/educ_llss_etds/117)
- Fillmore, C. J. (1985). Frames and the semantics of understanding. *Quaderni di Semantica*, 6(2), 222-254. <http://www.icsi.berkeley.edu/pubs/ai/framesand85.pdf>
- Fillmore, C. J., & Atkins, B. T. S. (1992). Toward a frame-based lexicon: The semantics of RISK and its neighbors. In A. Lehrer, & E. Kittay (Eds.), *Frames, fields and contrasts: New essays in semantic and lexical organization* (pp. 75- 102). Lawrence Erlbaum Associates. <http://www.icsi.berkeley.edu/pubs/ai/towarda92.pdf>
- Fillmore, C. J., & Baker, C. (2015). A frames approach to semantic analysis. In B. Heine, & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (pp. 791-816). Oxford University Press. <http://doi.org/10.1093/oxfordhb/9780199544004.001.0001>

- Gilardi, L., & Baker, C. (2018, May). Learning to align across languages: Toward multilingual FrameNet. In *Proceedings of the International FrameNet Workshop* (pp. 13-22). [http://lrec-conf.org/workshops/lrec2018/W5/pdf/11\\_W5.pdf](http://lrec-conf.org/workshops/lrec2018/W5/pdf/11_W5.pdf)
- Granger, S., & Paquot, M. (2009). In search of a general academic vocabulary: A corpus-driven study. *Linguistics (VESPA-LING)*, 58, 519-109. [shorturl.at/cvTYZ](http://shorturl.at/cvTYZ)
- Halliday, M. A. K., & Webster, J. J. (Eds.). (2009). *Bloomsbury companion to systemic functional linguistics*. Bloomsbury Publishing.
- Hanks, P. (2006). Metaphoricity is gradable. In A. Stefanowitsch & S. Gries (Eds.), *Corpus-based approaches to metaphor and metonymy* (pp. 17-35). Berlin, New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110199895.17>
- Hsiao, M. C., & Chung, S. F. (2018). Are they arguing or not: A corpus-based study. *Proceedings of the 32nd Pacific Asia Conference on Language, Information and Computation*. <https://aclanthology.org/Y18-1023.pdf>
- Jawaid, B., Kamran, A., & Bojar, O. (2014, May). A tagged corpus and a tagger for Urdu. *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)* (pp. 2938-2943). [http://www.lrec-conf.org/proceedings/lrec2014/pdf/544\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2014/pdf/544_Paper.pdf)
- Koul, O. N. (2008). *Modern Hindi grammar*. Dunwoody Press.
- Leech, G. (1992). Corpora and theories of linguistic performance. In J. Svartvik (Ed.), *Directions in Corpus Linguistics: Proceedings of the Nobel Symposium 82, Stockholm, 4-8 August 1991* (pp. 105–22). Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110867275.105>
- Liu, M., & Zhang, Z. (2021). Exploring the role of verb frames in assessing semantic difficulty. *ExLing 2021: Proceedings of 12<sup>th</sup> International Conference of Experimental Linguistics, 11-13 October 2021* (pp. 153-154). [shorturl.at/vzAJT](http://shorturl.at/vzAJT)
- Manetta, E. (2019). Verb-phrase ellipsis and complex predicates in Hindi-Urdu. *Natural Language & Linguistic Theory*, 37(3), 915-953. <https://doi.org/10.1007/s11049-018-9429-9>
- McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
- Meteyard, L., & Vigliocco, G. (2018). Lexico-semantics. In S-A. Rueschemeyer & M. G. Gaskell (Eds.), *The Oxford handbook of psycholinguistics* (pp. 71-90). <https://doi.org/10.1093/oxfordhb/9780198786825.013.4>
- Moon, R. (2007). Sinclair, lexicography, and the Cobuild Project: The application of theory. *International Journal of Corpus Linguistics*, 12(2), 159-181. <https://doi.org/10.1075/ijcl.12.2.05moo>
- Renouf, A., & Sinclair, J. (1991). “Collocational frameworks in English”. In K. Aijmer & B. Altenberg (Eds.), *English corpus linguistics* (pp. 128-143). Longman.

