LEXICAL-SEMANTIC FEATURES OF SOCIOCULTURAL REALITIES IN ENGLISH AND UZBEK: AN EMPIRICAL CORPUS-BASED COMPARATIVE STUDY

Burxonova Gulmira Olimjon qizi

Abstract: This article investigates the lexical-semantic features of sociocultural realities in English and Uzbek through a comparative corpus-based and linguocultural analysis. The study focuses on culture-bound lexical units whose meanings extend beyond denotative reference and encode culturally specific social practices, values, interpersonal norms, and collective experience. The Uzbek data include ***mahalla***, ***hashar***, ***andisha***, and ***mehmondo'stlik***, while the English data include ***privacy***, ***personal space***, ***fair play***, ***queueing***, and ***stiff upper lip***. The empirical material was examined using the Uzbek National Corpus and its electronic thesaurus, together with evidence from the British National Corpus and Cambridge English Corpus. The analysis combines componential, contextual, corpus-based, and comparative methods. The findings indicate that sociocultural realities display several recurring lexical-semantic patterns: partial equivalence, semantic asymmetry, distributed equivalence, culture-specific connotation, and pragmatic extension. Corpus evidence demonstrates that Uzbek ***andisha*** is represented through a semantic network involving consideration, reflection, modesty, shame, and social restraint, whereas English ***personal space*** occurs in contexts explicitly associated with interpersonal distance and cultural variation. The analysis of ***stiff upper lip*** further demonstrates how an idiomatic expression may develop a culturally recognizable association with emotional restraint without constituting a universal characteristic of all speakers. The study argues that direct lexical equivalence is insufficient for describing sociocultural realities because their meanings are structured by culturally shared associations and recurrent discourse practices. The research contributes to comparative lexical semantics, linguoculturology, corpus linguistics, and intercultural communication.

Keywords: sociocultural realities, lexical semantics, Uzbek, English, corpus linguistics, linguoculturology, culture-bound vocabulary, cultural keywords, semantic asymmetry, intercultural communication.

1. Introduction

Language provides one of the principal means through which communities encode and transmit culturally significant knowledge. Lexical units do not merely identify objects, actions, and abstract concepts; they may also preserve information about social relationships, cultural practices, historical experience, values, and

expectations concerning appropriate behavior. This relationship between lexical meaning and cultural knowledge has been particularly influential in studies of cultural keywords and linguocultural semantics.

Wierzbicka (1997) argues that languages contain culturally significant "key words" through which central cultural concepts and values can be investigated. Her approach demonstrates that lexical semantics can provide an important methodological instrument for the study of culture because words frequently encode culturally specific assumptions that are not immediately visible to outsiders. Academic+1 More recent computational work has likewise operationalized the notion of cultural keywords by examining frequency, semantic domains, phraseological clustering, and culturally revealing collocations across languages (Lim, 2024)

The problem becomes particularly significant when languages with different historical and sociocultural backgrounds are compared. English and Uzbek provide an appropriate comparative field because culture-bound lexical units in the two languages frequently exhibit different combinations of denotative, connotative, associative, and pragmatic meanings.

For example, *mahalla* denotes a local community or neighborhood but may additionally evoke collective identity, mutual assistance, social responsibility, and community-based interaction. Uzbek *hashar* refers to a culturally established form of collective voluntary labor and therefore cannot always be represented adequately by a single English lexical equivalent. Similarly, *andisha* contains semantic components associated with consideration, reflection, modesty, social restraint, and sensitivity to others. These meanings are explicitly documented in the electronic Uzbek educational corpus and the Uzbek corpus thesaurus.

English provides comparable examples of culturally salient lexical units. *Privacy* and *personal space* encode notions of individual boundaries and interpersonal distance, while *fair play* has been analyzed as an important Anglo cultural value. The purpose of this article is to identify the lexical-semantic properties of selected sociocultural realities in English and Uzbek and to establish how cultural knowledge is represented through their semantic structures.

The study addresses four research questions:

1. What lexical-semantic components characterize selected sociocultural realities in English and Uzbek?
2. What kinds of semantic correspondence and asymmetry occur between the twolanguages?
3. How do corpus contexts reveal culturally specific associations of the selected lexical units?

