ISSN NO: 2770-8608
February 2025

A corpus-based analysis of perception-related vocabulary in English and Uzbek

Karabaeva Barno Bobir qizi
PhD in Philology, doctorate,
Uzbekistan State World Languages University
b.karabayeva@gmail.com

Annotation. This article explores the lexical and semantic features of perception-related vocabulary in English and Uzbek through a corpus-based approach. Previous research has highlighted the significance of perception verbs and adjectives for understanding cross-linguistic conceptualization of sensory experiences (Biber, Conrad, & Reppen, 1998; Nazarova, 2018). Quantitative and qualitative analyses reveal convergences and divergences in frequency, collocational patterns, semantic extension, and grammatical behavior of perception words (such as *see*, *hear*, *feel*, *taste*, *smell* in English; *ko'rmoq* (to see), *eshitmoq* (to hear), *sezmoq* (to feel), *tatib ko'rmoq* (to taste), *hidlamoq* (to smell) in Uzbek). While direct perceptual verbs exhibit high frequency across both languages, notable differences appear in their metaphorical extensions. The results illuminate how cultural and linguistic structures mediate perception words, yielding implications for translation, language teaching, and lexicography.

Keywords: corpus linguistics, perception verbs, English, Uzbek, frequency, semantic extension

Introduction

Perception-related vocabulary forms a crucial subset of language. Across cultures and languages, humans categorize and label their sensory experiences differently, influenced by cognition, culture, and linguistic structure (Evans & Green, 2006). Verbs such as *see*, *hear*, and *feel* in English represent direct modes of sensory perception, while in Uzbek the corresponding perception terms (*ko'rmoq*, *eshitmoq*, *sezmoq*) embody overlapping yet distinct semantic nuances. The complexity and nuances embedded in these verbs demand rigorous linguistic investigation that moves beyond mere bilingual dictionaries.

In the domain of cognitive linguistics, perception verbs provide insight into how speakers of different languages conceptualize reality (Wierzbicka, 1985). These verbs are not solely denotative; they also acquire various connotations and metaphorical extensions. For instance, English uses *see* to express not just visual perception but also understanding (*I see your point*), whereas Uzbek cannot employ *ko'rmoq* in similar extended contexts (*Men sening fikringni tushundim*, literally "I understood your opinion," but can sometimes be expressed using a verb meaning 'see' in figurative usage).

Despite general acknowledgement of the importance of perception-related vocabulary, very few empirical studies have investigated these terms in a systematic, corpus-based manner in Uzbek (Hasanova, 2020; Nazarova, 2018). By employing corpora, researchers can extract authentic usage patterns, frequencies, and collocational behaviors that might not be evident through introspection alone (Biber et al., 1998). This article, therefore, addresses the need for a comparative, corpus-based study focusing on English and Uzbek perception-related vocabulary. Specifically, it aims to answer the following questions:

1. What are the most frequent perception verbs in English and Uzbek, and how do their frequencies compare?
2. What collocational and semantic patterns do these verbs exhibit?
3. How do cultural and linguistic constraints shape the metaphorical extensions of perception-related vocabulary in English and Uzbek?

Methodology

This study utilizes two main corpora. The English dataset is derived from the British National Corpus (BNC), containing approximately 100 million words of spoken and written English (BNC Consortium, 2007). For Uzbek, the Uzbek National Corpus (UNC) is employed, consisting of approximately 10 million words

ISSN NO: 2770-8608
February 2025

from literary, journalistic, and academic texts (Hasanova, 2020). Although smaller in size, the UNC provides a representative snapshot of modern Uzbek usage, including dialectal variations.

Five primary perception verbs in each language were selected, corresponding to the basic senses:

- **English**: *see*, *hear*, *feel*, *taste*, *smell*
- **Uzbek**: *ko'rmoq* ("to see"), *eshitmoq* ("to hear"), *sezmoq* ("to feel"), *tatib ko'rmoq* ("to taste"), *hidlamoq* ("to smell")

Using corpus software (AntConc 3.5.9), the study extracted raw frequency counts of each verb in both corpora. Following standard corpus-linguistic practice, we normalized frequencies per million words to facilitate cross-corpus comparison (Biber et al., 1998).

Subsequently, word sketches and concordance lines were examined to identify typical collocations—co-occurring words within a five-word span. The analysis distinguished between literal and metaphorical usages, relying on contextual clues.

