**ISSN Online:** 3006-4708
**ISSN Print:** 3006-4694

*SOCIAL SCIENCE REVIEW ARCHIVES*
https://policyjournalofms.com

# Influence of Urdu Language While Learning English Formulaic Expressions

**Farooq Ahmed[1], Zahid Yousaf[2], Saba Chaman[3], Saba Iram[4]**

[1] Lecturer, Department of English, University of Bhimber. Email: farooq@ajkuob.edu.pk
[2] Lecturer, Department of English, University of Bhimber. Email: zahid.usaf@gmail.com
[3] Lecturer, Department of English, University of Bhimber. Email: saba@ajkuob.edu.pk
[4] Department of English, University of Bhimber. Email: sabairum79@gmail.com

**DOI: https://doi.org/10.70670/sra.v3i4.1264**

## Abstract
Formulaic expressions play a vital role in enhancing language fluency and communicative competence. These expressions—such as idioms, proverbs, collocations, and phrasal verbs—are often language-specific and culturally embedded, posing challenges for second language (L2) learners. This study investigates the extent to which Urdu, as a first language (L1), influences the acquisition of English formulaic expressions among L2 learners. The research was conducted with a sample of 150 students from selected secondary and higher secondary institutions in Bhimber, Azad Kashmir. Data were collected using structured questionnaires and analyzed through a categorization framework that included: (1) English-specific idioms and proverbs, (2) conceptually similar English equivalents of Urdu expressions, and (3) word-for-word English translations of Urdu expressions. Findings reveal a significant influence of Urdu on the learning and use of English formulaic expressions. Learners demonstrated varied levels of understanding across the categories, with notable interference from their native language, especially when relying on literal translations. The results suggest that Urdu-speaking L2 learners face substantial challenges in acquiring native-like fluency in English formulaic language. These findings highlight the need for targeted pedagogical strategies that explicitly address formulaic language acquisition in L2 classrooms. Future research should explore instructional approaches to mitigate L1 interference and enhance learners' exposure to authentic English usage.

**Keywords:** Formulaic Expressions, L1 Interference, Urdu Influence, Second Language Acquisition, Idioms, Proverbs, Collocations, Phrasal Verbs

## Introduction
Formulaic language refers to fixed, often non-literal expressions that are commonly used in communication, including idioms, proverbs, collocations, and phrasal verbs. These expressions are stored and retrieved holistically from the mental lexicon and play a vital role in enhancing fluency, pragmatics, and naturalness in speech (Wray, 2008; Conklin & Schmitt, 2012). Native speakers acquire these expressions effortlessly through repeated exposure, but second language (L2) learners often struggle due to limited input and interference from their first language (L1).
Research shows that formulaic language contributes to faster processing, greater fluency, and improved communicative competence (Boers et al., 2006; Ortactepe, 2013). However, for L2 learners—particularly those with typologically and culturally distant L1s such as Urdu—the acquisition of such expressions is more complex. The metaphorical and culture-specific nature of many formulaic expressions further complicates the learning process for non-native speakers (Conklin

& Schmitt, 2012; Hama, 2010).

The role of L1 in acquiring L2 formulaic expressions has been widely discussed. While some studies emphasize the usefulness of recognizing similarities between L1 and L2 formulaic forms (e.g., Yamashita & Jiang, 2010), others argue that such comparisons may lead to errors due to non-congruencies (Bahns & Eldaw, 1993; Nesselhauf, 2003). Consequently, L2 learners may resort to direct translation, resulting in awkward or incorrect usage.

This study investigates the extent to which Urdu, as an L1, influences the acquisition of English formulaic expressions among secondary and higher secondary students in Bhimber, Azad Kashmir. By examining learners' understanding of idioms, proverbs, collocations, and phrasal verbs across three categories—English-specific expressions, conceptually similar expressions, and literal translations—this study seeks to provide insights into the challenges faced by Urdu-speaking learners and the pedagogical implications for English language instruction.

Statement of the Problem

Despite growing awareness of the importance of formulaic language in achieving fluency, limited research has examined how Urdu influences the learning of English formulaic expressions. Existing studies present conflicting views: some suggest that cross-linguistic similarities support learning, while others highlight the problems caused by L1 interference. However, few studies have explored this issue specifically in the context of Urdu and English.

Given the increasing emphasis on native-like fluency in L2 classrooms and the widespread use of formulaic expressions in natural communication, it is essential to investigate how L1 Urdu affects learners' ability to comprehend and produce English formulaic language. This study addresses this gap by evaluating the influence of L1 on the comprehension of different types of English formulaic expressions and provides recommendations for teaching strategies that can mitigate the effects of negative L1 transfer.

## Literature Review

Formulaic expressions are fixed sequences of words that are commonly used and recognized as a single unit play a pivotal role in achieving fluency and naturalness in any language. These expressions include idioms, collocations, proverbs, phrasal verbs, and other multi-word expressions that often carry meanings not predictable from their individual components (Wray, 2002; Schmitt, 2010). In second language acquisition (SLA), the ability to recognize, understand, and appropriately use formulaic language is regarded as a marker of advanced language competence (Pawley & Syder, 1983; Nation, 2001).

### Formulaic Language in Second Language Acquisition

Pawley and Syder (1983) were among the first to argue that native speakers' fluency is largely due to their ability to draw on a vast repertoire of formulaic sequences. These prefabricated chunks reduce the processing load and facilitate fluent speech. Subsequent studies (e.g., Nattinger & DeCarrico, 1992; Wood, 2009) have reiterated that formulaic language supports both language production and comprehension by enabling speakers to bypass the need for grammatical analysis of individual words in real-time communication.

For second language learners, however, the acquisition of formulaic expressions presents a distinct challenge. This is primarily because such expressions are often non-compositional and culturally embedded. Learners may encounter difficulties in both recognizing and using these expressions accurately. Research by Conklin and Schmitt (2012) has shown that native speakers process formulaic sequences faster than novel phrases, while non-native speakers typically do not exhibit the same processing speed or accuracy, highlighting a significant gap in formulaic competence.

