"INVESTIGATING THE ROLE OF FORMULAIC LANGUAGE IN ESTABLISHING COMMON GROUND AND SEEKING SALIENCY IN AN INTERCULTURAL ENVIRONMENT: ESSAY WRITING SKILLS OF UNDERGRADUATE STUDENTS"

[1]*Nadia Urooj*
PhD Scholar, Minhaj University, Lahore.
Email: nadiaurooj8@gmail.com
[2]*Prof. Dr. Zafar Iqbal Bhatti*
Professor, Minhaj University, Lahore.
Email: zafar.eng@mul.edu.pk
[3]*Rakhshanda Furnaz*
Lecturer, ELC, University of Central Punjab, Lahore.
Email: rakhshanda.fawad@ucp.edu.pk

Abstract
The aim of this research is to examine the central role formulaic language has in the essay writing competence of undergraduate students with special focus on its double role of establishing common ground and enhancing saliency in an intercultural context. Formulaic language, as defined as habitualized expressions, lexical bundles, and repeated phrases in various genres of communication, is tested against its capacity for creating mutual understanding, coherence, and cohesion in essay writing among students. Using Kecskes' Formulaic Continuum to the study of Intercultural Communication, this research combines qualitative analysis with quantitative measurement to describe the process by which formulaic expressions facilitate the achievement of mutual understanding between writers and readers in academic writing. In addition, it examines the impact of formulaic language on the clarity, persuasiveness, saliency, and overall effectiveness of undergraduate essays. Through the examination of the intricate relationship between the use of formulaic language, intercultural communication, and writing development, this research offers useful insights into pedagogical strategies that can improve the academic writing of students. Simultaneously, it adds to the refinement of theoretical knowledge of linguistic devices employed in written intercultural communication.

1. Introduction:

1.1. Role of English Language

Intercultural engagement is a common occurrence as a result of the global trend of students crossing borders to further their education. As a result, there has been a noticeable increase in the usage of Lingua Franca, especially when communicating across cultural boundaries. Non-native English speakers, who have common goals in academic settings with a wide range of ethnic backgrounds, frequently use formulaic language when expressing themselves. Graddol (2006) noted that the idea of a standardized language is called into question by this fact. Furthermore, as suggested by Kecskes (2000), more research is required to determine the impact that formulaic language plays in obtaining native-like proficiency because there aren't any native English speakers in the group. Given that English has become a lingua franca and is widely used by students from culturally varied backgrounds in universities, Kecskes (2007) cites Wray (2002) to highlight the differences in language preferences between groups and languages. According to him, linguistic inventiveness improves efficacy of communication at the oral and written levels.

1.2. Linguistics Saliency and Formulaic Language

These innovative linguists provide their users with a psychological saliency through the formulaic language they employ. Based on current research, it is generally agreed that formulaic language is salient because it is primarily composed of functional units and represents a significant amount of pertinent information. It's also well known that formulaic language reduces cognitive processing and seems native-like, even when used by persons from different cultural backgrounds. Many scholars concur that formulaic language is frequently used in both written and oral communication (McEnery 1996; Biber, 1999).

The formulaic language used by these creative linguists imparts to its users the quality of psychological saliency. According to the consensus of recent research, formulaic language is salient because it consists mostly of functional units and constitutes a substantial reservoir of relevant information. It is also known that, even when used by people from various cultural backgrounds, formulaic language appears to be native-like and lessens cognitive processing. The assertion that formulaic language is often employed in spoken and written conversation is supported by numerous academics (McEnery 1996; Biber, 1999).

This linguistics creativity in the form of formulaic language imparts the feature of psychological saliency on the part of its users. The consensus of the recent research is that formulaic language gives saliency because they are mostly functional units and are significant reservoirs of meaningful information. It is also acknowledged that formulaic language reduces cognitive processing and leaves an impression of native-like users despite a different cultural background. Many researchers second the claim that formulaic language tends frequently (McEnery 1996; Biber, 1999) used in written and spoken discourse. A very authentic description of formulaic language is presented by Simpson-Vlach and Ellis these occur in collocational strings of words and merged in the form of patterns (2010). In another research, the concept of "idiom principle" showed that "semi-pre-constructed phrases" facilitate the user in the form of a complete unit coined by different segments of the sentence (Sinclair, 2004). A prominent feature of this principle is not based on the criteria of frequency which dilutes the boundary (Biber, 1999) between conventionalized word sequence and the least psychologically salient combination of frequently occurring word strings.