Data Analysis

The methodological framework combined both quantitative and qualitative analyses. The quantitative component involved counting raw and normalized frequencies, identifying distribution patterns across text types (e.g., fiction, newspapers, academic texts). The qualitative component focused on collocational patterns, idiomatic expressions, and semantic extensions of each verb. Instances of metaphorical usage were coded to gauge cross-linguistic similarities and differences.

The data were triangulated by cross-referencing other relevant studies. For English, previous analyses of BNC data on perception verbs (Biber et al., 1998) guided the categorization of semantic functions. For Uzbek, research by Nazarova (2018) and Hasanova (2020) on perception verbs informed the classification of collocations and metaphorical meanings.

Results

Table 1 presents the normalized frequency (per million words) for the perception verbs in English and Uzbek.

Language	Sight	Hearing	Touch/Feel	Taste	Smell
English	see (2,100)	hear (1,320)	feel (2,050)	taste (210)	smell (180)
Uzbek	ko'rmoq (900)	eshitmoq (600)	sezmoq (840)	tatib ko'rmoq (70)	hidlamoq (60)

English shows higher normalized frequencies for *see* and *feel*, possibly reflecting their dual literal and figurative uses. Uzbek's overall frequencies are lower, partially due to the smaller size of the UNC and possible morphological variations. Notably, *ko'rmoq* shows a robust presence, aligning with its extensive semantic range that includes both literal and figurative meanings (Nazarova, 2018).

Collocational Patterns and Semantic Extensions

English

1. ***see***: Common collocations include "*see the difference*," "*see the point*," and "*see if*," indicating both literal vision and metaphorical sense of comprehension. Examples from the BNC reveal extended usage in contexts such as "*I see what you mean*" (BNC, Text AQA 286).
2. ***hear***: Literal contexts frequently involve direct objects such as "*hear a noise*," "*hear the music*." Metaphorical usage includes "*hear you out*," or "*I hear that*" used to signal rumor or second-hand knowledge (BNC, Text CRS 112).
3. ***feel***: Demonstrates high frequency, appearing in emotional contexts ("*feel better*," "*feel sad*") and sensory contexts ("*feel the texture*"), as well as evaluation contexts ("*feel strongly about*") (BNC, Text GVR 104).
4. ***taste***: Primarily literal usage in relation to food ("*taste the soup*," "*taste the sweetness*"), with occasional metaphors ("*taste success*").
5. ***smell***: Often used literally ("*smell the roses*," "*smell the perfume*"), and less commonly in figurative contexts ("*smell trouble*").

Uzbek

1. ***ko'rmoq* ("to see")**: Beyond literal vision ("*Derazadan tashqarini ko'rmoq*" – "to look out the window"), it is employed metaphorically to express understanding or realization, often in idiomatic expressions ("*Kelajagingni ko'rmoq*" – "to envision your future") (Nazarova, 2018, p. 47).

ISSN NO: 2770-8608
February 2025

2. ***eshitmoq* (“to hear”)**: Collocates with sounds, news, and rumors; for example, “*yomon xabarni eshitmoq*” (“to hear bad news”), and can also appear in figurative contexts related to paying attention or acknowledging advice.
3. ***sezmoq* (“to feel/perceive”)**: Often used in emotional and physiological contexts (“*yurak urishini sezmoq*” – “to feel the heartbeat”), and also found in contexts of sensing intangible elements (e.g., “*xavfni sezmoq*” – “to sense danger”) (Hasanova, 2020, p. 35).
4. ***tatib ko‘rmoq* (“to taste”)**: Primarily literal, as in “*ovqatni tatib ko‘rmoq*” (“to taste the food”). Extended metaphorical use appears in contexts like “*hayot shirinligini tatib ko‘rmoq*” (“to taste the sweetness of life”).
5. ***hidlamoq* (“to smell”)**: Mostly restricted to literal usage. Collocations are typically concrete, e.g., “*gulni hidlamoq*” (“to smell the flower”). Metaphorical usage is comparatively rare in Uzbek.

Discussion

The results confirm the centrality of perception verbs to both English and Uzbek, albeit with nuanced differences in frequency, collocations, and semantic extensions. **English** demonstrates a comparatively higher normalized frequency for the verbs of sight and touch/feel, consistent with previous corpus research indicating the prominence of these verbs for general experiential framing (Biber et al., 1998). The multi-functionality of *see* and *feel* likely contributes to their high frequency because they extend into abstract domains of cognition, emotion, and inference (e.g., “*I see your argument*,” “*I feel this is unfair*”).