### Typology of Formulaic Language

Wray (2002) provides a comprehensive typology of formulaic language, classifying it based on structural and functional criteria. Idioms, such as "spill the beans," and collocations, such as "strong

coffee," are structurally different but functionally similar in their fixedness and predictability. Proverbs like "A stitch in time saves nine" carry cultural and moral meanings that are often difficult to translate across languages. Phrasal verbs such as "give up" or "look into" are particularly challenging for learners from non-Germanic language backgrounds, including Urdu, due to syntactic and semantic opacity (Celce-Murcia & Larsen-Freeman, 1999).

Bardovi-Harlig (2012) highlights that phrasal verbs are particularly problematic for ESL learners because their meaning cannot be inferred directly from the individual words. Additionally, they often have multiple meanings depending on context (e.g., "take off" meaning to remove clothing or for a plane to depart). This underscores the need for context-rich learning approaches.

Theoretical Perspectives on Formulaic Acquisition

From a cognitive perspective, formulaic expressions are stored and retrieved as single units, making them easier to use under time constraints (Ellis, 2003; Schmitt & Carter, 2004). Usage-based theories of language acquisition suggest that exposure and frequency play a crucial role in the internalization of such expressions. Learners acquire formulaic language not through isolated instruction but through repeated exposure in meaningful contexts (Bybee, 2006).

Connectionist models (e.g., Ellis, 2006) argue that language learners form associations between words based on their co-occurrence frequencies. This view is supported by corpus-based research (Martinez & Schmitt, 2012), which has shown that learners who are frequently exposed to multi-word expressions in natural discourse tend to acquire them more effectively than those who rely solely on textbook examples.

**Challenges in Acquiring Formulaic Expressions**

Despite their importance, formulaic expressions are underrepresented in many ESL/EFL curricula. Researchers like Boers and Lindstromberg (2012) argue that traditional vocabulary instruction often ignores formulaicity, focusing instead on individual words and rules. This approach overlooks the fact that native-like fluency heavily depends on the ability to use fixed expressions appropriately.

Hama (2010) points out that learners' limited awareness and exposure to formulaic sequences, coupled with negative transfer from the L1, often lead to errors. Similarly, Yamashita and Jiang (2010) argue that formulaic language processing is mediated by L1-specific mechanisms, which may hinder the acquisition of L2 expressions that lack direct equivalents.

This is particularly true for learners of English from Urdu-speaking backgrounds, where the syntactic and semantic structures of idioms and proverbs often differ significantly from those in English. For example, the Urdu expression "naach na jaane aangan tedha" literally translates to "unable to dance, blaming the floor," which has a conceptual counterpart in English ("a bad workman blames his tools"), but the lexical components differ completely.

**L1 Influence on Formulaic Language Acquisition**

Cross-linguistic influence, particularly from the learner's L1, is one of the most influential factors in acquiring formulaic language in an L2. Kellerman (1979) introduced the concept of "transferability," suggesting that learners rely on their L1 structures to make inferences about the L2. When L1 and L2 share similar forms and meanings, positive transfer can facilitate learning. However, when differences exist—especially in idiomatic and figurative language—negative transfer can result in incorrect usage or misinterpretation.

Bahns and Eldaw (1993) conducted a study on collocation errors and found that learners often incorrectly assume that L1 collocations can be translated word-for-word into L2, resulting in unnatural or incorrect expressions. This tendency is particularly strong among learners of English from languages like Urdu, where direct translations often do not carry the intended idiomatic meaning in English.

Kecskes (2007) further emphasizes that L2 learners initially conceptualize meaning based on their L1 cognitive frameworks. Over time, with sufficient exposure and interaction, these frameworks may shift to incorporate L2-specific conceptualizations, but the initial stages of learning are often marked

by heavy L1 interference.

**Formulaic Expressions in the Context of Urdu and English**

Urdu and English belong to different language families—Indo-Aryan and Germanic, respectively—and exhibit considerable variation in syntax, morphology, and figurative language. Many English idioms and proverbs lack direct Urdu equivalents, and vice versa. For instance, the English idiom "kick the bucket" (to die) has no equivalent in Urdu, and if translated literally, it would be meaningless. Likewise, the Urdu idiom "munh mein paani bhar aana" (to salivate due to desire) may not be easily comprehensible to English speakers.

Studies focusing on South Asian learners of English have confirmed the difficulty learners face with idiomatic usage. Rahman (2002) analyzed English language textbooks used in Pakistan and concluded that idioms, phrasal verbs, and collocations are underrepresented. He also noted that the cultural content of idioms is rarely explained, leading to confusion among learners.

Mahmood, Shahzad, and Mahmood (2014) explored the errors made by Pakistani learners in using English idioms and found that most mistakes were due to literal translation from Urdu. Learners tend to assume that an Urdu phrase will retain its meaning when translated word-for-word into English, leading to semantic errors and communication breakdown.

**Teaching Formulaic Expressions: Approaches and Strategies**

To mitigate the challenges in acquiring formulaic expressions, researchers have advocated for a more focused and context-based pedagogical approach. Boers and Lindstromberg (2009) propose the use of metaphor awareness and etymological elaboration to help learners understand the conceptual underpinnings of idioms. For example, explaining the cultural and historical background of "kick the bucket" can make the expression more memorable.

Lewis (1993) advocates for a lexical approach in language teaching, emphasizing the teaching of chunks over isolated vocabulary items. This approach has been found particularly effective in helping learners notice and internalize formulaic sequences. Corpus-based tools such as concordancers can also aid learners in discovering how expressions are used in authentic contexts (Boulton, 2012).

In the Pakistani context, integrating formulaic expressions into classroom instruction remains a challenge due to outdated syllabi and insufficient teacher training. However, some recent initiatives have started incorporating more communicative and corpus-informed approaches to ESL instruction (Shamim & Rashid, 2019).

