

Terminological Equivalence and Harmonization Strategies in a Specialized Parallel Corpus of UN Documents

Gafarova Zumrad Zoxirjonovna

Asian International University, Head of Department, Associate Professor (PhD)
Independent researcher of Bukhara State University

Abstract

Terminological equivalence is one of the central issues in multilingual institutional communication, since the accuracy of a translated term directly influences the interpretation of the concept that it represents. This issue is especially important in United Nations documents, where legal, political, diplomatic, humanitarian and administrative concepts must be rendered consistently across languages. The problem investigated in this study arises from the fact that a single institutional concept may be expressed through several target-language equivalents, while these variants do not always demonstrate identical semantic scope, functional value or institutional status. Consequently, terminological variation may gradually develop into terminological inconsistency. The present study examines the nature of terminological equivalence and proposes harmonization strategies based on a specialized parallel corpus of UN documents. The research integrates corpus linguistics, terminology studies, translation theory, semantic analysis and contrastive linguistics.

Keywords: terminological equivalence, terminology harmonization, parallel corpus, United Nations documents, institutional discourse, terminology variation, corpus linguistics, translation studies, multilingual terminology.

Introduction

The expansion of multilingual institutional communication has increased the importance of terminological accuracy in international organizations. In the discourse of the United Nations, terminology performs a function that goes considerably beyond the ordinary lexical naming of objects or phenomena. A term often represents a concept that has already acquired a specific legal, political, diplomatic or administrative interpretation within an institutional framework. For this reason, the translation of such terminology should preserve not only the general lexical meaning of the source expression but also its conceptual structure, institutional function and established contextual usage.

The principal problem arises when one source-language term is translated by several alternative expressions in the target language. Such variation may appear natural at the lexical level, since languages rarely correspond to each other through absolute one-to-one equivalence. However, institutional terminology requires a higher degree of semantic stability than general vocabulary. If the variants used for the same source

term differ in conceptual scope, stylistic status or institutional function, translation variation may produce terminological inconsistency. [6:167]

Therefore, the problem can be formulated concisely as follows: *one institutional concept may acquire several translation equivalents, but these equivalents may not preserve the same conceptual and institutional meaning.*

This problem becomes particularly important in specialized parallel corpora. A parallel corpus makes visible the relationship between a source term and the range of target-language forms used for its translation. In an ordinary bilingual dictionary, a term may be represented by one or two conventional equivalents. In a parallel corpus, by contrast, the same source expression may reveal a much more complex pattern of variation across different documents, translators, institutional departments and historical periods. [1:24-25]

Such a situation raises several methodological questions. If three or four target-language equivalents are identified for the same source term, which one should be regarded as the preferred form? Can the most frequent equivalent automatically be treated as the most accurate? Should lexical synonymy always be eliminated? When does acceptable variation become terminological inconsistency? These questions demonstrate that terminological harmonization is not a simple process of replacing several forms with one standardized term. [3:29]

The concept of harmonization should therefore be interpreted more broadly. Terminological harmonization involves establishing correspondence not only between linguistic forms but also between the concepts represented by those forms. In institutional communication, a term is appropriate only when it remains semantically precise, contextually stable and institutionally recognizable. Thus, terminological harmonization constitutes both a linguistic and a conceptual procedure.

Corpus linguistics provides a productive methodological environment for investigating this process. A specialized parallel corpus enables the researcher to compare source and target texts systematically, retrieve recurrent equivalents, analyse their contextual distribution and identify patterns that may not be visible through isolated textual analysis. Corpus evidence therefore contributes an empirical dimension to terminology studies by showing how a term actually functions in authentic institutional discourse. [4:190]

The aim of the present research is to analyse terminological equivalence and develop harmonization strategies in a specialized parallel corpus of United Nations documents. Particular attention is paid to the relationship between equivalence, variation and standardization. The study assumes that terminological equivalence becomes reliable only when lexical similarity is supported by conceptual, functional, collocational and institutional correspondence.

Materials and Methods

The methodological basis of the research is a specialized parallel corpus consisting of United Nations institutional documents and their corresponding translations. Such documents may include resolutions, declarations, conventions, international

agreements, reports, policy documents, human rights materials, development programmes and administrative texts.