In contrast, **Uzbek** shows moderate frequencies for *ko‘rmoq*, *eshitmoq*, and *sezmoq*, reflecting a balanced distribution among sight, hearing, and touch/feel. Importantly, the Uzbek verb *ko‘rmoq* is versatile, encoding both immediate visual perception and abstract understanding (Nazarova, 2018). However, the data suggest that metaphorical extensions of Uzbek perception verbs, while present, are somewhat less varied or less frequently used than their English counterparts. This finding aligns with Hasanova’s (2020) observation that Uzbek tends to preserve more literal meanings in perception vocabulary compared to English, potentially linked to cultural-linguistic preferences for direct forms of expression.

Another salient point concerns the lower frequency of taste and smell verbs across both languages. Perception via taste and smell is generally less frequently discussed in everyday discourse, leading to lower corpus frequencies (Viberg, 1984). Nonetheless, both languages display occasional figurative usage (e.g., “*taste success*” in English, “*hayot shirinligini tatib ko‘rmoq*” in Uzbek), suggesting that these senses, albeit less central, do participate in metaphoric processes.

From a **cross-cultural** perspective, these differences and similarities shed light on the interplay between language structure and cognition. Metaphors such as “*I see*” for understanding in English and “*ko‘rmoq*” in Uzbek highlight shared tendencies to use sight as the dominant sense for conceptualizing knowledge acquisition. However, the frequency and scope of these metaphors can vary, reflecting broader cultural and linguistic patterns.

In addition, grammatical behavior offers insights: English frequently omits objects when using perception verbs metaphorically (“*I see*,” “*I feel*”), whereas Uzbek often retains explicit morphological and syntactic markers. For example, in Uzbek, “*Men sening fikringni tushundim*” literally states “I have understood your opinion,” while in English, a speaker might simply say “I see.” These morphological markers can influence how frequently a verb is used metaphorically.

Conclusion

A corpus-based analysis of perception-related vocabulary in English and Uzbek reveals both convergence and divergence shaped by cultural, cognitive, and linguistic factors. English perception verbs exhibit a wide array of metaphorical expansions, contributing to higher overall frequencies for certain verbs such as *see* and *feel*. Uzbek perception verbs like *ko‘rmoq* and *sezmoq* similarly display extended uses but within a slightly narrower range.

By examining authentic corpus data, this study supports the notion that while basic perception concepts are universal, their linguistic realizations are culturally mediated (Evans & Green, 2006; Nazarova, 2018). These findings carry several implications. First, from a **translation** perspective, direct equivalents may not always capture the full semantic spectrum, requiring nuanced approaches. Second, in **language teaching**, educators should highlight the various collocational and metaphorical dimensions of perception verbs,

ISSN NO: 2770-8608
February 2025

ensuring learners appreciate both literal and extended usages. Finally, for **lexicography**, corpus-driven evidence underlines the need to document subtle shifts in meaning, collocation, and usage frequency.

Future research could expand the scope by incorporating perception adjectives (e.g., *visible*, *audible* in English, and their Uzbek equivalents), investigating dialectal variations within Uzbek, and exploring additional languages of Central Asia for a broader typological perspective. Nevertheless, the current study underscores the rich interplay of cognition, culture, and linguistic form in shaping how speakers of different languages perceive—and talk about perceiving—the world around them.

References

1. Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use.* Cambridge University Press.
2. BNC Consortium. (2007). *British National Corpus (Version 3, BNC XML Edition).* Oxford University Computing Services.
3. Evans, V., & Green, M. (2006). *Cognitive Linguistics: An Introduction.* Edinburgh University Press.
4. Hasanova, D. (2020). *Lexical Variation in Uzbek: A Corpus-Based Study*. Tashkent: Uzbekistan State World Languages University Press.
5. Nazarova, G. (2018). *Cognitive Aspects of Perception Verbs in Uzbek.* Tashkent: Fan va Turmush.
6. Viberg, Å. (1984). The verbs of perception: A typological study. In B. Butterworth, B. Comrie, & Ö. Dahl (Eds.), *Explanations for Language Universals* (pp. 123–162). Mouton.
7. Wierzbicka, A. (1985). *Lexicography and Conceptual Analysis.* Karoma Publishers.