4. To what extent can sociocultural realities be adequately represented through direct lexical equivalents?

2. Theoretical Background

2.1. Language, culture, and lexical meaning

The relationship between language and culture has been investigated from cognitive, anthropological, pragmatic, and semantic perspectives. From a lexical-semantic viewpoint, words are particularly valuable because they provide conventionalized access to culturally shared knowledge.

Wierzbicka (1997) proposes that culturally significant words can function as analytical windows into cultural values. Her work demonstrates that semantic analysis may reveal culturally specific configurations even when the lexical forms appear superficially equivalent across languages.

Wierzbicka's later work emphasizes that English itself contains historically transmitted cultural assumptions and that culturally specific norms of interaction may contribute to misunderstanding in intercultural communication (Wierzbicka, 2010).

This is particularly relevant to the present study because sociocultural realities are not treated merely as translation problems but as semantic structures embedded in discourse.

2.2. Sociocultural realities as culture-bound lexical units

In this study, a sociocultural reality is defined as a lexical unit or stable expression that refers to a socially or culturally recognizable phenomenon and whose interpretation involves culturally specific knowledge.

Such units may refer to:

social institutions;
interpersonal relationships;
traditions and rituals;
culturally regulated behavior;
collective values;
culturally specific objects and practices;
socially meaningful emotional or evaluative concepts.

Their semantic structure can therefore be represented as:

lexical form → denotative meaning → connotative meaning → associative meaning → pragmatic function → cultural interpretation.

This model allows the analysis to move beyond dictionary equivalence.

3. Materials and Methods

3.1. Data sources

The Uzbek component of the study was based primarily on the **Uzbek National Corpus**and its associated electronic linguistic resources. The corpus contains written

and oral materials and incorporates scientific, official, publicistic, conversational, literary, educational, and other subcorpora. It also provides concordance, parallel-corpus, thesaurus, and other linguistic resources.

The Uzbek electronic thesaurus provides particularly relevant empirical evidence for *andisha*. It identifies several semantic interpretations and provides authentic examples, including contexts in which *andisha* denotes consideration of consequences, reflection, and a sense of shame, modesty, or adherence to *or-nomus*.

For English, the study draws on the British National Corpus tradition and Cambridge English Corpus evidence. The BNC2014 represents contemporary British English and is explicitly designed as a corpus resource for research on spoken and written British English.

3.2. Selection of lexical units

The units were selected according to the following criteria:

1. the lexical unit refers to a socially or culturally recognizable phenomenon;
   its meaning contains culturally significant semantic components;
   its interpretation is dependent, at least partially, on cultural context
   corpus or authoritative lexicographic evidence is available;
   the item permits meaningful comparison with the other language.

The Uzbek sample consists of:

mahalla;
hashar;
andisha;
mehmondo'stlik.

The English sample consists of:

privacy;
personal space;
fair play;
queueing;
stiff upper lip.

The selection is qualitative rather than statistically exhaustive. Therefore, the findings should be interpreted as a semantic case study rather than as a quantitative measurement of the entire lexical systems of English and Uzbek.

3.3. Analytical procedure

Four levels of semantic analysis were applied:

Denotative analysis identifies the basic referential meaning.

Componential analysis identifies smaller semantic components within the lexical meaning.

Contextual analysis examines the meaning of the lexical unit in authentic corpus environments.

Linguocultural analysis interprets culturally specific associations, values, social practices, and pragmatic implications.

4. Results

4.1. Uzbek *andisha*: a multidimensional semantic structure

Among the Uzbek examples, *andisha* provides especially strong empirical evidence for semantic multidimensionality.

The Uzbek educational corpus defines *andisha* through several related meanings. One definition refers to thinking about consequences or acting with consideration for another person's feelings. Another identifies the word with thought or reflection. A further interpretation connects *andisha* with modesty, shame, and adherence to *or-nomus*. The resource also records phraseological usage such as *andisha qilmoq* and the proverb *Andishaning otini qo'rqoq qo'ymoq*.

The corpus thesaurus similarly presents *andisha* as a positive quality contributing to family integrity and well-being and as a protective mechanism against causing emotional harm. It also provides literary examples such as:

"...cho'chitib yubormaylik degan andisha bilan ohista so'radik."

and:

"Qandaydir andisha tutdi tilini."

These authentic examples indicate that *andisha* functions not simply as an abstract noun but as a semantic-pragmatic mechanism regulating interpersonal behavior.