**Research Gap and Rationale for the Present Study**

While numerous studies have examined formulaic language acquisition in ESL learners globally, limited research exists on how Urdu, as a first language, influences the acquisition of English formulaic expressions in Pakistan—particularly in rural or semi-urban contexts like Bhimber, Azad Kashmir. Most prior research focuses on either grammar or vocabulary acquisition in general, often overlooking multi-word expressions or their unique challenges.

Furthermore, little is known about how students perceive, process, and apply formulaic language in real-life communication. The role of regional dialects, exposure to English media, and the sociolinguistic environment also remain underexplored. This study seeks to fill this gap by examining how Urdu-specific conceptualizations influence English idiom and proverb usage among secondary and higher secondary students in Bhimber. So the literature reviewed above underscores the complexity and importance of formulaic language in second language acquisition. It highlights the central role of L1 influence, the processing advantages of native speakers, and the pedagogical neglect of formulaic expressions in many ESL settings. For Urdu-speaking learners, the challenges are compounded by structural and cultural differences between Urdu and English. Understanding how learners navigate these challenges is essential for developing effective instructional strategies. This study aims to contribute to this understanding by examining the cross-linguistic influence of Urdu on the acquisition of English formulaic expressions in a specific sociolinguistic context.

## Research Methodology

The purpose of this study was to investigate the influence of Urdu language while learning English formulaic expressions namely, idioms, proverbs, collocations and phrasal verbs. The data was collected through the questionnaire method. The data was interpreted to address the following research questions;

1. How much does the Urdu influence while learning English formulaic expressions at secondary level?
2. How does the availability of an equivalent in L1 cause difficulty in learning the L2 formulaic language?

## Research paradigm

The research is based on mixed method research approach. The view point of pragmatists is playing a major role in the research paradigm. The research is combining the both, positivist and constructivist principles. The research tends to use both qualitative and quantitative methods to diagnose the research problem. Qualitative research involves collecting and analysing non-numerical data to understand the concepts, opinions, or experiences of people. While the quantitative research involves collecting and analysing numerical data for statistical analysis.

On the other hand mixed method research combines elements of both qualitative and quantitative research in order to answer the research questions. It help us to attain a complete picture than a standalone quantitative or qualitative study. Because it gives the benefits of both methods. It allows the researcher to gain a depth understanding of the research problem. Due to all these reasons the researcher is using a mixed method approach to gain the more reliable results of the research.

## Qualitative and Quantitative Aspects of the Study

The overall data for this study was collected through questionnaires. Non numerical data was collected to understand the concepts, and opinions of the students. For statistical analysis, the data was converted into Microsoft excel to make it quantitative.

## Population of the Study

The present study was conducted in four major, secondary and higher secondary institutions of Bhimber Azad Kashmir, including, the learners, read foundation, insight Model College and Army public school and college. Both male and female students participated in the research questionnaires. Those students participated in the research, whose native language was Urdu.

## Sample of the Study

As it is difficult to study the whole population, so considering both time and budget, the data was collected from 150 students. The data was collected from the students of matric and intermediate level. The reasoning behind choosing these classes is, formulaic language like idioms, proverbs etc., are the part of syllabus during matric and intermediate level. At this stage students learn to use the formulaic expression in their daily life. The total sample were 150.

## Data Collection Tool

Questionnaire method of research was used in the present research. The samples were collected through different questionnaires. The researcher selected the questionnaire method to provide ease for the respondents. It became easy for the students if they have to respond in less time. So, considering the level of students, easy and simple questions were used to analyse the current research.

## Questionnaire

A questionnaire is a type of research tool used to collect data from respondents for a survey or statistical analysis. It consists of a set of questions (or other forms of prompts). Typically, a research

questionnaire will have both closed-ended and open-ended questions. Long-term, open-ended inquiries provide the respondent the chance to go into more detail. The Statistical Society of London created the research questionnaire in 1838.

Questionnaire is important for this study because in comparison to certain other survey instruments, questionnaires are more advantageous, they are less expensive, do not need as much effort from the respondent as verbal or telephone surveys, and frequently include standardised answers that make it easy to gather data. Such standardised responses, however, could irritate users because they do not exactly reflect their expected responses. The requirement that respondents be able to read the questions and reply to them severely restricts the use of questionnaires. Therefore, conducting a survey through questionnaire may not be practically possible for some demographic groups.

The research questionnaire was consisted of two parts. The first part was about collocation and phrasal verbs of English language. The participants were asked to provide the equivalent of that words in their native language. These collocations and phrasal verbs were used in a paragraph to provide students with the context. The second part was about the English proverbs and idioms. The participants were given the three possible options for that statements in their native language. The three categories are

1. Correct
2. Interchangeable
3. Incorrect

The first category deals with the non-literal meanings while the second category deals with the literal meanings and the third category is just an incorrect option.

**Validity**

The validity of all the instruments of research was ensured by the opinions of different experts. Different samples from different institutions were taken to make our research more valid. The current research is providing accurate and good quality findings. Statistical techniques are used to analyse the data in a more effective way.

**Reliability**

One of the main purpose of this research is to provide the reliable results. The present findings are reliable and maintain stability of the results. A chi square test is a good statistical test to make the data reliable by comparing observed results with expected results. The data is analysed through this test.

**Data Analysis**

The study implemented two questionnaires to address these research questions. The participants were 150 students from the different institutions of Bhimber Azad Kashmir. The participants were given the questionnaires to assess their ability to produce the formulaic sequences of English into their mother tongue. The mother tongue counterparts were in three categories.

- Category 1, idioms and proverbs specific to the English language
- Category 2, conceptually similar English versions of the idioms and proverbs used in Urdu
- Category 3, word for word English translations of the idioms and proverbs used in Urdu

**Idioms**

In order to analyse the English idioms, the data was collected statistically. The total samples were 150. The overall data from the responses was gathered in the following form.

