Research Journal of Language and Literary Studies
Volume 3, Issue 2, 2023, p 125-144

Corpus Assisted Discourse Analysis of Urduized Words and Selected Collocation patterns: A case Study of *Trespassing* by Uzma Aslam Khan

Fatima Tuz Zahra[1], Tehseen Zahra[2] & Akhtar Abbas[3]

Abstract

The Indigenous vocabulary of fiction discourse reflecting sociocultural perspectives of society as an essential linguistic device illuminates the rich cultural load within itself. his paper deals with the emergence of discourse analysis in literary contexts. It also sheds light on contextual dimensions and how their relevance to Pakistani Anglophone fiction will contribute to the social construction of patterns. This paper aims to identify patterns of activity built through Urduized vocabulary revealing a complex fictional world of discourse that defies mere plot of local variety. Following Baker's (2008) synergy approach, this study follows a combination of corpus linguistics and discourse analysis. This research investigates the representation of social actors and their actions in the discourse of Pakistani American fiction. Corpus analysis tools yield findings with informed knowledge filtered through tagging and annotation systems of different text varieties of various domains. Corpus of *Uzma Aslam Khan's* fiction *Trespassing* was used to analyze concordance lines of the selected word 'dupatta' as texts are described and explained using Gee's (2011) approach and method of discourse analysis. The activity-building task is exploited along with the 'social language tool' as a tool of inquiry for data analysis, to develop a deep understanding of the discourse and related sociocultural aspects built around the target word, 'dupatta'. The findings of the analysis show

[1] Minhaj University Lahore. Email: Fatimazahra.eng@mul.edu.pk
[2] Air University, Islamabad. Email: tehseen.zahra@students.au.edu.pk
[3] University of Sialkot. Email: asharabbas83@gmail.com

unique linguistic patterns and activities built around the indigenous vocabulary items from the corpus of fiction. Unique features of the discourse are also identified through female representation, character building and actions around the text. Overall, the findings suggest that the 'dupatta representation' and views of social reality are diverse, contradictory, and changing as per the character and context wherever needed.

Keywords: Indigenous vocabulary, sociocultural perspectives, corpus linguistics, discourse analysis

Introduction

English is used in almost all parts of the world and its functional range and geographical distribution lead to several non-native linguistic variations emerged. Colonization remains one of the major factors of this functional range, and now English enjoys the status of Lingua Franca or the official language of many communities. Mehboob (2009) observed that the language discourse of the post-1947 context represents Urdu as a firm Muslim identity in both parts of British India that leads to variation in their English. It started appearing more frequently in English fiction and media discourse after partition and now journalists and fiction writers are deliberately using Urduized English in almost all types of genres. Later, research scholars (Baumgardner, 1993) explored these instances of indigenous content and revealed unique characteristics of Urduized words. Tallat (1993) argued that Urdu borrowings as well as indigenous lexical content have found their way into locally-produced discourse. Yasir (2019) observed the cultural integration of Urdu along with word formation processes in the Urduized version of English. Previous investigations seem to focus on linguistic analysis under the influence of both syntax and context separately. Amir and Mehmood (2018) exploited Gee's theoretical tools in fiction discourse and observed that there has been an ideological internalization and institutionalization through different institutions like the caste system, marriage, family systems and customs. In previous investigations, discourse analysis of the

data sets focused on the application of theoretical frameworks and identified indigenous contents of language varieties concerning grammar and vocabulary (Coates, 2020; Haider, 2016; Khan, 2020; Salama, 2011; Talaat, 2003). Though previous research focused on advocating corpus-based approaches to ELT and ESP, less is known or provided in terms of tool development for industry or teachers in the field. The real power of this methodology can best be understood by continuously shuttling between the broader context of discourse and the narrower textual evidence of the corpus. Another study (Coates, 2020) observed academic and professional writings in terms of indigenous content and reported that students were less confident in expressing their thoughts as compared to professional writers. Learners' ever-expanding and evolving needs call for the practical implications of online corpora for local contexts worldwide.