Its semantic structure may therefore be represented as:

ANDISHA → reflection + consideration + restraint + modesty + social sensitivity + awareness of consequences.

This structure demonstrates **distributed semantic correspondence** with English. Depending on context, *andisha* may be approximated by *consideration*, *tact*, *discretion*, *modesty*, or *restraint*, but no single lexical item reproduces the complete semantic configuration.

Thus, *andisha* represents a case of **lexical-semantic asymmetry** rather than simple lexical non-equivalence.

4.2. *Mahalla*: from spatial designation to social institution

The semantic structure of *mahalla* extends beyond the notion of a residential area. Research on Uzbek collectivist culture identifies *mahalla*, *hashar*, weddings, *gap-gashtak*, hospitality, and community opinion as elements of Uzbek social organization and collective cultural practice . The cultural content of *mahalla* may therefore be represented as:

MAHALLA → locality + community + mutual assistance + collective identity + social interaction + communal responsibility.

The English words *neighborhood* and *community* overlap with individual semantic components but do not automatically encode the entire culturally specific institutional structure.

This represents **partial lexical equivalence combined with cultural-semantic asymmetry**.

The distinction is important because translating *mahalla* merely as *neighborhood* may preserve its spatial component while reducing its social and institutional content.

4.3. *Hashar*: collective action and communal responsibility

Hashar is another highly culture-bound Uzbek lexical unit. Research describing Uzbek collectivist culture explicitly identifies *hashar* as voluntary group work and places it alongside *mahalla*, hospitality, weddings, and other collective practices.

Its semantic configuration can be formulated as:

HASHAR → voluntary participation + collective labor + mutual assistance + communal purpose + social solidarity.

English expressions such as *community work*, *voluntary group work*, or *community volunteering* can represent components of the concept, but they do not necessarily activate the same historical and sociocultural associations.

The semantic relationship can therefore be described as:

one lexical item → several explanatory equivalents → incomplete cultural equivalence.

4.4. *Mehmondo'stlik*: hospitality as a social value

The Uzbek lexical item *mehmondo'stlik* represents a culturally important concept associated with receiving and treating guests. Research specifically devoted to the Uzbek concept identifies *mehmondo'stlik* as a linguoculturally significant semantic field and analyzes its representation in Uzbek proverbs (Eshankulova, 2024).

Comparative research on Uzbek, English, and Turkish hospitality proverbs likewise treats hospitality as a linguocultural concept rather than merely a dictionary equivalent.

The Uzbek semantic structure may be summarized as:

MEHMONDO'STLIK → guest + welcome + generosity + respect + food-sharing + social bonding.

English *hospitality* has substantial semantic overlap. However, the corpus and cultural contexts in which the two terms occur determine which associations become prominent.

Consequently, *mehmondo'stlik/hospitality* represents a case of **high lexical correspondence but potentially different cultural-semantic profiles**.

5. English Sociocultural Realities

5.1. *Personal space* and interpersonal distance

A particularly useful empirical example is the English expression *personal space*.

A corpus-based study using the British National Corpus reports 38 occurrences of *social distance* in the BNC and notes that *personal space* appears in a context explicitly discussing cultural differences in interpersonal distance. The cited BNC context describes differences in how closely people stand to one another in the UK and explicitly frames the phenomenon as culturally variable.

This evidence is important because it demonstrates that the semantic field surrounding *personal space* is not purely spatial. It also contains social and cultural components:

PERSONAL SPACE → physical distance + interpersonal boundary + comfort + social norm + cultural variation.

Thus, its meaning extends from physical space to culturally regulated interpersonal behavior.

The concept differs from Uzbek interpersonal-distance norms not necessarily because one culture "has" personal space and the other does not, but because the lexical salience and discourse representation of the concept may differ.

5.2. *Fair play*: fairness as a cultural-semantic configuration

Wierzbicka's analysis of English identifies *fair* and *fairness* as culturally significant concepts and specifically discusses *fair play* as part of the semantic and historical development of the Anglo value of fairness.

The semantic structure can be represented as:

FAIR PLAY → fairness + rules + equality + ethical behavior + respect for others.

The important observation is that *fair play* is not simply equivalent to *fairness*. The collocation activates a culturally developed frame involving rules, competition, social expectations, and morally acceptable behavior.

This is an example of **phraseological semantic specialization**.