**Table 01: Use of English Idioms as a Formulaic Expression**

| Idioms | Correct | Interchangeable | Incorrect |
|---|---|---|---|
| **A bird's eye view** | 57 | 73 | 20 |
| **To pocket an insult** | 27 | 66 | 57 |
| **Kick the bucket** | 36 | 92 | 22 |

| | | | |
|---|---|---|---|
| **Bear in mind** | 85 | 46 | 19 |
| **To poison one's ear** | 49 | 88 | 13 |
| **Add insult to injury** | 61 | 21 | 68 |
| **Hit the nail on the head** | 13 | 85 | 52 |
| **To fight shy of** | 36 | 71 | 43 |
| **Total** | 364 | 542 | 294 |
| **Mean** | 45.5 | 67.75 | 36.75 |
| **StD** | 22.43 | 23.89 | 20.82 |

Table 01 suggests that the Urdu-speaking English formulaic expression learners used the idiom in different ways.

First category deals with the Urdu idioms specific to the English language.

Second category is conceptually similar Urdu version of the English idioms.

Third category is linked with word for word Urdu translation of the idioms.

The mean and standard deviations of the gain scores between all the three categories is mentioned above in table 4.1. As shown in table, the sample shows the greatest mean score with the category 2 idioms, followed by category 1, and finally category 3.

**Category 1**

Category 1, idioms and proverbs specific to the English language

**Correct**

**Table 02: Application of Chi Square Test on Category 1**

| O | **E** | **(O-E)** | **(O-E)2** | **(O-E)2/E** |
|---|---|---|---|---|
| 57 | 45.5 | 11.5 | 132.25 | 2.9 |
| 27 | 45.5 | -18.5 | 342.25 | 7.52 |
| 36 | 45.5 | -9.5 | 90.25 | 0.2 |
| 85 | 45.5 | 39.5 | 1560.25 | 34.29 |
| 49 | 45.5 | 3.5 | 12.25 | 0.269 |
| 61 | 45.5 | 15.5 | 240.25 | 5.28 |
| 13 | 45.5 | -32.5 | 1056.25 | 23.219 |
| 36 | 45.5 | -9.5 | 90.25 | 1.98 |
| | | | | 75.658 |

In table 02, *O* stands for observed value, while *E* stands for expected value. After applying chi square test, the value of X2 is 75.658.

*X2 calculated = 75.658*

After finding the value of calculated X2, now we will compare the calculated and tabular value of X2.

*X2 tabular:* to find the tabular value of X2, we have to check the degree of freedom,

*Degrees of freedom*= (*total rows-1*) (total columns-*1*)

(8-*1*) (3-*1*)

(7) (2)

=014

(0.05) (14) = 23.68

*X2 tabular = 23.68*

*Calculated value > tabular value*

As the value of calculated x2 is greater than the tabular value of x2, null hypothesis is rejected and the alternative hypothesis is accepted here. It means that the idioms depend upon these three variables.

**Category 2**

Category 2, conceptually similar English versions of the idioms and proverbs used in Urdu

**Interchangeable**

**Table 03: Application of chi square test on category 2**

| O | E | (O-E) | (O-E)2 | (O-E)2/E |
|---|---|---|---|---|
| **73** | 67.75 | 5.25 | 27.5625 | 0.406 |
| **66** | 67.75 | -1.75 | 3.0625 | 0.045 |
| **92** | 67.75 | 24.25 | 588.0625 | 8.6798 |
| **46** | 67.75 | -21.75 | 473.0625 | 6.982 |
| **88** | 67.75 | 20.25 | 410.0625 | 6.052 |
| **21** | 67.75 | -46.75 | 2185.56 | 32.259 |
| **85** | 67.75 | 17.28 | 298.5984 | 4.4 |
| **71** | 67.75 | 3.25 | 10.5625 | 0.1559 |
| | | | | 58.9797 |

In table 03, *O* stands for observed value, while *E* stands for expected value. After applying chi square test, the value of X2 is 75.658.

*X2 calculated = 58.9797*

After finding the value of calculated X2, now we will compare the calculated and tabular value of X2.

*X2 tabular:* to find the tabular value of X2, we have to check the degree of freedom,

*X2 tabular*

*Degrees of freedom*= (*total rows-1*) (total columns-*1*)

(8-*1*) (3-*1*)

(7) (2)

=014

(0.05) (14) = 23.68

*X2 tabular = 23.68*

*Calculated value > tabular value*

As the value of calculated x2 is greater than the tabular value of x2, null hypothesis is rejected and the alternative hypothesis is accepted here. It means that the idioms depend upon these three variables.

**Category 3**

Category 3, word for word English translations of the idioms and proverbs used in Urdu

**Incorrect**

**Table 04: Application of chi square test on category 3**

| O | E | (O-E) | (O-E)2 | (O-E)2/E |
|---|---|---|---|---|
| **20** | 36.75 | -16.75 | 280.5625 | 7.6343 |
| **57** | 36.75 | 20.25 | 410.0625 | 11.158 |
| **22** | 36.75 | -14.75 | 217.5625 | 5.92 |
| **19** | 36.75 | -17.75 | 315.0625 | 8.573 |
| **13** | 36.75 | -23.72 | 562.6384 | 15.309 |
| **68** | 36.75 | 31.25 | 976.5625 | 26.573 |
| **52** | 36.75 | 15.28 | 233.4784 | 6.353 |
| **43** | 36.75 | 6.25 | 39.0625 | 1.0637 |
| | | | | 82.577 |

In table 4.4, *O* stands for observed value, while *E* stands for expected value. After applying chi square test, the value of X2 is 82.577.

*X2 calculated = 82.577*

After finding the value of calculated X2, now we will compare the calculated and tabular value of X2.

*X2 tabular:* to find the tabular value of X2, we have to check the degree of freedom,

*Degrees of freedom*= (*total rows-1*) (total columns-*1*)

(8-*1*) (3-*1*)

(7) (2)
=014

(0.05) (14) = 23.68

*X2 tabular = 23.68*

*Calculated value > tabular value*

As the value of calculated x2 is greater than the tabular value of x2, null hypothesis is rejected and the alternative hypothesis is accepted here. It means that the idioms depend upon these three variables.