1.3. Formulaic Language Perception in the Past

Language learning theories in the second language domain highlight the importance of frequency in the instructional strategies of the classroom setting. Learners register the formulaic language in writing and speaking after noticing the occurrence of specific word strings according to either the situation or sentence types (Ellis et al., 2008). It is observed that years after years of instructions in the classroom setting determine the use of formulaic language by the non-native users however this is not the measure for accuracy and efficiency. Another very important element is the need to understand the coherence and this is the reason most second language learners avoid using formulaic language because they don't see the logic according to their cultural knowledge. So they determine the use of formulaic language based on their frequency whereas selection matters because of association that provides a psychological saliency in a contextual situation. While using formulaic language functionality is an important aspect that is the reason to immerse in the culture to make the nonnative understand those functions (Kecskes, 2004).

The first research question aims to see the frequency of formulaic language used by the nonnative users whereas for the second research question the purpose is to examine how students establish common ground and seek psychological saliency.

1.4. Kecskes' Formulaic Continuum

The concept behind the term "Formulaic Language" is the strings of words in the form of collocations, idioms, phrasal verbs, metaphors that are mostly fixed expressions or situation-bound utterances or grammatical units or lexical bundles that frequently occur together (Howarth, 1998) and deliver holistic meaning contrary to its formation having either literal meaning or replete with pragmatic meaning and the interesting thing it behaves "… as a single semantic unit (Gairn and Redman, 1986)".

Depending on the type of formulaic language, few categories are observed to have conventional meanings and are frequently used in discourse. Kecskes developed a formulaic continuum with varying functional aspects starting with the grammatical unit and ending with pragmatic expressions. Furthermore, language is the combination of different parts as emphasized by the recent linguistic approach. There is no denying the truth that formulaic language compromised rules and systems of language however, first and second-language learning starts with lexical bundles, phrases and linguistic routines that are stored in the long-term memory and retrieved when needed (Wray, 2002).

1.5. Threefold Reasons to use Formulaic Language

Kecskes (2004) asserted that formulaic language imparts the expression of native users. Wray observed that "people have preferred ways of saying things (2002). This means that native users utilize prefabricated expressions to exhibit idiomatic language in phrasal utterances. Kecskes asserted that the reason behind using formulaic language is threefold. First of all, it reduces cognitive processing (Miller and Weinert, 1998) and psycholinguistically it is proved that formulas help in saving mental efforts (Sinclair, 1991) and yet seem to have a cognitive impression. This also provides a real-time saliency in the discourse on the part of the speaker/writer. Secondly, formulaic language lends framing power to the discourse. Goffman defined frames as primary cognitive frames to perceive and represent reality (1974). Frames can capture the attention of the listeners or readers. However prefabricated chunks are not specially constituted but rather retrieved from the subconscious mind during a discourse. These frames work like conceptual metaphors (Lakoff, 1992). So first both features are about psychological saliency which helps in describing the writer's point of view. Thirdly formulaic language has the powerful nature of providing common ground in establishing a comprehension level between reader and writer and similarly between speaker and hearer. It means formulaic language can be interpreted based on shared experiences. Formulaic language originates from a common cultural background and is expected to be understood in the same way as expressed by the writer or speaker (Tannen and Oztek, 1981) hence communication occurs. The dilemma with the nonnatives is that they lack common ground which causes a communication gap if not use the common language.

1.6. Research gap in Formulaic Language

Formulaic language in pragmatic research has been avoided in the recent past (Kecskes, 2004) because of deviating from the principles and procedures followed while writing. However, the importance of the formulaic language cannot be denied and there is a need to bring this topic into consideration and think out of the box for its presence and out of the boundaries of 'neo-Gricean' and 'relevance' types of theories to examine the impact of formulaic language in academic writing. Kecskes already justified the native-like impression of formulaic language in daily conversation. Other researchers are working in the area of academic writing. In the stream of those researches, this study would be an add-on to examine the frequency of formulaic language use by nonnatives having English as a second language for studies. Their essay writing activity will demonstrate the

frequency of formulaic language in their writing through 'Sketch Engine software' and a corpus comparison will be done through COCA to see whether they are successful in establishing common ground and seeking psychological saliency in their communication or not. After data analysis, recommendations will be given according to the results.

