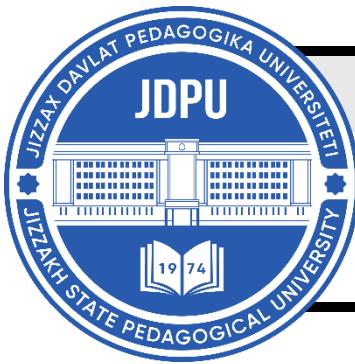


# MENTAL ENLIGHTENMENT SCIENTIFIC – METHODOLOGICAL JOURNAL



## MENTAL ENLIGHTENMENT SCIENTIFIC – METHODOLOGICAL JOURNAL

<http://mentaljournal-jspu.uz/index.php/mesmj/index>



### REVEALING SOCIAL IDENTITY THROUGH DIALECT- SPECIFIC LEXICAL UNITS: A CORPUS-BASED ANALYSIS OF ENGLISH AND UZBEK

***Kamola Malikovna Jalolova***

*Assistant at Karshi State Technical University*

[kamola210188@gmail.com](mailto:kamola210188@gmail.com)

*Karshi, Uzbekistan*

***Shakhnoza Jurayeva***

*Assistant at Karshi State Technical University*

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

*Karshi, Uzbekistan*

***Mukhayyo Tadjiyeva***

*Senior Lecturer at Karshi State Technical University*

[muxayyotadjiyeva2@gmail.com](mailto:muxayyotadjiyeva2@gmail.com)

*Karshi, Uzbekistan*

#### ABOUT ARTICLE

**Key words:** sociolinguistics, dialectal lexis, social identity, corpus linguistics, English language, Uzbek language.

**Received:** 21.01.26

**Accepted:** 22.01.26

**Published:** 23.01.26

**Abstract:** This article provides a comparative analysis of the sociolinguistic characteristics of dialect-specific lexical units in English and Uzbek and their role in shaping and expressing social identity. The study is conducted on the basis of corpus linguistics and sociolinguistic approaches, using data from the British National Corpus (BNC), the Corpus of Contemporary American English (COCA), and the Global Web-Based English (GloWbE) for English, as well as electronic texts, dialectological sources, and samples of authentic spoken discourse for Uzbek. The results of the study show that dialect-specific lexical units in both languages serve as important linguistic indicators of a speaker's regional, social, and cultural affiliation. While dialect vocabulary in English is more closely

associated with age and social group factors, in Uzbek it primarily functions to reinforce regional and cultural identity. The scientific novelty of the article lies in interpreting dialect-specific lexis as an active constructor of social identity and in applying a comparative corpus-based approach to English and Uzbek.

---

**Introduction.** In the twenty-first century, the intensification of globalization and migration processes, the expansion of digital communication, and the rapid development of intercultural contacts require a fundamental reconsideration of the role and functions of language in society. Language is no longer viewed solely as a means of information exchange, but rather as a complex social mechanism that shapes and expresses social identity [1]. In particular, the variable forms of language—dialects, vernaculars, and sociolects—are increasingly becoming significant markers that reflect an individual's position in society, group affiliation, and cultural experience [2].

Although the phenomenon of dialect has traditionally been explained in sociolinguistics primarily through territorial differentiation, recent studies interpret it as a phenomenon closely connected with social consciousness, identity, and discourse [3]. According to contemporary scholarly views, dialect-specific lexical units reveal not only a speaker's geographical origin but also their age, profession, social status, cultural capital, and communicative strategies [4]. For this reason, dialectal lexis is today recognized as one of the most important objects of sociolinguistic analysis.

In world linguistics, the study of the relationship between language and social identity entered scientific discourse with W. Labov's theory of variationist sociolinguistics. Subsequently, J. Fishman explored the relationship between language and ethnic identity, P. Trudgill focused on dialect and social stratification, and R. Wardhaugh systematically examined the interrelations between language, culture, and society [5, 6, 7]. In these studies, dialects are not regarded as "lower" or "deviant" forms of language, but rather as linguistic resources endowed with social meaning. In particular, the expression of identity through regional and social dialects of English has been extensively investigated on the basis of broad empirical data.

In recent years, the integration of sociolinguistics with corpus linguistics has been steadily increasing. A corpus-based approach makes it possible to identify the actual use of linguistic units in real communicative environments, as well as to analyze their contextual and pragmatic features. In English linguistics, large corpora such as the British National Corpus (BNC), the Corpus of Contemporary American English (COCA), and the Global Web-Based English (GloWbE) have enabled in-depth analysis of the social functions of dialect-specific

lexical units, their role in shaping identity, and their discursive load [8]. These studies demonstrate that dialectal lexis is often used as a means of creating social closeness and solidarity or, conversely, of establishing social distance.