**Proverbs**

In order to analyse the English proverbs, the data was collected statistically. The total samples were 150. The overall data from the responses was gathered in the following form.

**Table 05: Use of English proverbs as a formulaic expressions**

| Proverbs | Correct | Interchangeable | Incorrect |
|---|---|---|---|
| **Barking dogs seldom bite.** | 25 | 111 | 14 |
| **Once bitten twice shy.** | 46 | 68 | 36 |
| **A wolf in sheep's clothing.** | 21 | 52 | 77 |
| **Wolves may lose their teeth but not their nature.** | 58 | 71 | 21 |
| **A black sheep spoils the whole flock.** | 93 | 39 | 18 |
| **The nearer the church, the farther from heaven.** | 50 | 43 | 57 |
| **Let us see in which way the wind blows.** | 9 | 102 | 39 |
| **Beauty requires no paint.** | 54 | 83 | 13 |
| **Total** | 356 | 569 | 275 |
| **Mean** | 44.5 | 31.12 | 34.37 |
| **StD** | 24.62 | 24.7 | 21.42 |

As shown from the table 05, there are also three categories for analysing the use of English proverbs.
First category deals with the Urdu proverbs specific to the English language.
Second category is conceptually similar Urdu version of the English proverbs.
Third category is linked with word for word Urdu translation of the proverbs.
The mean and standard deviations of the gain scores between all the three categories is mentioned above in table 4.5. As shown in table, the sample shows the greatest mean score with the category 1 proverbs, followed by category 3, and finally category 2.

**Category 1**

First category deals with the Urdu proverbs specific to the English language.

**Correct**

**Table 06: Application of chi square test on category 1**

| O | E | (O-E) | (O-E)2 | (O-E)2/E |
|---|---|---|---|---|
| **25** | 44.5 | -19.5 | 380.25 | 8.54 |
| **46** | 44.5 | 1.5 | 2.25 | 0.05 |
| **21** | 44.5 | -23.5 | 552.25 | 12.41 |
| **58** | 44.5 | 13.5 | 182.25 | 4.095 |
| **93** | 44.5 | 48.5 | 2352.25 | 52.859 |
| **50** | 44.5 | 5.5 | 30.25 | 0.679 |
| **9** | 44.5 | -35.5 | 1260.25 | 28.32 |
| **54** | 44.5 | 9.5 | 90.25 | 2.028 |
| | | | | 104.89 |

In table 4.6, *O* stands for observed value, while *E* stands for expected value. After applying chi square test, the value of X2 is 104.89.

*X2 calculated = 104.89*

After finding the value of calculated X2, now we will compare the calculated and tabular value of X2.
*X2 tabular:* to find the tabular value of X2, we have to check the degree of freedom,
*Degrees of freedom*= (*total rows-1*) (total columns-*1*)
(8-*1*) (3-*1*)
(7) (2)
=014
(0.05) (14) = 23.68
*X2 tabular = 23.68*
*Calculated value > tabular value*
As the value of calculated x2 is greater than the tabular value of x2, null hypothesis is rejected and the alternative hypothesis is accepted here. It means that the proverbs depend upon these three variables.

**Category 2**
Second category is conceptually similar Urdu version of the English proverbs.
**Interchangeable**
**Table 07: Application of chi square test on category 2**

| O | E | (O-E) | (O-E)2 | (O-E)2/E |
|---|---|---|---|---|
| **111** | 71.125 | 39.875 | 1590.01 | 22.35 |
| **68** | 71.125 | -3.125 | 9.76 | 0.137 |
| **52** | 71.125 | -19.125 | 365.76 | 5.14 |
| **71** | 71.125 | -0.125 | 0.015 | 0.0002 |
| **39** | 71.125 | -32.125 | 1032.01 | 14.509 |
| **43** | 71.125 | -28.125 | 791.015 | 11.12 |
| **102** | 71.125 | 30.875 | 953.26 | 13.4 |
| **83** | 71.125 | 11.875 | 141.015 | 1.98 |
| | | | | 68.64 |

In table 4.7, *O* stands for observed value, while *E* stands for expected value. After applying chi square test, the value of X2 is 68.64.
*X2 calculated = 68.64*
After finding the value of calculated X2, now we will compare the calculated and tabular value of X2.
*X2 tabular:* to find the tabular value of X2, we have to check the degree of freedom,
*Degrees of freedom*= (*total rows-1*) (total columns-*1*)
(8-*1*) (3-*1*)
(7) (2)
=014
(0.05) (14) = 23.68
*X2 tabular = 23.68*
*Calculated value > tabular value*
As the value of calculated x2 is greater than the tabular value of x2, null hypothesis is rejected and the alternative hypothesis is accepted here. It means that the proverbs depend upon these three variables.

**Category 3**
Third category is linked with word for word Urdu translation of the proverbs.
**Incorrect**
**Table 08 : Application of Chi Square Test on Category 3**

| O | E | (O-E) | (O-E)2 | (O-E)2/E |
|---|---|---|---|---|
| **14** | 34.375 | -20.375 | 415.14 | 12.077 |
| **36** | 34.375 | 1.625 | 2.64 | 0.077 |
| **77** | 34.375 | 42.625 | 1816.89 | 52.85 |
| **21** | 34.375 | -13.375 | 178.89 | 5.2 |

| | | | | |
|---|---|---|---|---|
| **18** | 34.375 | -16.375 | 268.14 | 7.8 |
| **57** | 34.375 | 22.625 | 511.89 | 14.89 |
| **39** | 34.375 | 4.625 | 21.39 | 0.62 |
| **13** | 34.375 | -21.375 | 456.89 | 13.29 |
| | | | | 106.8 |

In Table 08, *O* stands for observed value, while *E* stands for expected value. After applying chi chi-square test, the value of X2 is 106.8.

*X2 calculated = 106.8*

After finding the value of calculated X2, now we will compare the calculated and tabular value of X2.