2. LITERATURE REVIEW

In this section, relevant literature about formulaic language and its use in writing is observed concerning past research. Kecskes suggested a formulaic continuum starting from grammatical items and ending on idiomatic expressions.

Academic writing is composed of various linguistic patterns, one of which is formulaic language. Formulaic language is observed as sequences of words or expressions that are commonly used in specific contexts and provide the reason for common ground to the foreign writer to be part of a particular discourse group. The purpose of using formulaic language in academic writing is to cater to several functions, for example, creating coherent patterns, communicating complex ideas in a meaningful way, and establishing salience by exhibiting the writer's familiarity with situational conventions. The objective of the literature review is to shed light on the use and implication of formulaic language in academic writing. The use of formulaic language is witnessed in spoken discourse communities in research. However, there is less work conducted to see its use in writing discourse. This study aims to identify and fill this gap and provide evidence that formulaic language helps convey ideas according to the given context.

Abdalhussein (2022) authenticated / established the ground (Wei & Ying, 2011) for formulaic language in the area of language learning. As it is substantiated that formulaic language plays a vital role in different stages of the language learning process (Wood, 2002). It was reasoned by Wood (2002) that expressions or strings of words in the form of SBUs (situation-bound utterances, phrasal verbs, grammatical units, and fixed semantic units that are different types of formulaic language make the processing of language easier and native-like fluent and coherent. Furthermore, it supported the storing and retrieving (Wei & Ying, 2011) method to use the language effortlessly convincing.

In continuation with the previous studies, there is more evidence in support of the usefulness of formulaic language with reference to the time (Conklin & Schmitt, 2008) variable affecting the process of native and non-native speakers. Research supported that both groups performed better while using formulaic expressions as it took less time to produce language as compared to non-formulaic expressions. To further support this argument, another initial research explored that multiple types of formulaic expressions facilitate the production process of language by new learners (Wood, 2006). One more research specifically focusing the use of discourse markers by language learners saving time by reducing cognitive load hence fluent and coherent communication. However, all these researches focus on the fluency and coherence of spoken language. Here, an important point to consider is that coherence is also a feature of textual content. For this, there is linguistic theory presented by Halliday & Hassan supplemented textual coherence in which the writer uses the signposts for the readers' guide to navigate through the text.

Academic writing demands cohesion and coherence among which coherence facilitates the readers to derive meaning from the text. And as discussed above formulaic language is used to create coherence in the text. Since both impart the same purpose of bringing sense in writing hence interdependent. Hayland (2012) researched to examine the use of formulaic language to convey meaning and purpose in writing. A research based on the argumentative essay written by Chinese Undergraduate students (Yang & Sun, 2012) revealed that irrespective of the similarities and

difference in the use of cohesive devices, there is a positive correlation in their overall use. However, there is a dearth of literature claiming the effectiveness of formulaic language in creating coherent writing specifically in the intercultural context. There is also less research in the area of different types of formulaic language used in writing. It has been observed that the proficiency of students using cohesive devices correctly should be checked to authenticate the effectiveness of formulaic language.

There are different types of formulaic language categorized by different researchers (Sinclair, 1991; Wray, 2002; Kecskes, 2003; Schmitt, 2010). Now it is obvious that the concept of formulaic language is active however, to identify it in writing is questionable. Even after various definitions, there are chances that the identification is complicated. Different research address this subject as Schmitt (2010) asserted that formulaic language poses challenges in the form of range and size. The acquisition approach is one such effort by Schmitt (2010) to identify the formulaic language. The spoken language supports this psycholinguistics approach that the predesigned language units bring fluency and work sometimes as fillers. Additionally, it helps in scanning the text with quick eye movement (Shaoul & Westbury, 2011) yet there are some impediments due to deviations of certain kinds. The phraseological approach is also a technique used because of its transparent and replaceable conditions based on collocation rules. Besides it acceptance the manual identification process is laborious. After a comprehensive examination, Schmitt supports corpus analysis as the most suitable approach to identify formulaic language. The more or less frequent occurrences of formulaic language help research to identify and analyze the nature of its use and effectiveness. However, the frequency parameter is not sufficient to identify FL because combinations vary and need proper reasoning to be identified as FL. Generally, people trust their intuition which is a non-scientific approach (Wray, 2002) however a more logical way to identify FL is the use of corpora (Wray, 2002). Chomsky (1965) and Sinclair (1991) opposed the idea of trusting intuition because of no authentic reason that should be further verified by the available corpus sources (Wray, 2002). There are other issues in detecting the use of FL in writing for example different members have different criteria of considering certain chunks as FL due to the variation in their frequency. Another issue is their very composition which can be regular yet idiomatic or genre-specific. Hyland highlighted this variation and argued that categorizing general academic writing is not supportive of FL while subject-specific is possible (2008). Simpson-Vlach and Ellis asserted that there is a notable volume of FL that makes a broad range of academic writing (2010). Similarly, there are other forms of categorization, such as research-, text-, and participation-oriented. Here, Dontacheva-Navaratilova argued that such kind of categorization could be ambiguous (2012).