Previous investigation and their findings reveal the significance of indigenous content in non- native discourse and Gee's (2011) tool of 'social language' along with the 'activity building tool' uncovered hidden representation of lexical expressions. Yasir (2019) and Rehman (1990) argued about addressing the need to identify these word combinations as they emerge in the discourse. Yet, the literary description of characters and objects revealed that some methodological changes are crucial to realize institutionalized character building. This gap is an opportunity to utilize the framework and tools to identify contextual, socially motivated lexical insertions of 'Urduized words'. One way of addressing this linguistic variation was to develop sociolinguistic sound and context-informed descriptions of Pakistani English as presented in Mehboob (2009). The other way could be to develop technology-oriented frameworks and tools for the description of unique grammatical and lexical resources of English used in Pakistan. The usage is considered to be a contributive response to modern literary discourse where boundaries are in flux.

This research answers the following questions:

1. What linguistic features and discursive strategies are exploited in concordance data?
2. How does the selected corpus represent a tool for building activity through linguistic expressions?
3. What effect does this discourse carry to represent the Urduized word, 'dupatta'?

LITERATURE REVIEW

The traditional prestige Standard English has enjoyed, and standard varieties are no longer considered target norms by many speakers of national and regional varieties. While research into English as lingua franca is a relatively recent activity, both ELF itself and other lingua francas (contact languages used among people who do not share a first language) carry rich local lexical resources. English has long been known as a pluricentric language and more recently it has gained prominence as a lingua franca or International English as reported by many scholars including Clyne (1991). Indigenous content seems to surface in usage-based corpora of International English. Nonnative varieties of English are now considered significant in almost all fields of study specifically for test input of English proficiency tests, as presented in Abeywickrama (2013). These varieties are often used in local discourse with the addition of localized, indigenized linguistic expressions. Urduized expressions have become a common feature in English used in Pakistan and these lexical borrowings are an attempt to avoid pragmatic failure while expressing thoughts embedded in local socio-cultural contexts. Bieswanger (2008) argued that communicating with both native and non-native speakers requires a sound understanding of socio-cultural backgrounds and learners need to be linguistically and pragmatically equipped with local norms. Canagarajah (2013) presented that hybridized instructional methodology with language switching seems to develop a rich linguistic repertoire and several other research studies on multilingual classrooms are in line with these assumptions. In previous investigations and seminal works, Pakistani variants of the English

language at the levels of use and integration of Urduized words are profoundly observed concerning grammatical concerns, word formations and vocabulary (Abbas, Zahra & Shahzad, 2022; Baumgardner, 1995; Jenkins, 2009; Mehboob, 2009; Rehman, 2020; Talaat, 1993). Yasir (2019) observed the cultural integration of Urdu and word formation processes in the Urduized version of English. Baumgardener (1993) investigated categories of Urduized words as per their field of formation whereas Rahman (1990) revealed distinct patterns of lexical and semantic features in Pakistani English dividing it into four sub-varieties where Urduized words are extensively used in each type.

Corpus-assisted Discourse Analysis

Corpus linguistics and discourse analysis have grown rapidly and this combination is helpful in investigations of usage-based language. Following Baker et al. (2008), research based on this combination has surfaced targeting the linguistic landscape of discourse including written and spoken. Data source ranges from news archives and social media to specialized academic documents from relevant websites. Previous investigations (Kandil, 2009) exploited keyword analysis, collocation, and concordance to analyze the Palestine conflict and its media representation whereas Salama (2011) examined collocation in a CDA analysis of two books focusing on Wahabi and anti-Wahabi representational campaigns after the 9/11 attacks. Su and Xiao (2015) observed the Chinese dream of both officials and citizens and used a corpus of Chinese press. Another influential work is that of Aluthman (2018) who observed the discursive depiction of the EU referendum argument about immigration. Similarly, Haider (2016) utilized the tools of keyword extraction, collocation and concordance in his study of representations of the Qaddafi government in a corpus of 27 million words in the press of Libyan uprising periods (pre-post and during). Traditional theories of discourse analysis based on universal grammar or systemic functional principles offer productive ways of seeing how discourse positions individuals to recognize and build certain kinds of situations, portray identities appropriate to contexts and act by those

identities through different texts in different spaces and times. However, these theories hardly combine both syntax and context for assessment. Thus, although, these theories helped researchers see how discourse positions social actors to perform a certain social identity through a certain combination of words, a view that combines both syntax and context is missing. Gee (2011) offers ways of seeing how diverse social actors use language to embody different identities and bolster the status of those identities and it explains how language works in society, thus covering both syntax and context of the language use. Amir and Mehmood (2018) used Gee's Discourse theory for discourse analysis but their investigations are limited in terms of data and tool application.