In Uzbek linguistics, the social nature of language has a strong scholarly tradition. Scholars such as A. Abduazizov, N. Uluqov, Z. Xolmanova, A. Nurmonov, and B. Mengliyev have theoretically examined the relationship between language and society [9, 10, 11, 12]. Uzbek dialects have mainly been studied from dialectological, historical, and phonetic perspectives, and their lexical composition, territorial variation, and sources of formation have been identified [2, 10]. However, from the standpoint of modern sociolinguistics, issues such as the role of dialectal lexis in expressing social identity, its functions in communicative strategies, and its place within social discourse have not yet been comprehensively investigated.

In particular, a comparative sociolinguistic study of English and Uzbek is of great relevance. As these languages belong to typologically different systems, the social functions of dialect-specific lexical units may display both similarities and differences. This, in turn, can lead to important scientific insights into linguistic universals and cultural specificity. At the same time, such comparative analysis contributes to advancing sociolinguistic research in Uzbek linguistics to a new level.

**Methodology.** This study was conducted by integrating sociolinguistic and corpus-linguistic approaches. The methodological framework of the research is grounded in contemporary sociolinguistic theories aimed at identifying the functional characteristics of linguistic units that express identity and the social variability of language.

Research material. The research material consists of dialect-specific lexical units in English and Uzbek. Empirical data for English were compiled on the basis of the following open and widely used language corpora [13]:

- British National Corpus,
- Corpus of Contemporary American English,
- Global Web-Based English.

From these corpora, lexical units related to regionally and socially marked dialectal vocabulary (for example, local words, units characteristic of informal speech, and lexis associated with specific social groups) were selected.

Materials for Uzbek were collected from the following sources:

- existing electronic and written text databases of the Uzbek language,
- dialectal units recorded in dialectological studies,

- samples of dialect-specific speech found in mass media, social networks, and literary texts.

These sources made it possible to observe the actual use of dialectal lexis in Uzbek and to identify its social context.

Research methods. During the research process, the following scientific methods were employed:

The sociolinguistic analysis method was used to identify the relationship between dialect-specific lexical units and social factors such as age, social group, regional affiliation, and communicative situation. Through this method, the role of lexical units in expressing social identity was revealed [14].

The corpus linguistics method was applied as the primary method for analyzing the frequency of use, context, and functional characteristics of the selected lexical units in authentic speech. Corpus data served to determine the activity of dialectal units in contemporary language use.

The comparative (contrastive) method was employed to identify similarities and differences between dialect-specific lexical units in English and Uzbek and to compare their sociolinguistic functions [15]. This method made it possible to identify the mechanisms of identity expression in the two language systems.

Contextual-semantic analysis was used to determine the meaning of dialectal units in speech, their pragmatic load, and their discursive functions. Through this analysis, the functions of lexical units in expressing social closeness, group affiliation, or informality were identified.

Selection criteria. Dialect-specific lexical units were selected on the basis of the following criteria:

1. the unit is regionally or socially marked;
2. the unit is restricted by the literary norm or characteristic of informal speech;
3. the unit is actively used in expressing social identity;
4. the unit is contextually identifiable in corpus materials.

Methodological novelty. The methodological novelty of the study lies in interpreting dialect-specific lexical units as linguistic indicators of social identity and examining them on the basis of a comparative corpus-based approach using English and Uzbek as examples. Unlike traditional dialectological studies, this approach makes it possible to identify the social functions of dialectal lexis in authentic language use.

**Results.** The results of the present study made it possible to identify the sociolinguistic characteristics of dialect-specific lexical units in English and Uzbek in expressing social identity.

The analysis conducted on the basis of corpus materials demonstrated that dialectal lexis is used regularly and functionally in actual speech practice.

Materials drawn from the British National Corpus (BNC), the Corpus of Contemporary American English (COCA), and the Global Web-Based English (GloWbE) indicate that dialect-specific lexical units in English are predominantly used in informal communication, everyday conversations, and contexts expressing social closeness. These units often serve to express the speaker's regional affiliation (for example, British or American regions), social group membership, and cultural identity.