The selection of a specialized corpus rather than a general multilingual corpus is methodologically justified by the domain-specific nature of institutional terminology. A lexical unit may possess a broad general-language meaning but acquire a narrower and more specialized interpretation within UN discourse. For instance, terms related to peacekeeping, humanitarian assistance, sustainable development or international legal obligations are not merely descriptive expressions; they are integrated into established conceptual systems.

The corpus should ideally preserve detailed metadata, including document title, publication date, institutional body, document type, source language, target language and thematic field. Such metadata allows the researcher to determine whether terminological variation is associated with genre, institutional department, historical period or translation context. [2:238]

The first analytical stage consists of terminological extraction. Candidate terms may be identified through frequency lists, keyword extraction, n-gram analysis and manual examination of recurrent specialized units. Multi-word expressions are particularly important because institutional terminology frequently consists of complex nominal structures rather than isolated words. Expressions such as *sustainable development*, *international peace and security*, *human rights mechanisms*, *peacekeeping operations* and *internally displaced persons* demonstrate the structural complexity of institutional terminology. [5:79]

After term extraction, corresponding target-language forms are identified through sentence-level or segment-level alignment. At this stage, the researcher is no longer analysing the source term in isolation but examining its translation behaviour across multiple authentic contexts. This process makes it possible to establish whether the source term is translated consistently or whether it generates a set of competing equivalents.

The next stage is semantic verification. Here, the researcher evaluates whether the identified target-language forms represent the same institutional concept as the source term. Lexical similarity alone is insufficient. Two expressions may appear formally close while differing in conceptual scope. Conversely, a structurally different translation may preserve the institutional meaning more accurately. [8:35]

For this reason, terminological equivalence is analysed through several dimensions. Lexical equivalence concerns correspondence at the level of dictionary meaning. Conceptual equivalence evaluates whether the source and target terms represent the same specialized concept. Functional equivalence considers whether they perform the same communicative role in institutional discourse. Collocational equivalence examines whether the target term combines naturally with recurrent institutional phraseology. Institutional equivalence evaluates whether the term corresponds to established professional or official usage.

These dimensions are analytically interdependent. A target term may demonstrate strong lexical equivalence but weak institutional equivalence. Another variant may be lexically less direct but conceptually and functionally more accurate. Consequently, the evaluation of terminology requires an integrated approach.

The research also distinguishes three general levels of equivalence. Full equivalence occurs when the source and target terms represent substantially the same concept. Partial equivalence occurs when the concepts overlap but differ in semantic scope or contextual restriction. Non-equivalence appears when the target expression fails to preserve the essential institutional meaning of the source term. [9:29-30]

Terminological variation is then classified according to its linguistic and conceptual origin. Lexical variation arises when different lexical items are used for the same concept. Morphological variation involves derivational or grammatical differences. Syntactic variation concerns changes in the structure of multi-word terms. Orthographic variation reflects graphic differences. Translation variation occurs when translators select different equivalents for the same source unit. Conceptual variation is the most serious form, since apparently synonymous expressions may in fact represent different concepts. [10:28-30]

The final analytical stage consists of corpus validation. Competing equivalents are evaluated according to frequency, contextual distribution, semantic accuracy, collocational stability, translation consistency and institutional acceptability. This stage is necessary because high frequency does not automatically guarantee terminological correctness.

Results

The analysis demonstrates that terminological equivalence in institutional discourse is inherently multidimensional. One of the most important findings is that direct lexical correspondence cannot serve as the sole criterion for determining a successful equivalent. Institutional terms are embedded in conceptual and functional systems, and their meanings are shaped by recurrent patterns of professional use.

This means that a term should be evaluated through a sequence extending from lexical meaning to specialized concept, from specialized concept to institutional function and from institutional function to contextual stability. Such an approach reveals why some literal translations may be linguistically acceptable but terminologically inadequate.

A second important result concerns the relationship between source terms and target-language variants. In a specialized parallel corpus, a single source term may correspond to several target-language forms. These forms do not necessarily have identical status. One may represent the institutionally established equivalent, another may function as a contextually acceptable synonym, while a third may reflect inconsistent translation practice.

This finding demonstrates that terminological variation itself should not automatically be interpreted negatively. Variation is a natural characteristic of language and may arise from differences in grammatical structure, stylistic preference or contextual adaptation. The crucial question is whether the variation preserves conceptual stability.

If several variants refer to the same institutional concept without altering its semantic boundaries, they may be considered acceptable alternatives. However, if the variants introduce different interpretations, narrow or broaden the source concept, or create ambiguity, the variation becomes terminological inconsistency.