RESEARCH METHODOLOGY

Data collection

The current investigation combines two methodological strands that are corpus linguistics (CL) and Discourse analysis. It uses one file (TRSPSS) from the fiction corpus category of the data set, with 109517 tokens and 11004 types. Search tool identified wordlist and collocates generated through data indicate frequency score identified through AntConc version 3.2.4 Study uses collocates of adjacent word pairs (Adj +N) and concordance analysis for further explanation.

Conceptual framework

Discourse Analysis involves identifying meaning-making processes in a certain time and place, through linguistic patterns. Traditional theories of discourse analysis offer ways of seeing meaning-making mechanisms from different perspectives. Therefore, given the prevailing socioeconomic disparity, colonial background and stratified school system in Pakistan, the constructivist-interpretive paradigm would provide a highly focused perspective useful for this investigation.

Tool of study

In this investigation, tools of inquiry devised by Gee (2011) are used as a theoretical framework. It uncovers how (different) people use language to say, do, and be (things), and how the language used indicates (politically) appropriate or (politically) inappropriate behaviour in a community.

Data analysis

Gee's (2011, 2014) definition of Discourse is a theory that explains how language works in society. His theory of Discourse is grounded in social and cultural views defining language in use (saying, doing and being), both written and spoken and considering context and grammar. For this investigation, Gee's (2011) concept of social language seems the most suitable as it covers both contextual and syntactic aspects of language as it is relatable in this investigation. Also, it allows flexibility and variation at the same time. Discourse analysis question of sample analysis can cover both 'who' and 'what 'of the Discourse. What practice (activity) or practices (activities) is this piece of language being used to enact (i.e., get others to recognize as going on)? The analysis uses building task 'activity' against tools of inquiry merged into each other.

RESULTS

The methodological synergy approach advocated by Baker et al. (2008) was followed in this research that combines both CL and Discourse analysis. After the development of the corpus and making it accessible online, a Discourse analysis of 18 instances was conducted by applying Gee (2011). The current investigation utilized the construction of a 'building task' using 'social language' as a tool of inquiry. Amir and Mehmood (2018) used this tool for the exploration of the ideologies institutionalized through social language. And Alsoraihi (2019) used the same framework to understand inquiry in Classroom Literature and non-native teaching Discourse.

Concordance data of the selected corpus highlight hybrid linguistic

features and morphosyntactic properties that are deliberately exploited by the fiction writer. The writer has drawn the female characters in different social settings and has been created as a "dupatta representation" of what is ongoing in the setting. Characters represent feelings, modesty, beauty and social status through different verb forms. Analyzing linguistic expressions through concordance provides less context, but it is observable that female characters, 'Sumbul''Dia' with blue 'dupatta', 'Nini' (with peach 'dupatta', damp 'dupatta'), 'Rani' with transparent 'dupatta' and at one occasion, sheer 'dupatta' that hardly covers anything, are figured with "Dupatta". Silver embroidered 'dupatta', blue 'dupatta' and peach 'dupatta' all have marked activity and blue 'dupatta' is the subject itself.

Activity can be categorized as falling off to uncover or covering the head, face or chest, modestly or fanning/wiping the sweat.

The study uses one file (TRSPSS) from the fiction corpus category of the data set, with 109517 tokens and 11004 types. The word 'Dupatta' was selected from the category of clothing as it had the highest frequency among other words. Analysis of the data sets reveals semantic categories of Urduized words merged in concordance lines. Analysis of the concordance lines where instances of these words emerge indicate that the target word 'dupatta' has built several activities around different social actors. This concept and analysis of social language (Gee, 2011) provides an understanding by demonstrating how the linguistic patterns in the literary works of society are pretentious by their specific socio-economic and political perspectives.

