

SEMANTIC EQUIVALENCE, TERMINOLOGICAL VARIATION AND
HARMONIZATION MECHANISMS IN A SPECIALIZED PARALLEL
CORPUS OF UNITED NATIONS DOCUMENTS

Gafarova Zumrad Zoxirjonovna

Head of Department, Associate Professor (PhD)
Asian International University

ABSTRACT	KEYWORDS
The increasing role of multilingual communication in international organizations has intensified the need for systematic terminological harmonization, particularly in institutional and legal discourse. The United Nations produces a substantial volume of multilingual documentation in which terminological accuracy, conceptual correspondence, and consistency are essential for maintaining uniform interpretation across languages. This study investigates semantic equivalence, terminological variation, and harmonization mechanisms within a specialized parallel corpus of United Nations documents. The research combines corpus linguistics, terminology studies, translation studies, and contrastive-semantic analysis to establish a methodological framework for identifying and evaluating multilingual terminological correspondences. Particular attention is given to lexical, conceptual, functional, collocational, and institutional equivalence.	Semantic equivalence, terminological variation, terminological harmonization, parallel corpus, United Nations documents, institutional discourse, corpus linguistics, terminology management, multilingual terminology, translation equivalence.

Introduction

Multilingual communication constitutes one of the fundamental conditions for the effective functioning of international institutions. The United Nations operates through six official languages Arabic, Chinese, English, French, Russian, and Spanish and most major UN documents are produced or translated across these languages. Such multilingualism creates a complex linguistic environment in which the accurate transmission of institutionally recognized concepts is of particular significance.

Terminology represents one of the most sensitive components of institutional translation. Terms occurring in resolutions, conventions, declarations, reports, agreements, and policy documents often refer to concepts whose interpretation has political, legal, diplomatic, or administrative significance. Therefore, inappropriate terminological variation may create not only stylistic inconsistency but also conceptual ambiguity. [6:149]

The institutional importance of terminology is demonstrated by the existence of UNTERM, the official global terminology portal of the United Nations. It provides terminology and nomenclature relevant to the work of the Organization in its six official languages and is explicitly designed to promote consistent use of official UN terminology throughout the system.

The problem becomes even more significant when UN terminology is transferred into languages that are not official UN languages. In such cases, several translation variants may coexist without an established terminological standard. For example, one English institutional term may acquire alternative equivalents in Uzbek because of direct translation, Russian-mediated translation, national terminological traditions, lexical synonymy, or differing interpretations of the underlying concept.

These circumstances require a methodology capable of distinguishing between legitimate terminological variation and problematic inconsistency. [9:180]

Corpus linguistics provides an effective empirical framework for such research. Corpus-based approaches to translation studies make it possible to examine recurring linguistic patterns through authentic textual evidence rather than relying exclusively on intuition or isolated examples. Baker's foundational work established corpus-based methodology as an important direction in translation studies, while subsequent research has expanded its application to specialized translation and terminological investigation. [3:129]

Specialized corpora are particularly valuable in terminology research because terms acquire their precise meanings in specific disciplinary and institutional contexts. Bowker and Pearson demonstrate that specialized corpora allow researchers and translators to study domain-specific terminology, usage patterns, collocations, and recurring linguistic structures systematically. [4:210]

The aim of the research is to develop a corpus-based framework for identifying, classifying, and harmonizing terminological equivalents in multilingual United Nations institutional discourse.

The study seeks to answer the following research questions:

1. How can semantic equivalence between source and target institutional terms be identified through a specialized parallel corpus?
2. What major types of terminological variation occur in multilingual institutional discourse?
3. How can legitimate variation be differentiated from terminological inconsistency?
4. Which corpus-based criteria can be applied to select a preferred terminological equivalent?
5. How can terminological harmonization be modelled systematically within an English–Uzbek or multilingual UN parallel corpus? [4:212]

The hypothesis of the study is that terminological harmonization becomes more reliable when translation equivalents are evaluated not only according to lexical similarity but according to their conceptual, contextual, functional, collocational, and institutional correspondence.

Materials and Methods

The research is based on a specialized parallel corpus of United Nations documents, including resolutions, conventions, declarations, reports, policy papers, human rights texts, legal documents, and sustainable development materials. Source and target texts are aligned at sentence or segment level to identify terminological equivalents systematically.