*X2 tabular:* to find the tabular value of X2, we have to check the degree of freedom,

*Degrees of freedom*= (*total rows-1*) (total columns-*1*)

(8-*1*) (3-*1*)

(7) (2)

=014

(0.05) (14) = 23.68

*X2 tabular = 23.68*

*Calculated value > tabular value*

The value of calculated x2 is greater than the tabular value of x2, null hypothesis is rejected and the alternative hypothesis is accepted here. It means that the proverbs depend upon these three variables.

**Collocations**

**Table 09 : Use of English Collocations as a Formulaic Expression**

| Collocations | Correct | Interchangeable | Incorrect |
|---|---|---|---|
| **As a result** | 137 | 13 | 0 |
| **Make a difference** | 131 | 10 | 9 |
| **Strong belief** | 129 | 18 | 3 |
| **Pretty sure** | 10 | 78 | 62 |
| **Total** | 407 | 119 | 74 |
| **Mean** | 101.75 | 29.75 | 18.5 |
| **StD** | 30.96 | 28 | 25.32 |

Table 09 suggests that the Urdu-speaking English formulaic expression learners also used the collocations in different ways.

First category deals with the Urdu collocations specific to the English language.

Second category is conceptually similar Urdu version of the English collocations.

Third category is linked with word for word Urdu translation of the collocations.

As shown in table 09, the first category of collocations is showing the greatest value of mean and standard deviation, followed by second category, and finally the third category with least mean and standard deviation.

**Category 1**

First category deals with the Urdu collocations specific to the English language.

**Correct**

**Table 10**

*Application of chi square test on category 1*

| O | E | (O-E) | (O-E)2 | (O-E)2/E |
|---|---|---|---|---|
| **137** | 101.75 | 35.25 | 1242.56 | 12.21 |
| **131** | 101.75 | 29.25 | 855.56 | 8.408 |
| **129** | 101.75 | 27.25 | 742.56 | 7.297 |
| **10** | 101.75 | -99.75 | 9950.06 | 97.789 |
| | | | | 125.704 |

In table 10, *O* stands for observed value, while *E* stands for expected value. After applying chi square test, the value of X2 is 125.704.

*X2 calculated = 125.704*

After finding the value of calculated X2, now we will compare the calculated and tabular value of X2.

*X2 tabular:* to find the tabular value of X2, we have to check the degree of freedom,

*Degrees of freedom*= (*total rows-1*) (total columns-*1*)

(4-*1*) (3-*1*)

(3) (2)

=06

(0.05) (06) = 12.59

*X2 tabular = 12.59*

*Calculated value > tabular value*

The value of calculated x2 is greater than the tabular value of x2, null hypothesis is rejected and the alternative hypothesis is accepted here. It means that the collocations also depend upon these three variables.

**Category 2**

Second category is conceptually similar Urdu version of the English collocations.

**Interchangeable**

**Table 11 : Application of Chi Square Test on Category 2**

| O | E | (O-E) | (O-E)2 | (O-E)2/E |
|---|---|---|---|---|
| **13** | 29.75 | -16.75 | 280.56 | 9.43 |
| **10** | 29.75 | -19.75 | 390.06 | 13.11 |
| **18** | 29.75 | -11.75 | 138.06 | 4.66 |
| **78** | 29.75 | 48.25 | 2328.06 | 78.25 |
| | | | | 105.45 |

In table 11, *O* stands for observed value, while *E* stands for expected value. After applying chi square test, the value of X2 is 105.45.

*X2 calculated = 105.45*

After finding the value of calculated X2, now we will compare the calculated and tabular value of X2.

*X2 tabular:* to find the tabular value of X2, we have to check the degree of freedom,

*Degrees of freedom*= (*total rows-1*) (total columns-*1*)

(4-*1*) (3-*1*)

(3) (2)

=06

(0.05) (06) = 12.59

*X2 tabular = 12.59*

*Calculated value > tabular value*

The value of calculated x2 is greater than the tabular value of x2, null hypothesis is rejected and the alternative hypothesis is accepted here. It means that the collocations also depend upon these three variables.

**Category 3**

Third category is linked with word for word Urdu translation of the collocations.

**Incorrect**

**Table 12: Application of chi square test on category 3**

| O | E | (O-E) | (O-E)2 | (O-E)2/E |
|---|---|---|---|---|
| **0** | 18.5 | -18.5 | 342.25 | 18.5 |
| **9** | 18.5 | -9.5 | 90.25 | 4.878 |
| **3** | 18.5 | -15.5 | 240.25 | 12.986 |

| | | | | |
|---|---|---|---|---|
| **62** | 18.5 | 43.5 | 1892.25 | 102.28 |
| | | | | 138.641 |

In table 12, *O* stands for observed value, while *E* stands for expected value. After applying chi square test, the value of X2 is 138.641.

*X2 calculated = 138.641*

After finding the value of calculated X2, now we will compare the calculated and tabular value of X2.

*X2 tabular:* to find the tabular value of X2, we have to check the degree of freedom,

*Degrees of freedom*= (*total rows-1*) (total columns-*1*)

(4-*1*) (3-*1*)

(3) (2)

=06

(0.05) (06) = 12.59

*X2 tabular = 12.59*

*Calculated value > tabular value*

The value of calculated x2 is greater than the tabular value of x2, null hypothesis is rejected and the alternative hypothesis is accepted here. It means that the collocations also depend upon these three variables.

**Phrasal verbs**

**Table 13: Use of English Phrasal Verbs as a Formulaic Expression**

| Phrasal verbs | Correct | Interchangeable | Incorrect |
|---|---|---|---|
| **Fed up** | 112 | 14 | 24 |
| **Growing up** | 62 | 78 | 10 |
| **Give up** | 46 | 87 | 17 |
| **Hold up** | 51 | 56 | 43 |
| **Go ahead** | 102 | 21 | 27 |
| **Total** | 373 | 256 | 121 |
| **Mean** | 74.6 | 51.2 | 24.2 |
| **StD** | 27.14 | 29.38 | 11.08 |

First category deals with the Urdu phrasal verbs specific to the English language.