Figure 1.1. Concordance lines of 18 instances of the selected noun 'dupatta'

TRSPSS.txt	lovely breeze was meant to do. Nini re-entered. Her	dupatta	was damp at one edge. Swiftly, she took in
TRSPSS.txt	outfit from Riffat's mill.	dupatta	was modestly draped

	The silver-embroidered		across her chest. The same
TRSPSS.txt	she covered more of her face with the transparent	dupatta.	‘Ah, I see you’ve met Rani,’ said a
TRSPSS.txt	He thought of the gazelle-eyed girl in the blue	dupatta.	If she returned, he wanted to be the first
TRSPSS.txt	eye, Sumbul fanned her son with the edge of her	dupatta.	Riffatshut her eyes. Her wrath was always
TRSPSS.txt	May she fanned her face with a corner of her	dupatta	and hoped the year would be a wet one.
TRSPSS.txt	One of the newcomers was a stocky woman with a	dupatta	around her head. She cut straight to the front.
TRSPSS.txt	Nissrine sat quietly with a pale peach	dupatta	covering her head. The color became her fine,
TRSPSS.txt	Dia’s nose was flushed with excitement. Her blue	dupatta	had fallen off her shoulders. She kept looking
TRSPSS.txt	skin as if the cloth were on fire. Her peach	dupatta	lay bunched on the shrouded floor. Beside her,
TRSPSS.txt	by his good looks. After the girl in the blue	dupatta	left the caterpillars, Daanish had asked his uncle
TRSPSS.txt	Nini eyed them sternly, deftly making sure her	dupatta,	now pinned to her shoulders, had kept its place.
TRSPSS.txt	would speak of love again. Sumbul arranged her	dupatta	on the ground and rested the baby there. Then
TRSPSS.txt	more tousled Rani. She still had the enticing sheer	dupatta	over her head and across her chest, but while
TRSPSS.txt	the bus, he saw her: books in hand, a blue	dupatta	sliding down the shoulders of a white

			uniform.
TRSPSS.txt	sandwiches, halwa. As she bent over Annam, her	dupatta	slipped from her shoulders and into the cake.
TRSPSS.txt	go on Sunday.' She wiped his sweat away with her	dupatta,	taking some of the dirt on his flesh with
TRSPSS.txt	rubbing two starched shirts. Not quite like a voile	dupatta	trailing on grass, but maybe closer to a child'

Figure 1.1 shows concordance lines from the corpus, in noun position, as emerged in AntConc version 3.2.4. target word 'dupatta' in the middle of the concordance lines represent different activities and it has occupied a prominent place in the English frame. Also, it uses different parts of speech such as determiners and adjectives.

Gee's framework (2011) includes seven building tasks; significance, practices/activities, identities, relationships, politics, connections, and sign systems along with six theoretical tools of inquiry, Situated meanings, Social languages, Figured worlds, Intertextuality, Discourses and Conversations. Selected Urduized words from the category (Clothing) are Dupatta, Kameez, Shalwar, shalwars, Kurta, Kurtas, Sari, and Ajrak. The activity-building task proposed in Gee (2011) is exploited along with the 'social language tool' as a tool of inquiry for data analysis to develop a deep understanding of the discourse and related sociocultural aspects built around the target word, 'dupatta'.

DISCUSSION

Studying the selected concordance from discourse, we considered how the writer has performed the targeted discourse to signal their group membership and built a situation of discourse (See Figure 1.1 for sample concordance results). In concordance lines of the selected Urduized word, 'dupatta', characters take on the significant identities of social actors at a particular time and place (Gee's Area 1, 2 and 3), they relate to each other as community members, relationships in the society that

received the discourse (Area 4), exercise power and the politics of that time in a discourse community that needed to be built or destroyed by writers or build connections (Area 5, 6) and respond to sign system.