Corpus structure: *UN documents* → *Source corpus* ↔ *Target corpus* → *Term extraction* → *Equivalence analysis* → *Variation analysis* → *Harmonization*.

For English–Uzbek research, English functions as the source language and Uzbek as the target language, while Russian may be used as an auxiliary language.

2.1. Corpus Selection

Corpus documents should have official institutional authorship, corresponding source and translated versions, and sufficient metadata such as title, year, UN body, document type, subject field, language, and translation status. The corpus may cover human rights, security, sustainable development, migration, environment, law, education, health, and economic development.

2.2. Research Methods

The study applies corpus, contrastive, semantic, contextual, translation-equivalence, collocational, frequency, and terminological analyses. These methods make it possible to examine terminology in authentic institutional contexts.

2.3. Term Extraction

Terms are identified through frequency and keyword analysis, n-gram extraction, collocational analysis, and manual verification. Typical UN terms include sustainable development, peacekeeping operation, international humanitarian law, and Member States. Candidate terms are verified through context and institutional terminology resources such as UNTERM.

2.4. Semantic Equivalence

Terminological equivalence is evaluated at five levels: lexical, conceptual, functional, collocational, and institutional equivalence. Conceptual correspondence is especially important because literal translation does not always preserve the institutional meaning of a term.

2.5. Equivalence Types

Terminological equivalents are divided into:

Full equivalence — source and target concepts coincide;

Partial equivalence — concepts overlap but differ in scope or usage;

Non-equivalence — the target term does not adequately represent the source concept.

2.7. Terminological Variation

Terminological variation may be orthographic, morphological, lexical, syntactic, translational, or conceptual. The existence of several variants does not necessarily mean that the translation is incorrect. [1:245]

Results

The analysis shows that terminological equivalence is not limited to literal translation. It depends mainly on conceptual meaning, institutional function, contextual use, and terminological status.

One source term may have several target-language equivalents. Such variation may result from context, synonymy, different translation traditions, translator choice, historical development, mediated translation, or conceptual misunderstanding. Therefore, variation should not automatically be treated as inconsistency.

Terminological variation is acceptable when different forms preserve the same concept and institutional function. Terminological inconsistency, however, occurs when translations distort concepts, create ambiguity, or replace established terminology without justification.

Frequency alone cannot determine the best equivalent. A preferred term should be selected according to frequency, distribution, conceptual accuracy, contextual and collocational stability, institutional acceptance, and translation consistency.

UN terminology also includes multi-word terms, collocations, phraseological units, and institutional formulas. Therefore, harmonization should cover not only individual words but also stable expressions used in institutional discourse.

After analysis, variants can be classified as preferred terms, admitted variants, or non-preferred variants. The preferred form is the one with the strongest semantic, corpus-based, contextual, and institutional justification. The proposed harmonization mechanism consists of eight stages:

Extraction → Alignment → Equivalent identification → Semantic verification → Variation classification → Corpus validation → Harmonization → Codification.

For English–Uzbek terminology, terms may enter Uzbek either directly from English or through English → Russian → Uzbek mediated translation. Parallel corpus analysis helps identify these transfer mechanisms and their influence on Uzbek institutional terminology. [8:121]

Discussion

The results indicate that the harmonization of UN terminology should be understood as a systematic process involving both linguistic description and terminological regulation.

The proposed approach corresponds with the general principle established by ISO 860:2007, according to which terminology harmonization involves not merely terms but also concepts, concept systems, and definitions. This principle is particularly important in multilingual institutional discourse because lexical forms cannot be standardized effectively unless the underlying concepts have first been compared.

The findings also confirm the methodological value of corpus-based translation research. Corpus studies make it possible to investigate recurring translation patterns empirically and to compare source and translated discourse systematically. [2:245]

In traditional terminology work, equivalence may be established by dictionary comparison or expert judgment. Although these approaches remain valuable, parallel-corpus evidence adds an important empirical dimension. It demonstrates:

- which equivalents are actually used;
- how often they occur;
- in which document types they appear;
- which collocations they form;
- whether their usage remains stable across contexts. [7:198]

Therefore, corpus methodology establishes a relationship between terminology theory and real institutional language practice.