There are many researches, especially with reference to academic writing, shedding light in this particular area of language learning. Biber argued that language learning is easier when it comes to focus on the sequential formation of certain textual formats (2001). This strategy works in the form of implementing the first stage following the use of spoken chunks and then transferring it into an academic writing piece by reinforcing its use in a specific sequence or situation. This technique is supported by many researchers as Staples, Gray, Biber, & Egbert 2016. Another study in the domain of academic writing argued that formulaic language is useful in teaching certain patterns and formats that are the demand of different types of writing hence frequent in delivering university lectures (spoken) as compared to textbooks, handouts, and worksheets written (Biber and Barbieri, 2007). The authenticity of this study is acknowledgeable as data is collected from written and spoken academic writing samples of TOEFL (2000). The lower frequency of the lexical bundles was due to topic and objective specificity.

Observing the patterns, phrases, and expressions, it is obvious that FL is useful in communication hence its importance in language learning has been observed by many researchers. There is no denying the fact that FL is an effective reservoir concerning contextual communication and provides the opportunity to create salience with the help of productive performance in an event of classroom activity or other academic context of writing. However, imparting the concept of using FL in writing and adapting its use are both quite perplexing, especially in the academic context. There are many reasons behind this challenging nature of the FL that are discussed above. There are also some past researches available for reflection and guidance to identify the gap in it. Biber and Barbieri pointed an important recognition is the utilization of reference linguistic items was preferred by the students in writing whereas sequential and positional & illustrative expressions were frequent in spoken discourse (2007). Role-play activities and dialogues were utilized by Rajeswasan (2019), in which first-year students of undergraduate performed. Both role-play and writing assessments were checked. The results showed that multiword units (FL) are effective in learning language and improve the students' performance in writing and show a correlation. In another study, explicit teaching was endorsed as compared to the conventional vocabulary teaching approach. It was observed (Le-Thi et al., 2017) that contextual learning was effective instead of making them memorize chunks of linguistic items. Then there was another study (Masoni, 2020) conducted in Italian University using picture books and asking students to describe the situation. It was acknowledged as an effective strategy to stimulate classroom interaction in a better and communicative way which is helping in acquiring language. Martinez (2013) highlighted the challenge faced by the teacher in selecting and teaching MWU or formulaic language.

As mentioned above different researchers categorized formulaic language according to their study design. However, this study uses Kecskes' formulaic continuum to identify the frequency of formulaic language used by undergraduate students in expository writing. Students from different countries study at the private university in Lahore, Pakistan. They create an intercultural environment that demands common ground for communication in an academic context yet salient.

2.1. Research Question
1. What is the frequency of different elements of formulaic language in the essays written by undergraduate students?
2. How does the language of written essays deviate from the elements of formulaic language?

3. METHODOLOGY
3.1. Framework

This study employed a mixed-method approach (see Creswell, 2009) with one data set of 150-word essays collected from 15 participants on the given topic. The topic of the essay was selected from the HSSC level. The qualitative data is analyzed with the help of SketchEngine software in the form of variations of N-gram from 3-5 words and quantitative data in the form of frequency is taken with the help of "the Corpus of Contemporary American English" (COCA) (Davies, 2008). The corpus of data was collected from the undergraduate BS level students of a private university studying communication skill course (CEFR Level-B2). The purpose of the study was to see the functional aspects of the formulaic language in written communication. Conventionalized meanings in language sequences vary in nature across different fixed expressions, supporting the idea of a continuum due to their distinct functional aspects in various situations (Kecskes, 2014).

The purposively sampled 15 participants of this study were selected from the International Students Service (ISS) center at UCP from, Saudi, Chinese, Sri Lankan, and Pakistani nationalities. All of them are final-semester BS-level students, studying in the Faculty of Information &

Technology, Faculty of Management Sciences, Faculty of Mas & Media Communication. Purposive sampling is useful for a qualitative or mixed-method approach where the researcher seeks to obtain in-depth insight of the phenomenon (Kecskes, 2007) mainly the aspect of using formulaic language.