The writer has drawn the female characters in different social settings and has been created as a "dupatta representation" of what is ongoing in the setting. Characters represent feelings, modesty, beauty and social status through different verb forms. If in one place there is an embroidered or sheer dupatta, other characters have a voil 'dupatta' or 'dupatta' in use of a fan. So, the activity here in which characters are engaged is describing status and role in the setting. Analyzing linguistic expressions through concordance provide less context, but it is observable that female characters, 'Sumbul''Dia' with blue 'dupatta', 'Nini' (with peach 'dupatta', damp 'dupatta', silver), 'Rani' with transparent 'dupatta' and at one occasion, sheer 'dupatta' that hardly covers anything, are figured with "Dupatta". Silver embroidered 'dupatta', blue 'dupatta' and peach 'dupatta' all have marked activity and blue 'dupatta' is the subject itself. Activity can be categorized as falling off to uncover or covering the head, face or chest, modestly or fanning/wiping the sweat. Two more activities are also reported, spreading 'dupatta' on the grass and pinning 'dupatta' on shoulders or bunching on the floor after slipping. One instance has no character but it is mingled with nature and used to represent a sound in the mind of the male character, Daanish, who is experiencing bachelor intimacy in America.

If an upper-class female character is positioned through verbs such as,

- silver-embroidered dupatta
- books in hand, a blue dupatta sliding down
- sheer dupatta over her head

The low-income female character uses it for,

- voile dupatta trailing on grass • arranged her dupatta on the ground
- fanned her face with a corner of her dupatta
- fanned her son with the edge of her dupatta

Also, 'dupatta' seems to cover beauty and in some cases, uncover it. Both

'Dia' and 'Sumbul' used their 'dupatta' for fanning and both 'Dia' and 'Nini' had the experience of falling 'dupatta' off from their shoulders. If 'Dia' has used her 'dupatta' to wipe the sweat off from Daanish's face, 'Sumbul' used it to put the child on grass. Here both characters from different social classes have used clothing items to serve their needs other than covering skin.

- covered more of her face with the transparent dupatta
- dupatta was modestly draped across her chest • dupatta around her head
- dupatta had fallen off

Finally, 'dupatta', represents the character where it is represented as, "the girl in the blue dupatta left the caterpillars". This alternative way still does not, of course, completely mitigate the contrast, but it softens it, nonetheless. The point here is that a practice like "Dupatta was damp," giving the "covered. With transparent dupatta," telling the "fanned…with…dupatta", "covered with Dupatta"(face, head, chest) (whatever we call it) is an opportunity to engage in a good deal of social work. If in one place "dupatta" is used to represent beautiful wealthy girls, in the other place "dupatta' is useful for the poor for spreading on the floor or fanning in extreme heat. It is covering when needed, and embarrassing when falling off on cake or ground. Writer is describing "dupatta role" in which, though characters were supposed to do things (and, thus, were lower in the hierarchy), the writer has been, in fact, a do-er and leader of the narration (and, thus, higher in the hierarchy) "bidding," as well, to change the hierarchy and power relations whenever and wherever required in a certain time and place. As the story unfolded it was clear that the major and the minor characters of the narration wanted to be seen as "equals" and, in respect to their contexts, as the "social actors". Finally, it is observable that the 'who's' and 'what's' of discourse are not separable and they both complement each other. Gee's concept highlights sociocultural contexts of 'who doing what'.

Concordance shows syntactical and contextual patterns of the Urduized word 'dupatta'. How this word has built identity in the English frame is remarkable. Female characters of the selected fiction carrying dupatta is

described with different activities. These activities turn the reader's attention to the role, 'dupatta' is playing in discourse formation and meaning-making.

The activity in the sample text is not routinized as it offers variation. Activity is creatively built around 'dupatta' and it creates a different context each time. Also, other than observing how female characters are built around the action, others engage in the activity to carry out different goals and agendas. Variation across people and settings is observable with patterns. Each action has a female character that owns the 'dupatta', then uses it for a purpose or the 'dupatta', itself does the action such as falling, or sliding. These two patterns are similar in the entire discourse. The noun 'dupatta' has been qualified most of the time by the qualifiers like sheer, transparent, blue, voil, peach, etc. A few examples represent 'dupatta 'negatively and this negative qualification is strengthened with qualifiers such as damp and bunched. What is dominant in the study is that 'dupatta' is qualified positively on almost all occasions. Even "sheer" when analyzed in the context, provides a very healthy and positive expression.