The findings are also closely related to the practical function of UNTERM. UNTERM contains official terminology relevant to the work of the United Nations and is maintained to support consistent terminology throughout the UN system. A specialized parallel corpus can complement such terminology resources by providing broader contextual evidence for individual terms. [3: 249]

An important theoretical implication of the research is the distinction between standardization and harmonization. Standardization tends to establish a formally preferred designation. Harmonization, by contrast, seeks compatibility between:

concepts + definitions + terms + languages + contexts.

Thus, harmonization should precede or support terminological standardization.

Another significant issue concerns the relationship between linguistic diversity and terminological uniformity. Complete elimination of variation is neither always possible nor desirable. Natural languages provide several grammatical and lexical possibilities for representing similar concepts.[10:210]

The objective should therefore not be:

variation → elimination

but rather:

variation → evaluation → functional classification → harmonization.

This model preserves legitimate target-language variation while reducing conceptual inconsistency.

For Uzbek terminology, this approach offers a particularly productive direction. English–Uzbek UN terminology could be systematically analysed through a specialized parallel corpus and compared with official multilingual UN terminology. Such research may contribute to:

- normalization of Uzbek institutional terminology;
- development of an English–Uzbek UN terminology database;
- preparation of specialized bilingual dictionaries;
- improvement of institutional translation;
- terminology-aware machine translation;
- translator training;
- linguistic standardization. [8:19]

Conclusion

The study demonstrates that semantic equivalence, terminological variation, and harmonization constitute closely interconnected dimensions of multilingual institutional communication. The analysis shows that terminological equivalence cannot be determined exclusively through lexical correspondence. Reliable equivalence requires comparison at lexical, conceptual, functional, collocational, contextual, and institutional levels.

The research also establishes an important distinction between terminological variation and terminological inconsistency. Variation may represent a natural and contextually motivated linguistic phenomenon, whereas inconsistency emerges when competing forms create conceptual instability or undermine institutional uniformity.

A specialized parallel corpus provides an effective methodological environment for distinguishing these phenomena. It enables researchers to retrieve translation equivalents, compare their distribution, analyse collocational patterns, evaluate conceptual correspondence, and establish preferred terminological forms.

The main methodological result of the study is an eight-stage harmonization mechanism: *terminological extraction → parallel alignment → equivalent identification → semantic verification → variation classification → corpus validation → harmonization → terminological codification.*

The research further demonstrates that corpus frequency should not be treated as the sole criterion for determining terminological preference. Frequency must be combined with conceptual accuracy, contextual stability, collocational behaviour, translation consistency, and institutional acceptance.

References

1. Baker, M. (1993). Corpus linguistics and translation studies: Implications and applications. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and Technology: In Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins. (pp. 233–250).
2. Baker, M. (1995). Corpora in translation studies: An overview and some suggestions for future research. *Target*, 7(2), 223–243.
3. Biel, Ł. (2014). *Lost in the Eurofog: The Textual Fit of Translated Law*. Frankfurt am Main: Peter Lang.
4. Bowker, L., & Pearson, J. (2002). *Working with Specialized Language: A Practical Guide to Using Corpora*. London/New York: Routledge.
5. Cabré, M. T. (1999). *Terminology: Theory, Methods and Applications*. Amsterdam/Philadelphia: John Benjamins.
6. Faber, P. (Ed.). (2012). *A Cognitive Linguistics View of Terminology and Specialized Language*. Berlin/Boston: De Gruyter Mouton.
7. ISO. (2007). *ISO 860:2007. Terminology Work — Harmonization of Concepts and Terms*. Geneva: International Organization for Standardization.
8. Laviosa, S. (2002). *Corpus-Based Translation Studies: Theory, Findings, Applications*. Amsterdam/New York: Rodopi.
9. McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
10. McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-Based Language Studies: An Advanced Resource Book*. London/New York: Routledge.
11. Pearson, J. (1998). *Terms in Context*. Amsterdam/Philadelphia: John Benjamins.
12. Sager, J. C. (1990). *A Practical Course in Terminology Processing*. Amsterdam/Philadelphia: John Benjamins.