For qualitative analysis, the SketchEngine Corpus software is utilized to identify the N-grams that will help in identifying the presence of formulaic language as define in formulaic continuum (Kecskes, 2000). The quantitative data is further analyzed by using the "COCA" tool to measure the frequency of formulaic expressions used by students in their writing. Kecskes (2000) asserted that there are many forms of formulaic sequences and presented a formulaic continuum table 1.1, starting from left with grammatical units, fixed semantic units, phrasal verbs, speech formulas, situation-bound utterances, and idioms. A non-native writer has a certain edge when using formulaic language, with benefits such as less processing time and effort, fixed expressions providing referential frames, and shared common ground to fulfill communicative functions.

Gramm. Units	Fixed Sem. Units	Phrasal Verbs	Speech Formulas	Situation-bound Utterances	Idioms
going to	as a matter of fact	put up with	going shopping	welcome aboard	kick the bucket
have to	suffice it to say	get along with	not bad you know	help yourself	spill the beans

Table 1: Formulaic Continuum

3.2. Participants

The research participants were International students of UCP studying at the Undergraduate level in different faculties. These participants are from Turkey, China, Sri Lanka, and Saudi Arabia. The average four students were from all the mentioned nationalities. It was certified that they had an Intermediate level of education at the time of admission in the university and specifically at the B1 level of English Language proficiency. They have basic vocabulary and grammar knowledge of English language and are eligible to '… communicate in most situations while traveling in an English-speaking area. They can write simple connected texts on familiar topics (British Council).'

3.3. Tools

An essay writing task was assigned to the students with detailed instructions set. The reason behind detailed instructions was to keep the students strictly topic-focused and a word limit (250 words) was assigned so that they should be conscious of vocabulary choice. It was a half-hour task. The topic was the narrative 'The Dream That Never Let Me Sleep!'. With the help of 'Sketch Engine Software' a frequency of 3-5 N-grams was measured and then the comparative analysis was made with COCA for the use of different types of formulaic language in academic writing.

4. DATA ANALYSIS

4.1. The 3-gram word List and Frequency

Starting with the 3-5 N-gram lists obtained from the Sketch Engine, here is a detailed discussion of the formulaic expression. First of all, look at the first table (1.1)

	N-gram	Frequency
1	be able to	7
2	I want to	6
3	dream of becoming	5

Table 2. 3-gram, word (items: 3, total frequency: 18)

There is a total of three 3-gram words were observed. Concerning Kecskes formulaic continuum the first two (be able to, I want to) 3-grams are most used grammatical units and are on the scale of least salient because they are the part of structural unit of a sentence. The third (dream of becoming) N-gram is a fixed semantic unit with contextual meaning referring to the topic of the essay. Using grammatical and fixed semantic units, the requirement of sharing common ground is fulfilled as both writer concept and reader interpretation is same as being structural usage.

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	629214	87878	72671	152015	141917	62203	48866	49449	14215	69584	80750	76001	77758	84003	80569
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	633.64	683.27	584.86	1,186.93	1,125.11	525.71	387.55	406.18	118.67	574.57	644.91	609.84	631.92	681.02	656.36

Want to

Table 3. Want to / wanna

The frequency of formulaic expression 'want to' is of significant number however they are not as salient in written discourse as in spoken. However, its use is based on principles of language hence reason to shared common ground.

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	127536	23910	20584	13692	23145	10263	14385	11881	9676	13297	13758	13911	13762	14237	14077
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	128.43	185.91	165.66	106.91	183.49	86.74	114.08	97.59	80.77	109.80	109.88	111.62	111.84	115.42	114.68
SEE ALL SUB-SECTIONS AT ONCE															

Table 4. be able to

The second grammatical unit (be able to) is frequently used in spoken and academic discourse and share common ground. This is also least salient but meeting the criteria of grammar rule to make a logical sentence.

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	335	47	36	34	32	35	74	61	16	30	34	52	50	46	40
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	0.34	0.37	0.29	0.27	0.25	0.30	0.59	0.50	0.13	0.25	0.27	0.42	0.41	0.37	0.33
SEE ALL SUB-SECTIONS AT ONCE															

Table 5. 'Dream of becoming'

The third N-gram (dream of becoming) is a fixed semantic expression that has framing power and is comparatively salient to point to the topic. The above table (1.2) is authenticating its most frequent use in spoken discourse whereas it is also used in academic discourse as well to assert saliency.