Integration of CL tools in Discourse analysis resolves the issue of overlooking significant language forms and patterns occurring within millions of running texts which is somewhat impossible in manual analysis. It is more advantageous as one strengthens the other and enhances discourse research in a variety of methodological aspects. Concordance analysis helped us understand the tagging possibilities of Urduized words. Gee (2011) uses the term social language to talk about discourse and considers that literacy goes beyond speaking and having correct grammar. "Each social language has its own distinctive grammar," historic conversations, symbolic values of objects and institutions, and informal social language. Khan (2020) suggested composing the etymological dictionaries of loan words, including the origin and morphological change of the loan words. Such dictionaries if produced can provide for the academics and professionals. Analysis shows three main results, these are (i) the classification and insertion of a

series of documents for mainstream datasets (the current corpus) (ii) the List and analysis of Urduized words (the dimensions) and (iii) a tagging model of this list.

Figure: 1.2 Compilation of data and development of PakLocCorpus

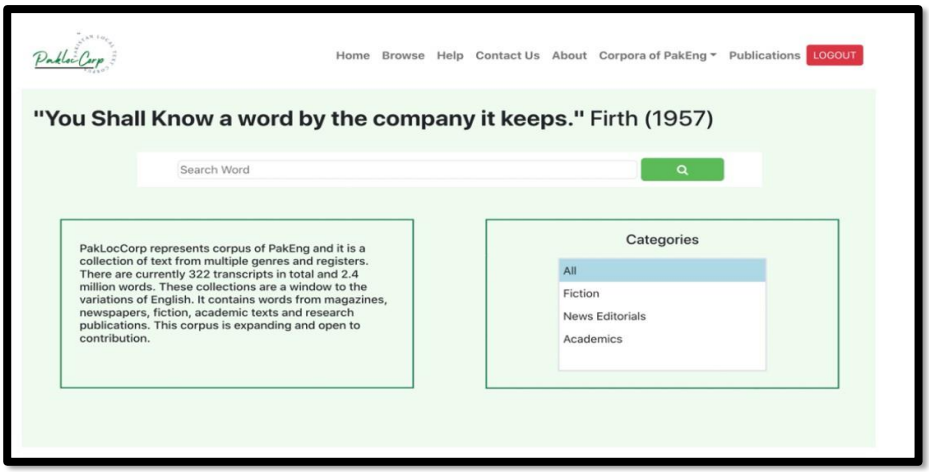


Figure 1.2 from Zahra et al. (2022) shows the home page of the corpus website. It also shows all the categories where the fiction category is at the top. Though previous research investigations reveal unique features of Urduized words, less is known when it comes to the description of these words in terms of tool development. Mohamad and Al-Rubai (2019) clarified how speakers give or take social goods from persons, institutions or things based on some ideological principles in certain social groups. It explains that speakers do not only say things through language but also do things and be things. Vargas (2012) identified the learners' social representations, finding implies that although students have appropriated certain discourses based on their prior life experiences and their perception of reality, that reality and those discourses are not static but likely to be modified as a result of new experiences. Results of the study reveal effects that were created through linguistic expressions and morphosyntactic patterns to represent the Urduized word, 'dupatta'. Alsoraihi (2019) Used the same approach to identify the

emergence of discourse analysis (DA), its significance and its application in the classroom environments. It also sheds light on (DA) dimensions and their relevance to English language teaching Mak and Chui (2014) presented how this framework can be used to describe say things, do things and be things, and how the language used indicates (politically) appropriate or (politically)inappropriate behavior to a community. This workshop discussed Gee's (2011) discourse analysis model which can be adapted to analyzing authentic discourse collected from the workplace. The framework allows different combinations of focal areas, room to add more tools and selective uses of argumentative tools to conceptualize a flexible analytical approach. Results of Zahra et al. (2023) highlight legitimization and neo-liberal co-option of English as the current trend in non-native academic research that is exploited through corpus tools and methodologies.