The corpus analysis revealed that dialect-specific words in English occur more frequently in youth speech, communication within local groups, and informal discourse. Certain lexical units are used to enhance social solidarity, create closeness between interlocutors, or emphasize belonging to a particular group. The results demonstrate that dialectal lexis in English constitutes one of the clear linguistic indicators of social identity.

The analysis of Uzbek-language materials confirmed the wide and active use of dialect-specific lexical units in authentic communicative environments. Based on data obtained from mass media, literary works, recordings of spoken discourse, and social networks, it was established that dialectal words in Uzbek are mainly used in oral communication, local settings, and informal communicative situations.

The results of the analysis show that in Uzbek, dialectal lexis expresses not only the speaker's regional origin but also cultural closeness, national traditions, and social solidarity. Certain dialect-specific units function to reduce social distance between interlocutors, enhance sincerity, or strengthen interaction within a group.

A comparative analysis of corpus materials in English and Uzbek revealed the following common and distinctive features. First, in both languages, dialect-specific lexical units function as important linguistic means of expressing social identity, indicating a speaker's affiliation with a particular social or regional group. Second, while dialectal lexis in English is more closely associated with social stratification and age-related factors, in Uzbek it is primarily linked to regional and cultural identity. Third, dialectal units in English are actively used in written and digital communication (online platforms, blogs, and social networks), whereas in Uzbek their primary domain of use remains oral speech.

The obtained results demonstrate that dialect-specific lexical units in English and Uzbek are not random speech phenomena but rather constitute a stable mechanism for shaping and expressing social identity. The corpus-based analysis provides clear empirical evidence of the social functions of dialectal lexis.

These findings make it possible to interpret dialectal lexis from a sociolinguistic perspective, to evaluate it as a social indicator, and to establish a scientific foundation for further analytical discussion.

**Discussion.** The results of this study confirm that dialect-specific lexical units in English and Uzbek constitute an important sociolinguistic tool for expressing social identity. The empirical data obtained on the basis of corpus analysis demonstrate the active participation of dialectal lexis in communicative processes and once again provide scientific evidence of the close interrelationship between language and society.

According to the variationist approach proposed by W. Labov in world sociolinguistics, the selection of linguistic units is not random but is determined by social factors. The findings of the present study support this theoretical perspective: dialect-specific lexical units in English vary depending on age, social group, and communicative situation, functioning as indicators that reveal the speaker's social identity. This observation is consistent with the views of P. Trudgill and R. Wardhaugh regarding the socially stratified nature of language.

The active use of dialectal lexis in written and digital communication in English confirms the trends of digital sociolinguistics highlighted in recent research. Through the use of dialect-specific units on social networks and online platforms, speakers consciously or unconsciously express their social affiliation, cultural closeness, or informal stance. This indicates that dialectal lexis has moved beyond the boundaries of traditional oral communication and has become a means of expressing identity in new communicative spaces as well.

The results obtained for Uzbek, in turn, indicate the need to expand traditional dialectological perspectives in national linguistics from a sociolinguistic standpoint. Although Uzbek dialects have often been studied primarily in terms of territorial differentiation, the present research demonstrates their close connection with social and cultural identity. In Uzbek speech, dialect-specific lexical units express not only geographical origin but also cultural closeness, tradition, and intra-group solidarity. This once again confirms the close link between language and national mentality.

A comparative analysis of English and Uzbek shows that although dialectal lexis in both languages operates through common sociolinguistic mechanisms in expressing social identity, their functional priorities differ. In English, dialectal units are more closely associated with social stratification, age, and subcultures, whereas in Uzbek they primarily serve to reinforce regional and cultural identity. This difference can be explained by the specific social roles of language and communicative traditions in the two societies.

One of the significant aspects of this study is the interpretation of dialect-specific lexical units as active constructors of social identity. In other words, dialectal words do not merely reflect existing social realities but actively shape and reinforce identity in the process of communication. This approach corresponds to contemporary sociolinguistic theories that emphasize the dynamic nature of identity.

Furthermore, the application of corpus linguistics methods made it possible to analyze dialectal lexis on the basis of authentic language data, free from subjective observation. This can be regarded as a methodological innovation in Uzbek sociolinguistics and may prove effective for the study of other social-linguistic phenomena in future research.

Overall, the discussed results contribute to a deeper understanding of the social role of dialect-specific lexical units in English and Uzbek and provide an opportunity to draw significant scientific conclusions about the complex relationship between language, society, and identity.