4.2. The 4-gram Word list and Frequency

	N-gram	Frequency		N-gram	Frequency		N-gram	Frequency
1	my dream come true	3	13	is clear to me	2	25	the drugs to patient	2
2	keeps me up at	3	14	but I believe that	2	26	three to four years	2
3	me up at night	3	15	chose me to come	2	27	dream of becoming an	2
4	my dream of becoming	3	16	Making my dream come	2	28	up with new ideas	2
5	my life and society	3	17	I can open variety	2	29	wo n't be easy	2
6	to adopt a daughter	3	18	completing my degree I	2	30	to achieve my dream	2
7	dream of becoming a	3	19	of my dream is	2	31	dream that does n't	2
8	would be able to	3	20	that does n't make	2	32	To achieve this dream	2
9	a passion within me	2	21	that keeps me up	2	33	I started to dream	2
10	after completing my degree	2	22	that my dream of	2	34	I believe that I	2
11	I would be able	2	23	This dream is n't	2			
12	everything is clear to	2	24	the beauty of stars	2			

Table 6. 4-grams, word (items: 34, total frequency: 76)

In the above table (1.5), a 4-gram word list is presented. Here, it is important to mention that phrases and word strings obtained through Sketch Engine are many but not all are formulaic expressions. We are selecting and discussing only those formulaic expressions that have the power of psychological saliency as given in the formulaic continuum. The first selected N-gram

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD		1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	28	7	5	1	2	4	5	2	2		2	3	4	2	2	3
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8		121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	0.03	0.05	0.04	0.01	0.02	0.03	0.04	0.02	0.02		0.02	0.02	0.03	0.02	0.02	0.02
SEE ALL SUB-SECTIONS AT ONCE																

Table 7. frequency of 'After completing my …'

The next 4-gram word (everything is clear …) is a fixed semantic unit and conveys its meaning as a whole. Pragmatically, it is charged and shares common ground based on its literal meaning.

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD		1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	37	5	3	12	1	6	3	5	2		5	4	4	4	5	7
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8		121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	0.04	0.04	0.02	0.09	0.01	0.05	0.02	0.04	0.02		0.04	0.03	0.03	0.03	0.04	0.06
SEE ALL SUB-SECTIONS AT ONCE																

Table 8. frequency of 'Everything is clear'

As per AI feedback, 'The sentence "Everything is clear" is correct and can be used in written English. It can be used when something has been explained or understood and there is nothing left to clarify. For example, "Thanks for walking me through this project, now everything is clear."'
It's a fixed semantic unit and has psychological saliency in addition to sharing common ground.

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	28	7	5	1	2	4	5	2	2	2	3	4	2	2	3
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	0.03	0.05	0.04	0.01	0.02	0.03	0.04	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02
SEE ALL SUB-SECTIONS AT ONCE															

Table 9. After completing my degree

The phrase 'after completing my degree' is correct and usable in written English. It can be used when someone is considering doing something after they have obtained a degree or diploma, and they want to mention it in their writing. For example: After completing my degree in Accounting, I am now looking for a job in the finance industry. It is a fixed semantic expression and shares common ground describing the literal meaning of the phrase. On the scale of formulaic continuum it is least salient phrase

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	833	198	125	74	249	30	67	53	37	111	127	85	84	53	50
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	0.84	1.54	1.01	0.58	1.97	0.25	0.53	0.44	0.31	0.92	1.01	0.68	0.68	0.43	0.41
SEE ALL SUB-SECTIONS AT ONCE															

Table 10. Frequency of 'but I believe that'

'but I believe that' is a perfectly valid part of a sentence in written English. You can use this phrase to introduce a statement of opinion or a conclusion that you have reached. For example: "I've looked into the issue, but I believe that the best approach is to wait and see what happens." According to the frequency table, it is more like a speech formula so at the highest level of psychological saliency. This authenticates the already established common ground and seeks a salient position.