5.3. *Stiff upper lip*: emotional restraint and cultural representation

The expression *stiff upper lip* is particularly useful for demonstrating how an idiom can acquire a culturally recognizable semantic association.

Cambridge English Corpus examples include:

> "She was to keep a stiff upper lip."

and:

> "Up to a point performers are exempted from self-restraint - from keeping a stiff upper lip."

These examples directly associate the expression with emotional self-restraintCambridge Dictionary

Corpus examples from Collins similarly show recurring patterns such as *keep a stiff upper lip*, *maintained a stiff upper lip*, and *British stiff upper lip*.

Academic research also examines *stiff upper lip* as a representation of British affect and social habitus, while emphasizing its relationship to social class and historical cultural representation (Morgan, 2023). Its semantic structure can therefore be described as:

STIFF UPPER LIP → emotional restraint + composure + endurance + controlled self-expression.

However, the expression should not be interpreted as empirical evidence that all British speakers are emotionally restrained. Rather, it represents a culturally recognizable discourse model or stereotype. This distinction is methodologically important.

6. Comparative Discussion

The empirical evidence reveals five principal types of lexical-semantic relationship.

6.1. Type I: High lexical correspondence with cultural divergence

Example:

mehmondo‘stlik ↔ hospitality

The words have considerable denotative overlap, but their associative fields may differ according to cultural practice.

6.2. Type II: Partial correspondence

Example:

mahalla ↔ community/neighborhood

The English expressions represent particular components of *mahalla*, but not necessarily its entire institutional and cultural structure.

6.3. Type III: Distributed equivalence

Example:

andisha ↔ consideration/tact/modesty/discretion/restraint

The semantic components of one Uzbek lexical item are distributed across several English lexical units.

6.4. Type IV: Lexical lacuna or strong culture-bound specificity

Example:

hashar

English can explain the concept through several expressions, but no single lexical unit necessarily reproduces its complete cultural-semantic content.

6.5. Type V: Phraseological-cultural specialization

Example:

stiff upper lip and **fair play**

The cultural meaning is not contained exclusively in the individual lexical components. It emerges from conventionalized phraseological usage and recurrent cultural associations.

7. Empirical Evidence Matrix

Reality	Language	Corpus/empirical evidence	Dominant semantic components	Type of correspondence
Andisha	Uzbek	Uzbek Educational Corpus; Uzbek Corpus Thesaurus	consideration, reflection, modesty, restraint, social sensitivity	Distributed equivalence
Mahalla	Uzbek	Uzbek Corpus; linguocultural research	community, locality, solidarity, collective identity	Partial equivalence
Hashar	Uzbek	Uzbek cultural/linguistic research	voluntary labor, solidarity, collective responsibility	Culture-specific asymmetry
Mehmon do'stlik	Uzbek	Uzbek proverb/linguocultural studies	hospitality, generosity, respect, social bonding	High but incomplete equivalence
Personal space	English	BNC-based corpus research	distance, boundaries, comfort, cultural norms	Conceptual correspondence with different salience
Fair play	English	Linguosemantic cultural analysis	fairness, rules, equality, ethical conduct	Phraseological-cultural specialization
Stiff upper lip	English	Cambridge English Corpus; Collins Corpus; academic research	restraint, composure, emotional control	Cultural/phraseological specialization

The table demonstrates that lexical-semantic equivalence exists on a continuum rather than as a binary relationship.

8. Discussion

The findings support the view that culture-bound lexical units should be analyzed as multidimensional semantic structures rather than isolated dictionary entries.

The case of *andisha* is especially revealing. Corpus and thesaurus evidence demonstrate that the word operates across several semantic domains: reflection, consideration, modesty, social restraint, and protection of interpersonal relations.

The case of *personal space* demonstrates a different process. Here, a spatial expression acquires a strong interpersonal and cultural dimension. Corpus-based research shows that discussions of personal space may explicitly invoke cultural differences in interpersonal distance.

The two cases therefore illustrate different directions of semantic development:

ANDISHA: abstract social value → interpersonal regulation

PERSONAL SPACE: physical/spatial concept → interpersonal and cultural norm

Similarly, *fair play* demonstrates how a relatively transparent phrase can become associated with a culturally meaningful moral framework. Wierzbicka's analysis places *fairness* within the historical development of Anglo cultural values.