- Rundell, M., & Fox, G. (Eds.). (2002). *Macmillan English dictionary for advanced learners*. Macmillan Education
- Safari, A., & Rahmatinejad, L. (2018). Frame semantics: Applying FrameNet principles to “touring” and “travel” semantic frames in Persian. *Biannual Journal of Applications of Languages Studies*, 2017-2018.
- Sanacore, D., Hathout, N., & Namer, F. (2019, September). Semantic descriptions of French derivational relations in a families-and-paradigms framework. In *Proceedings of the Second International Workshop on Resources and Tools for Derivational Morphology* (pp. 15-24). <https://aclanthology.org/W19-8503.pdf>
- Sardinha, T. B. (2019). Lexicogrammar. In C. A. Chapelle (Ed.), *The encyclopaedia of applied linguistics*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781405198431.wbeal0698.pub2>
- Schmidt, R. L. (1999). *Urdu: An essential grammar*. Routledge.
- Sajjad, H., & Schmid, H. (2009). Tagging Urdu text with parts of speech: A tagger comparison. In *Proceedings of the 12<sup>th</sup> Conference of the European Chapter of the ACL (EACL 2009)*. <https://aclanthology.org/E09-1079.pdf>
- Shafique, H. (2018). A corpus-based study of metadiscourse markers in English and Urdu. In *ExLing 2018: Proceedings of 9th Tutorial and Research Workshop on Experimental Linguistics, 28-30 August, Paris, France* (pp. 101-104). <https://doi.org/10.36505/ExLing-2018/09/0024/000357>
- Shahid, M. I., Qasim, H. M., & Hasnain, M. (2020). A cross-linguistic study of metadiscourse in English and Urdu newspaper editorials. *Corporum: Journal of Corpus Linguistics*, 3(1), 33-56. <https://journals.au.edu.pk/ojsrcc/index.php/crc/article/view/90>
- Shaoul, C., & Westbury, C. (2010). Exploring lexical co-occurrence space using HiDEx. *Behavior Research Methods*, 42(2), 393-413. <https://doi.org/10.3758/BRM.42.2.393>
- Sinclair, J., & Sinclair, L. (1991). *Corpus, concordance, collocation*. Oxford University Press, USA.
- Stringer, D. (2019). Lexical semantics: Relativity and transfer. In N. Erdogan & M. Wei (Eds.), *Applied linguistics for teachers of culturally and linguistically diverse learners* (pp. 180- 203). IGI Global. <https://doi.org/10.4018/978-1-5225-8467-4.ch007>
- Tanveer, B., Akhter, S., Awais, H. M., & Qureshi, A. H. (2021). Improving speaking skills through dialogues in EFL classroom at higher secondary level. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 18(4), 2861-2872. [shorturl.at/oIQ16](https://shorturl.at/oIQ16)
- Verdaguer, I. (2020). Semantic frames and semantic networks in the Health Science Corpus. *Estudios de Lingüística Del Español*, (Anejo 1), 117-155. [https://infoling.org/elines/anejo-1-2020/anejo1\\_20-2.pdf](https://infoling.org/elines/anejo-1-2020/anejo1_20-2.pdf)

Viebahn, E., & Vetter, B. (2016). How many meanings for ‘may’? The case for modal polysemy. *Philosopher's Imprint*, 16(10), 1-26.  
<http://hdl.handle.net/2027/spo.3521354.0016.010>

Yaqub, H., & Shakir, A. (2019). A corpus-based study of hypotactic and paratactic thematic relations in English and Urdu clause complexes. *International Journal of English Linguistics*, 9(5), 430-448. <https://doi.org/10.5539/ijel.v9n5p430>