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	5167	776	775	521	447	642	1018	545	443	571	695	667	595	601	487
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	5.20	6.03	6.24	4.07	3.54	5.43	8.07	4.48	3.70	4.71	5.55	5.35	4.84	4.87	3.97
SEE ALL SUB-SECTIONS AT ONCE															

Table 11. Frequency of 'the beauty of something)

According to the definition of the Cambridge dictionary, 'beauty of something' is an idiom. On the formulaic continuum, idioms are the most salient formulaic expressions and exhibit the native-like skill of using a second language in writing. Hence, the user is not only sharing common ground but also at the level of success in achieving psychological salience to get the attention of the readers. The frequency in the table (1.7) supports its academic use as well.

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD	1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	1341	235	230	27	61	23	206	64	495	160	136	142	136	148	154
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8	121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	1.35	1.83	1.85	0.21	0.48	0.19	1.63	0.53	4.13	1.32	1.09	1.14	1.11	1.20	1.25
SEE ALL SUB-SECTIONS AT ONCE															

Table 12. Frequency of 'To achieve this …'

The academic frequency of 'to achieve this …' is very high due to its being a grammatical unit. This means it shares the common ground but saliency level is very low. It means that structurally it is used to make grammatically correct sentences and mostly use as a motivational quote.

4.3. The 5-gram Word List and Frequency

	N-gram	Frequency
1	keeps me up at night	3
2	. to achieve this dream	2
3	. this dream is n't	2
4	, but i believe that	2
5	god chose me to come	2
6	i would be able to	2
7	after completing my degree i	2
8	making my dream come true	2
9	my dream of becoming a	2
10	my life and society .	2
11	to adopt a daughter .	2
12	that keeps me up at	2
13	everything is clear to me	2
14	. making my dream come	2
15	, after completing my degree	2

Table 13. 5-grams, word (lowercase) (items: 15, total frequency: 31)

The above table is 5-gram word list obtained through Sketch Engine. This is the lengthiest word string and that's why most of the expressions are having comma (,) or full stop (.) in the beginning or end of the string that is excluding those expressions from the criteria of formulaic expressions. Some of them are better in the form of 3-gram and 4-gram expressions and are already discussed. Now three expressions are unique. Look at the first formulaic expression 'keeps me up at night' which is an idiom and is therefore the most salient.

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD		1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	150	18	21	56	12	20	16	5	2		16	11	21	11	26	26
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8		121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	0.15	0.14	0.17	0.44	0.10	0.17	0.13	0.04	0.02		0.13	0.09	0.17	0.09	0.21	0.21

Table 14. Frequency of 'keeps me up …'

The variation in the frequency of 'keeps me up …' shows that pragmatically it is charged. Being an idiomatic expression, it is psychologically salient on the scale of the formulaic continuum and shows the writer's grip as a native-like language user.

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD		1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	13413	2613	2215	839	1305	1302	1549	1559	2031		1282	1447	1349	1487	1500	1520
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8		121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	13.51	20.32	17.83	6.55	10.35	11.00	12.28	12.81	16.95		10.59	11.56	10.82	12.08	12.16	12.38
SEE ALL SUB-SECTIONS AT ONCE																

Table 15. Frequency of 'God chose … to …'

The above frequency table from COCA software shows that "chose to" is a grammatical unit. Another crucial information about this 5-gram expression is that it is in the Bible. This verse is pragmatically charged with religious meaning. Contrary to its presence on the scale of the formulaic continuum at the least salient position, the appearance of this quote as a verse in the Bible makes it salient and gives native-like language use.

SECTION	ALL	BLOG	WEB	TV/M	SPOK	FIC	MAG	NEWS	ACAD		1990-94	1995-99	2000-04	2005-09	2010-14	2015-19
FREQ	7	0	1	3	1	1	0	1	0		0	1	1	3	0	1
WORDS (M)	993	128.6	124.3	128.1	126.1	118.3	126.1	121.7	119.8		121.1	125.2	124.6	123.1	123.3	122.8
PER MIL	0.01	0.00	0.01	0.02	0.01	0.01	0.00	0.01	0.00		0.00	0.01	0.01	0.02	0.00	0.01
SEE ALL SUB-SECTIONS AT ONCE																

Table 16. Frequency of 'making my dream come true'

In this last formulaic expression, it is clear that being an idiomatic expression, users are not using this in academic writing. However, it is evident (Bible verse, Romans 1:1-32 From Paul) that it is highly psychologically salient. If we examine the frequency of 'come true', we can see that academically it is being used. To get native-like expression there is a need to inculcate the knowledge of the formulaic language in second language learners.