Second category is conceptually similar Urdu version of the English phrasal verbs.

Third category is linked with word for word Urdu translation of the phrasal verbs.

The mean and standard deviations of the gain scores between all the three categories is mentioned above in table 4.13. As shown in table, the sample shows the greatest mean score with the category 2 phrasal verbs, followed by category 1, and finally category 3.

**Category 1**

First category deals with the Urdu phrasal verbs specific to the English language.

**Correct**

**Table 14: Application of Chi Square Test on Category 1**

| O | E | (O-E) | (O-E)2 | (O-E)2/E |
|---|---|---|---|---|
| **112** | 74.6 | 37.4 | 1398.76 | 18.75 |
| **62** | 74.6 | -12.6 | 158.76 | 2.128 |
| **46** | 74.6 | -28.6 | 817.96 | 10.96 |
| **51** | 74.6 | -23.6 | 556.96 | 7.46 |
| **102** | 74.6 | 27.4 | 750.76 | 10.06 |
| | | | | 49.358 |

In table 4.14, *O* stands for observed value, while *E* stands for expected value. After applying chi square test, the value of X2 is 49.358.

*X2 calculated = 49.358*

After finding the value of calculated X2, now we will compare the calculated and tabular value of X2.
*X2 tabular:* to find the tabular value of X2, we have to check the degree of freedom,
*Degrees of freedom*= (*total rows-1*) (total columns-*1*)
(5-*1*) (3-*1*)
(4) (2)
=08
(0.05) (8) = 15.51
*X2 tabular = 15.51*
*Calculated value > tabular value*
The value of calculated x2 is greater than the tabular value of x2, null hypothesis is rejected and the alternative hypothesis is accepted here. It means that the phrasal verbs also depend upon these three variables.

**Category 2**
Second category is conceptually similar Urdu version of the English phrasal verbs.
**Interchangeable**
**Table 15: Application of Chi Square Test on Category 2**

| O | E | (O-E) | (O-E)2 | (O-E)2/E |
|---|---|---|---|---|
| **14** | 51.2 | -37.2 | 1383.84 | 27.028 |
| **78** | 51.2 | 26.8 | 718.24 | 14.028 |
| **87** | 51.2 | 35.8 | 1281.64 | 25.03 |
| **56** | 51.2 | 4.8 | 23.04 | 0.45 |
| **21** | 51.2 | -30.2 | 912.04 | 17.81 |
| | | | | 84.346 |

In Table 15, *O* stands for observed value, while *E* stands for expected value. After applying chi square test, the value of X2 is 84.346.
*X2 calculated = 84.346*
After finding the value of calculated X2, now we will compare the calculated and tabular value of X2.
*X2 tabular:* to find the tabular value of X2, we have to check the degree of freedom,
*Degrees of freedom*= (*total rows-1*) (total columns-*1*)
(5-*1*) (3-*1*)
(4) (2)
=08
(0.05) (8) = 15.51
*X2 tabular = 15.51*
*Calculated value > tabular value*
The value of calculated x2 is greater than the tabular value of x2, null hypothesis is rejected and the alternative hypothesis is accepted here. It means that the phrasal verbs also depend upon these three variables.

**Category 3**
Third category is linked with word for word Urdu translation of the phrasal verbs.
**Incorrect**
**Table 16: Application of Chi Square Test on Category 3**

| O | E | (O-E) | (O-E)2 | (O-E)2/E |
|---|---|---|---|---|
| **24** | 24.2 | -0.2 | 0.04 | 0.0016 |
| **10** | 24.2 | -14.2 | 201.64 | 8.33 |
| **17** | 24.2 | -7.2 | 51.84 | 2.14 |
| **43** | 24.2 | 18.8 | 353.44 | 14.6 |
| **27** | 24.2 | 2.8 | 7.84 | 0.32 |

| | |
|---|---|
| | 25.3916 |

In table 16, *O* stands for observed value, while *E* stands for expected value. After applying chi square test, the value of X2 is 25.3916.

*X2 calculated = 25.3916*

After finding the value of calculated X2, now we will compare the calculated and tabular value of X2.

*X2 tabular:* to find the tabular value of X2, we have to check the degree of freedom,

*Degrees of freedom*= (*total rows-1*) (total columns-*1*)

(5-*1*) (3-*1*)

(4) (2)

=08

(0.05) (8) = 15.51

*X2 tabular = 15.51*

*Calculated value > tabular value*

The value of calculated x2 is greater than the tabular value of x2, null hypothesis is rejected and the alternative hypothesis is accepted here. It means that the phrasal verbs also depend upon these three variables.

**Discussion**

The primary objective of this study was to investigate the influence of the Urdu language on the acquisition of English formulaic expressions by L2 learners. Given the structural and lexical differences between Urdu and English, the study hypothesized that Urdu may exert a negative influence on learners' ability to accurately acquire and use English formulaic expressions. This hypothesis is supported by a number of prior studies in the field.

According to Conklin and Schmitt (2012), native speakers possess a significant advantage in processing formulaic language, while L2 learners often do not experience the same facilitation. This is particularly relevant in the case of Urdu-speaking learners of English. The learning of formulaic expressions in a second language is a gradual and effortful process, and L2 learners typically acquire these expressions at a much slower pace compared to native speakers (Kuiper, Columbus, & Schmitt, 2009; Bores & Lindstromberg, 2012).

Hama (2010) found that first language (L1) interference and limited exposure to formulaic language are major contributors to collocational errors in L2 learners. Similarly, Wray (2008) emphasized the strong influence of specific phrase associations in the acquisition of formulaic language. Because formulaic expressions are often culturally and linguistically bound, the divergence between Urdu and English in idioms, proverbs, collocations, and phrasal verbs poses a significant challenge. Each language has its own distinct lexical and semantic frameworks, which complicates direct translation or equivalence. Yamashita and Jiang (2010) also emphasized the role of L1 influence in the acquisition of L2 formulaic language.