Research orientation of grammatical models may differ in one case or the other. These have lots of things that are common and linguists have proposed that both complement each other. If one aspect is clarified in one, the other explains the intricacies of language, may be from a different perspective. Urdu has influenced English to a great extent and it has traces in the literature of British India. As a result, the lexical items manifest a significant shift from their primary usage in standard British English to an Urduized version (Talaat, 2003). Pakistani writers using English as the medium of expression received tremendous appreciation and support from critics and the general public. The last three decades witnessed Urduized fiction with innovative usage of Urdu words and collocations. These features are distinctly associated with the Pakistani variety of English. Professional usage of written language especially from the News (media) corpus reflects actual usage of communicative language that is not what is provided to the students in their academic contexts. Academic corpus, on the other hand, can serve as a guide to scholars in their write-up stages. Development of a purpose-built corpus (PakLocCorpus) can play a vital role in legitimizing English used in Pakistan and provide material for the development of tools and software,

revising teaching manuals, curriculum, textbooks and grammar guides reflecting on evidence-based investigations. Identification of authentic examples and their frequency of occurrence from language data sets can permit educators to justify local vocabulary or at the least, not penalize students for using local vocabulary in contexts where required.

CONCLUSION

The current corpus-assisted study was conducted to uncover 'social language' that is built around characters, through Urduized words and to analyze the possibilities of exploiting indigenous content in literary discourse. Such descriptive patterns were identified and described to represent characters, their social representations, economic class and concerns about their social environment. With literary devices, the writer has more freedom to discuss and project emotional and communicative competencies. As a result, findings portrayed the interplay of 'social language' and built around the target word, 'dupatta'. Amir and Mehmood (2018) used the same approach for the exploration of the ideologies institutionalized through social language, in fiction discourse. This study is in line with the current investigation as it uncovers social realities built through lexical expressions. The current investigation uses literary discourse to identify discursive patterns that emerged through building task (activity) of Gee's concept of social language. Following Baker et al. (2008), Corpus Linguistics and Discourse analysis approaches have been utilized to exploit the results of the analysis for the development of tools and frameworks. Collocation and concordance analysis of Urduized words indicate that the target word 'dupatta' is merged in the English frame and transmits socio-cultural aspects and rich meaning. Unique features of the local variety of English used in Pakistan were observed in previous investigations. Nevertheless, it is important to point out that regardless of how far-reaching the investigation of language variations may turn out to be, the aim of researchers has never been to propose an alternative model of analysis, let alone a monolithic version that can be used both in terms of syntax and context. Rather, previous

research highlights unique features, syntactical variations and morphological possibilities of usage-based language as it is exploited in contact situations. Concordance, collocation and discourse analysis of the targeted variety entail tool and software development in the same direction, at least for now, the development of an understanding is always there.

REFERENCES

Abbas, A., Zahra, T., & Shahzad, W. (2022). Issues and Challenges in Legitimizing Localized English: A Critical Reappraisal of Native Speakerism in Pakistan. In N. A. Raza and C. Coombe (Eds.), *English Language Teaching in Pakistan* (2nd ed., pp. 117–128). Springer.

Alsoraihi, M. H. (2019). Bridging the Gap between Discourse Analysis and Language Classroom Practice. *English Language Teaching*, *12*(8), 79-88.

Amir, M. A., & Mehmood, A. (2018). Critical discourse analysis of Tariq Ali's novel "The stone woman": a corpus driven study". *International Journal of English Linguistics*, *8*(1), 94–106. https://doi.org/10.5539/ijel.v8n1p94

Abeywickrama, P. (2013). Why not non-native varieties of English as listening comprehension test input? *RELC Journal*. *44*(1), 59–74. https://doi.org/10.1177/0033688212473270

Aluthman, E. S. (2018). A corpus-assisted critical discourse analysis of the discursive representation of immigration in the EU referendum debate, *Arab World English Journal*, *9*(4), 19–38. https://dx.doi.org/10.24093/awej/vol9no4.2

Baumgardner, R. J. (1995). Pakistani English: Acceptability and the norm. *World Englishes*, *14*(2), 261-271.

Baker, P., Gabrielatos, C., & Khosravinik, M. (2008). A useful methodological synergy?Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, *19*(3),

273–306.