5. CONCLUSION

To conclude the research based on the frequency of formulaic language and the deviation from using it in academic writing we are in the position to answer these questions about our results and analysis. Reflecting on the results about the frequency of different types of formulaic language on the scale of continuum, it is obvious that students do not prefer to use formulaic language or it seems they don't know about the importance and impact of its use in writing.

5.1. Discussion

The corpus of approximately 3500 words from 14 students of diverse cultural backgrounds is minor but has demonstrated that formulaic language is not their preferred way of communication. They feel difficulty because of not having a common culture which causes hindrances to choosing any such expression which could be ambiguous or difficult to interpret by the reader. The other possible explanation is that even after getting an education in the English language since primary grade, they don't have frequent chances to observe native users of this language. The results show that students were strictly using fixed grammatical units in their writing which they were taught during their academic learning. There are also some fixed expressions used by 2 students that could be justified that their knowledge of language is meeting CEFR Level B1 while the rest of the students do not have such mastery. The other formulaic language types like idioms and phrasal verbs seemed challenging. The students prefer using idioms instead of phrasal verbs as idioms have fixed meanings while phrasal verbs can be interpreted differently and hence ambiguous. The speech formulas and SBU were missing as according to learners' academic knowledge, they are

part of spoken discourse and their task was based on written discourse. This study acknowledged that students are only using those formulaic expressions which have a literal meaning and psychological salient in their intercultural environment.

5.2. Recommendation

The results of this study show that formulaic language is an important aspect of second language acquisition and should be part of the pedagogical framework. The use of social media has transformed the concept of communication and poses new challenges to being a competent and native-like user of the English language. The introduction of formulaic language should not be part of only speaking skills but also writing skills as well. This will improve communication and give an edge to the user to convey their message effectively.

REFERENCES

Gairns, R., & Redman, S. (1986). Working with Words. Cambridge: Cambridge University Press.

Goffman, E. (1974). Frame Analysis: An Essay on the Organization of Experience. New York: Harper & Row.

Graddol, D. (2006). English Next. London: The British Council.

Grice, P. H. (1972). Models of interaction of language and social life. In J. Gumperz & D. Hymes (Eds.), Directions in Sociolinguistics (pp. 35–71). New York: Holt, Rinehart and Winston.

Kecskes, I. (2000). A cognitive-pragmatic approach to situation-bound utterances. Journal of Pragmatics, 32(6), 605–625.

Kecskes, I. (2003). Situation-Bound Utterances in L1 and L2. Berlin: Mouton de Gruyter.

Kecskes, I. (2004). The role of salience in processing pragmatic units. Acta Linguistica Hungarita, 51(3–4), 309–324.

Kecskes, I. (2006). On my mind: Thoughts about salience, context, and figurative language from a second language perspective. Second Language Research, 22(2), 219–237.

Miller, J., & Weinert, R. (1998). Spontaneous Spoken Language: Syntax and Discourse. Oxford: Clarendon Press.

Recanati, F. (2001). What is said. Synthese, 128, 75–91.

Recanati, F. (2004). Literal Meaning. Cambridge: Cambridge University Press.

Searle, J. (1975). Indirect speech acts. In P. Cole & J. Morgan (Eds.), Syntax and Semantics, Vol. 3 (Speech Acts) (pp. 59–82). New York: Academic Press.

Searle, J. (1978). Literal meaning. Erkenntnis, 13, 207–224.

Seidlhofer, B. (2004). Research perspectives on teaching English as a lingua franca. Annual Review of Applied Linguistics, 24, 209–239.

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

Stalnaker, R. (1978). Assertion. In P. Cole (Ed.), Syntax and Semantics 9 (Pragmatics) (pp. 315–332). New York: Academic Press.

Tannen, D., & Oztek, P. C. (1981). Health to our mouths. Formulaic expressions in Turkish and Greek. In F. Coulmas (Ed.), Conversational Routine (pp. 516–534). The Hague: Mouton.

Wood, D. (2002). Formulaic language in acquisition and production: Implications for teaching. TESL Canada Journal, 20(1), 1–15.

Wray, A. (1999). Formulaic language in learners and native speakers. Language Teaching, 32, 213–231.

Wray, A. (2002). Formulaic language and the lexicon. Cambridge: Cambridge University Press.

Wray, A. (2005). Idiomaticity in an L2: Linguistic processing as a predictor of success. Paper submitted for Proceedings of 2005 IATEFL Conference.