Wray (2004, 2008) further highlighted the importance of formulaic language in the development of natural language fluency, regardless of a learner's proficiency level. Supporting this, Bahns and Eldaw (1993) cautioned against relying on L1–L2 collocational equivalents in language classrooms, as they often lead to confusion and errors in usage. These findings are consistent with the current study's conclusion that Urdu, as the L1, negatively influences the learning of English formulaic expressions among L2 learners.

This study was conducted across several major educational institutions in Bhimber, Azad Kashmir, including Read Foundation College, Army Public School and Colleges, Insight Model College, and The Learner College. A total of 150 students participated in the research. The focus was on assessing the influence of Urdu on the learning of four types of English formulaic expressions: idioms, proverbs, collocations, and phrasal verbs. Data were collected using questionnaires and analyzed using the Chi-square test.

Three categories were established for analysis:

1. Category 1: Idioms and proverbs unique to English

2. Category 2: English equivalents that are conceptually similar to Urdu expressions
3. Category 3: Literal, word-for-word English translations of Urdu expressions

The results revealed varying levels of difficulty across these categories. For idioms, Category 2 showed the highest mean and standard deviation, indicating better comprehension when conceptual similarities existed between Urdu and English. For proverbs, however, Category 1 had the highest scores, suggesting greater recognition of original English expressions. In terms of collocations, Category 1 again showed the highest values, followed by Category 2 and then Category 3. Similarly, for phrasal verbs, the highest mean and standard deviation were observed in Category 1, followed by Categories 2 and 3, respectively.

These findings highlight that Urdu and English formulaic expressions differ significantly, and this discrepancy hinders Urdu-speaking learners from acquiring English expressions with ease. The negative transfer from Urdu to English, especially when learners rely on literal translations, leads to errors and misinterpretations. Thus, it can be concluded that the Urdu language has a negative influence on the acquisition of English formulaic expressions among L2 learners.

Recommendations for Future Research

Future studies should investigate pedagogical interventions to mitigate the negative effects of L1 interference, such as providing explicit instruction in L2 formulaic expressions and increasing exposure to authentic English usage. Comparative studies involving learners of different L1 backgrounds may also shed further light on the universal versus language-specific challenges in formulaic language acquisition.

## Conclusion

The present study investigated the impact of the Urdu language on the acquisition and use of English formulaic expressions among L2 learners in selected educational institutions of Bhimber, Azad Kashmir. The findings provide clear evidence that Urdu, as the learners' first language, significantly influences their understanding and use of English formulaic language—particularly idioms, proverbs, collocations, and phrasal verbs.

Through the categorization of expressions into three distinct groups—(1) idioms and proverbs specific to English, (2) conceptually similar English expressions derived from Urdu, and (3) word-for-word English translations of Urdu expressions—the study demonstrated varying levels of comprehension and usage accuracy among learners. Category 2, which involved conceptually similar expressions, generally showed better performance, suggesting some level of positive transfer. However, Categories 1 and 3 revealed considerable challenges, with Category 3 especially reflecting negative transfer due to literal translation strategies that do not align with native-like English usage.

These results align with prior literature (e.g., Conklin & Schmitt, 2012; Wray, 2008; Hama, 2010) highlighting the complexities of formulaic language learning and the role of L1 interference. The influence of Urdu was particularly apparent in the students' tendency to rely on literal or direct equivalents, which often led to incorrect or awkward phrasing in English.

Overall, the study confirms that Urdu has a predominantly **negative influence** on the acquisition of English formulaic expressions, particularly when learners rely heavily on direct translation strategies. However, the presence of partial conceptual similarity between Urdu and English idioms suggests that **cross-linguistic awareness and targeted pedagogical interventions** could improve learners' competence in formulaic language.

## References

Ali, S., & Ullah, Z. (2015). Semiotic insight into cosmetic advertisements in Pakistani print media. International Journal of Humanities and Cultural Studies, 2(1), 689-705. http://www.ijhcs.com/index.php/ijhcs/article/ view/358/425

Badshah, I., Kausar, G., & Khan, R. N. (2017). Investigating Pakistani university students' beliefs about learning English. The Dialogue, 12(4), 415-432. https://pdfs. semanticscholar.org/56a3/b9aaaf4b930532be184440ebf900a91a0f9a.pdf

Beardsmore, B. H., (1982). Bilingualism: basic principles. Clevedon, Avon, England: Tieto. Berk, L.

(1999). The Noun Phrase. In English Syntax. New York: Oxford University Press
Ellis, R. (1985). The Role of First Language. In Understanding Second Language Acquisition (pp. 19-49). New York: Oxford University Press. Erarslan,
A, & Hol. D., (2014). Language Interference on English: Transfer on the Vocabulary, Tense and Preposition Use of Freshmen Turkish EFL Learners. ELTA Journal.
Fatemi, M. A., Sobhani, A., & Abolhasani, H. (2012). Difficulties of Persian learners of English in pronouncing some English consonant clusters. World Journal English Language, 2(4), 69-75.
Fries, C. (1945). Teaching and learning English as a foreign language. Ann Arbor: University of Michigan Press.
Fromkin, V., & Rodman, R. (1998). Language Acquisition. In An Introduction to Language (Sixth ed., p. 349). Christopher P.Klein.
Manan, N, A, A., & Raslee, N, N. (2016). First language (l1) interference in the writing of English as a second language (ESL) learners.
Masood, M. H., & Shafi, S. (2020). Exploring Marxist perspective amidst exploitation and false consciousness in Hosain's The Old Man. International Journal of Applied Linguistics & English Literature (IJALEL), 9(3), 18-24. http://dx.doi.org/10.7575/aiac.ijalel.v.9n.3p.18 Nemati.
M., & Taghizadeh. M. (2013). Exploring Similarities and Differences between L1. International Research Journal of Applied and Basic Sciences.
Odlin, T. (1989). Language transfer: Cross-linguistic influence in language learning. Cambridge University Press. Platts, J. T. (1874). A Grammar of the Hindustani or Urdu Language. London: Publishers of the India Office.