The expression *stiff upper lip* illustrates yet another semantic mechanism: an idiom becomes associated with a culturally recognizable model of emotional behavior. Corpus evidence confirms its recurrent association with emotional self-restraint, while academic research demonstrates that the expression can be interpreted through social class and historical context. Taken together, these cases indicate that sociocultural realities should be examined at the intersection of lexical semantics, phraseology, pragmatics, and cultural conceptualization.

9. Conclusion

The corpus-based comparative analysis demonstrates that sociocultural realities in English and Uzbek possess complex lexical-semantic structures that cannot be fully described through direct lexical equivalence. The Uzbek realities *andisha*, *mahalla*, *hashar*, and *mehmondo'stlik* encode culturally significant patterns of interpersonal consideration, community organization, collective activity, hospitality, and social solidarity. The English realities *personal space*, *fair play*, and *stiff upper lip* demonstrate the semantic representation of interpersonal boundaries, fairness, rule-governed conduct, emotional restraint, and culturally recognizable behavioral ideals.

The empirical evidence supports five major conclusions.

First, sociocultural realities frequently contain semantic components that extend beyond their denotative meanings.

Second, lexical equivalence and cultural equivalence are not identical. Two words may have similar dictionary definitions while differing substantially in their associative and pragmatic structures.

Third, some culture-bound concepts demonstrate distributed equivalence: one lexical unit in one language may correspond to several lexical units in another.

Fourth, corpus data are essential for revealing how cultural meanings are activated in authentic contexts. The Uzbek Corpus and its associated thesaurus demonstrate this particularly clearly for *andisha*, while BNC-based research provides evidence for the cultural and interpersonal semantics of *personal space*.

Fifth, culturally salient expressions should not be interpreted as universal descriptions of entire populations. For example, *stiff upper lip* represents a culturally recognizable discourse model associated with Britishness, but it should not be treated as evidence that all British people exhibit emotional restraint.

The study consequently proposes that the analysis of sociocultural realities should combine **lexical semantics, corpus linguistics, phraseology, pragmatics, and linguocultural interpretation**. Such an integrated approach can contribute to more precise comparative studies of English and Uzbek and can also improve intercultural communication and the teaching of culturally embedded vocabulary.

References

1. Eshankulova, S. (2024). O‘zbek tilidagi “mehmondo‘stlik” konseptli maqollarning lingvokulturologik xususiyatlari. *Eurasian Journal of Academic Research, 4*(5), 184–187. https://doi.org/10.5281/zenodo.11316245
2. Lim, Y. (2024). A computational approach to identifying cultural keywords across languages. *Cognitive Science*. https://doi.org/10.1111/cogs.13402
3. Love, R., Dembry, C., Hardie, A., Brezina, V., & McEnery, T. (2017). The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations. *International Journal of Corpus Linguistics, 22*(3), 319–344.
4. Morgan, A. (2023). Stiff upper lips: British affect and habitus in *It's a Sin*. *Journal of Popular Television, 11*(2), 135–171. https://doi.org/10.1386/jptv_00100_1 r
5. Nagórko, A. (2021). Anny Wierzbickiej key words a kulturemy: Porównywanie języków – porównywanie kultur? *Prace Filologiczne, 76*, 387–400. https://doi.org/10.32798/pf.878
6. Palmer, G. B. (2000). Understanding cultures through their key words: English, Russian, Polish, German, and Japanese and semantics: Primes and universals. *Journal of Linguistic Anthropology, 10*(2), 279–284. https://doi.org/10.1525/jlin.2000.10.2.279
7. Usmanova, S. (2022). The specialties of Uzbek collectivism culture. *Current Research Journal of Philological Sciences, 3*(6), 14–22. https://doi.org/10.37547/philological-crjps-03-06-04
8. Wierzbicka, A. (1997). *Understanding cultures through their key words: English, Russian, Polish, German, and Japanese*. Oxford University Press. https://doi.org/10.1093/oso/9780195088359.001.000 1 OUP Academic
9. Wierzbicka, A. (2006). *English: Meaning and culture*. Oxford University https://doi.org/10.1093/acprof:oso/9780195174748.001.0001
10. Wierzbicka, A. (2010). Cross-cultural communication and miscommunication: The role of cultural keywords. *Intercultural Pragmatics, 7*(1), 1–23. https://doi.org/10.1515/IPRG.2010.001