**Conclusion.** This study was aimed at identifying the significant sociolinguistic role of dialect-specific lexical units in shaping and expressing social identity in English and Uzbek. The comparative analysis conducted on the basis of corpus data confirmed that dialectal lexis is not a secondary or marginal phenomenon within the language system, but rather an active linguistic mechanism closely connected with social relations in society.

The results of the research show that dialect-specific lexical units in both English and Uzbek serve as stable indicators of a speaker's regional, social, and cultural affiliation. While dialectal lexis in English is more frequently associated with age, social group, and communicative situation, in Uzbek it primarily plays a dominant role in expressing regional and cultural identity. This difference can be explained by the distinct communicative traditions and social structures characteristic of the two languages and societies.

The scientific novelty of the article lies in interpreting dialect-specific lexical units as linguistic indicators of social identity and analyzing them through a comparative corpus-based approach using English and Uzbek as examples. This approach made it possible to reveal the functions of dialectal lexis in authentic language use on the basis of objective empirical data and to enrich traditional dialectological studies from a sociolinguistic perspective.

The findings of the study are of significant theoretical importance for sociolinguistics, dialectology, and corpus linguistics, as they contribute to a deeper understanding of the relationship between language and society. At the same time, the results have practical relevance for language policy, for considering linguistic variation in educational contexts, and for explaining the social characteristics of language in intercultural communication.

In conclusion, the sociolinguistic study of dialect-specific lexical units, particularly in relation to social identity, represents a promising direction in linguistics. Future research on this topic, based on larger corpora, new communicative spaces, and additional languages, may yield further significant scientific insights.

### **References:**

1. Yodgorov U. ELEKTRON LUG 'AT, AXBOROT-QIDIRUV TEZARUSI VA ONTOLOGIYALARNI ISHLAB CHIQISH VOSITALARI //COMPUTER LINGUISTICS: PROBLEMS, SOLUTIONS, PROSPECTS. – 2022. – T. 1. – №. 1.
2. Samixon A. O 'zbek dialektologiyasi //T.: Navro 'z nashriyoti. – 2016.
3. Sadigzade Z. Dialect Diversity and Social Change: New Approaches in Sociolinguistics //Journal of Azerbaijan Language and Education Studies. – 2025. – T. 2. – №. 3. – C. 91-107.
4. Kostadinova V. et al. I English Language //The Year's Work in English Studies. – 2020. – T. 99. – №. 1. – C. 1-185.
5. Fishman J. A., García O. (ed.). Handbook of language & ethnic identity. – Oxford University Press, 2010. – T. 1.
6. Trudgill P. Dialect typology: Isolation, social network and phonological structure //Towards a Social Science of Language: Papers in honor of William Labov. Volume 1: Variation and change in language and society. – John Benjamins Publishing Company, 2011. – C. 3-22.
7. Wardhaugh R., Fuller J. M. An introduction to sociolinguistics. – John Wiley & Sons, 2021.
8. Kostadinova V. et al. I English Language //The Year's Work in English Studies. – 2020. – T. 99. – №. 1. – C. 1-185.
9. Abduazizov A. A. On Some Typological Features in the Phonological Systems of Germanic and Turkic languages //Uzbekisch-Deutsche Studien. – 2005. – T. 1. – C. 3-18.
10. Yuldashev S. EXOTICISMS IN UZBEK LANGUAGE //EduVision: Journal of Innovations in Pedagogy and Educational Advancements. – 2025. – T. 1. – №. 5. – C. 1092-1095.
11. Xolmanova Z. O 'ZBEK SHEVALARI VA TURKIY TILLAR LEKSIKASIDAGI MUSHTARAKLIK //Conferences. – 2025. – T. 1. – №. 1.
12. Mengliev D. B. et al. Building a comprehensive Uzbek lexicon: bridging dialects for text standardization //2024 IEEE 25th International Conference of Young Professionals in Electron Devices and Materials (EDM). – IEEE, 2024. – C. 2440-2444.
13. Römer U. Using general and specialized corpora in English language teaching: Past, present and future //Corpus-based approaches to English language teaching. – 2010. – C. 18-35.

14. Milroy L., Gordon M. *Sociolinguistics: Method and interpretation*. – John Wiley & Sons, 2008.
15. Kholboboeva E., Tursunnazarov E. T. Comparative and contrastive analysis of phraseological units in english and uzbek languages //Академические исследования в современной науке. – 2024. – Т. 3. – №. 34. – С. 41-46.