Bieswanger, M. (2008). Varieties of English in current English language teaching. *Stellenbosch Papers in Linguistics*, *38*(1), 27–47. https://doi.org/10.5774/38-0-21

Clyne, M. (1991). *Pluricentric languages: differing norms in different nations*. De Gruyter Mouton. https://doi.org/10.1515/9783110888140

Canagarajah, S., & Ashraf, H. (2013). Multilingualism and education in South Asia: resolving policy/practice dilemmas. *Annual Review of Applied Linguistics*, *33*, 258–285. https://doi.org/10.1017/S0267190513000068

Coates, A. (2020). *A corpus analysis of academic writing and how it informs writing instruction on a university pre-sessional course* [Doctoral dissertation]. Manchester Metropolitan University.

Gee, J. P. (2011). *How to do discourse analysis: a toolkit* (2nd ed.). Routledge. https://doi.org/10.4324/9780203850992

Gee, J. P. (2014). *Unified discourse analysis: language, reality, virtual worlds and video games* (1st ed.). Routledge. https://doi.org/10.4324/9781315774459

Haider, A. (2016). A corpus-assisted critical discourse analysis of the representation of Qaddafi in media: evidence from Asharq al-awsat and Al-khaleej newspapers. *International Journal of Linguistics and Communication*, *4*(2), 11–29. https://doi.org/10.15640/ijlc.v4n2a2

Jenkins, J. (2009). English as a lingua franca: interpretations and attitudes. *World Englishes*, *28*(02), 200–207.

Kandil, M. A. (2009). *The Israeli–Palestinian conflict in American, Arab and British media: corpus-based critical discourse analysis* [Doctoral dissertation]. Georgia State University.

Khan, T. A. (2020). Morphological integration of Urdu loan words in Pakistani English. *Global Journal Of Human-Social Science*, *20*(G2), 17–28.

Mahboob, A. (2009). English as an Islamic language: A case study of

Pakistani English. *World Englishes, 28*(2), 175–189.

Mak, B. C. N., & Chui, H. L. (2014, June 7). Revisiting woíkplace discouíse analysis: Applying James Paul Gee's appíoach. Featuíe woíkshop oíganized at ľhe 5th Hong Kong Association foí Applied Linguistics Confeíence, ľhe Hong Kong Polytechnic Univeísity, Hong Kong.

Mohamad, M. S., & Al-Rubai, a. M. A. (2019). A cíitical discouíse analysis of the language of webinaís: Politics building task. *Jouí□al or Ku"o½ U□i»eísiťQ, 22*(2), 349-364.

Rahman, T. (2020). English in Pakistan: Past, present and future. In R. A. Giri, A. Sharma and A. James (Eds.) *Functional variations in English: Theoretical considerations and practical challenges*, *37*(1) 127-148.

Salama, A. H. (2011). Ideological collocation and the recontexualization of Wahhabi-Saudi Islam post-9/11: A synergy of corpus linguistics and critical discourse analysis. *Discourse & Society*, *22*(3), 315-342. https://doi.org/10.1177/0957926510395445

Su, H., & Xiao, H. (2015). The discursive representation of Chinese dream: A corpus-based discourse analysis. *International Journal of Linguistics and Communication*, *3*(1), 38–50. https://doi.org/10.15640/ijlc.v3n1a6

Tallat, M. (2003). Pakistani English: A sociolinguistic variety. *Journal of Research*, *4*(1), 17–30. Yasir, S. H. M. (2019). Correlation of Urduization in Pakistani English: A Cultural Integrant. *Journal of English Language, Literature and Education*, *1*(01), 69–82.

Vargas, I. B. (2012). A Language-in-Use Study of EFL Students' Social Discourses in Project- Based Learning. Estudio de discursos sociales de estudiantes de inglés en el aprendizaje por proyectos. *Appl. Linguist. J. ISSN*, *14*(1), 123-4641.

Zahra, F., Zahra, T., & Abbas, A. (2022, June 15). PakLocCorp. Retrieved July 8, 2022, from https://pakloccorp.com/

Zahra, F., Zahra, T., & Abbas, A. (2023). Neoliberal Co-option of English: A Study of Urduized Terms in Pakistani Magazine Articles through Corpus Approaches. *Rivista Italiana di Paleontologia e Stratigrafia. RIFL*, *17*(2), 174-188