Thus, the distinction between variation and inconsistency can be explained through the notion of conceptual stability. Acceptable variation preserves the concept; problematic inconsistency modifies or obscures it. [11:45]

The research also confirms that frequency is an important but insufficient criterion for terminological selection. A frequently occurring equivalent may reflect historical translation tradition rather than contemporary conceptual precision. In some cases, a less frequent form may more accurately represent the source concept.

This observation is especially important in languages where international terminology has developed through intermediary languages. In English–Uzbek translation, for instance, some institutional terms may have entered Uzbek through Russian rather than directly from English. Such mediated forms may become frequent because of historical administrative usage, but their conceptual correspondence with contemporary English institutional terminology still requires verification.

Therefore, frequency should be interpreted together with semantic precision, contextual stability and institutional acceptability. A preferred term should demonstrate not merely recurrence but also conceptual reliability.

Another result concerns the structure of the units requiring harmonization. Institutional terminology cannot be limited to isolated words. Many concepts are expressed through stable multi-word combinations and recurrent institutional formulas. In such cases, harmonization must involve phraseological and collocational structures as well.

This is particularly visible in recurrent constructions such as *in accordance with international law*, *within the framework of*, *reaffirming its commitment to*, or *Member States are encouraged to*. Such patterns function as relatively stable institutional discourse units. Their translation should therefore be evaluated not only word by word but also as formulaic structures.

The corpus-based analysis consequently supports a broader understanding of terminological units. The object of harmonization may range from a single lexical term to a multi-word term, institutional collocation, phraseological unit or recurring discourse formula. [1:233]

The study also makes it possible to classify target-language equivalents into three functional categories. The first category is the preferred term. This is the form that demonstrates the highest level of conceptual accuracy, contextual stability and institutional acceptability. The second category is the admitted variant, which may be used in certain contexts without changing the concept. The third category is the non-preferred form, which may be semantically unstable, contextually inconsistent or institutionally unsupported.

This classification reflects the actual complexity of harmonization. The objective is not to eliminate every alternative form but to assign each variant an appropriate terminological status.

Discussion

The results suggest that terminological harmonization should be understood as a process of conceptual coordination rather than simple lexical standardization. This distinction is particularly important in multilingual institutional communication.

Standardization normally aims to establish a preferred linguistic form. Harmonization, however, requires prior analysis of the conceptual relationship between competing forms. If two variants represent genuinely different concepts, selecting one as the standard equivalent would be methodologically incorrect. Therefore, conceptual differentiation must precede formal normalization.

Corpus analysis provides the evidence needed for such differentiation. Unlike intuition-based approaches, a specialized parallel corpus allows researchers to observe repeated translation decisions across multiple contexts. This makes it possible to identify stable patterns and distinguish them from occasional or idiosyncratic choices. The corpus approach is particularly valuable because it combines quantitative and qualitative analysis. Frequency data reveal which forms are recurrent, while contextual analysis explains why they occur. Collocational patterns show how the term interacts with surrounding lexical material, and institutional comparison reveals whether a particular variant is accepted within professional discourse. [2:235]

This integrated approach also makes terminological harmonization more transparent. Rather than selecting a preferred term solely on expert intuition, the researcher can justify the choice through observable linguistic evidence.

For English–Uzbek institutional terminology, the approach has significant practical implications. Uzbek translations of UN-related concepts may reflect different linguistic traditions. Some equivalents may be direct translations from English, while others may preserve structures established through Russian. A multilingual parallel corpus can help identify these layers and determine which forms are most compatible with contemporary Uzbek terminology. [13:90]

Such research can also contribute to the development of specialized bilingual or multilingual terminology resources. A terminology database based on corpus evidence could contain not only the source term and target equivalent but also definition, context, collocations, variant forms, preferred status and source document. This would make the resource more useful for translators, editors, terminologists and researchers. The broader methodological implication is that terminology management should move from isolated lexical matching toward contextualized, corpus-supported conceptual analysis. In this sense, the parallel corpus functions not merely as a collection of texts but as an empirical environment for observing how institutional concepts are linguistically realized.

Conclusion

The analysis confirms that terminological equivalence and harmonization in United Nations documents constitute complex linguistic and conceptual processes. The central problem lies in the fact that one source-language institutional concept may generate several target-language equivalents, while these forms do not always preserve identical semantic, functional or institutional content.

The study demonstrates that terminological equivalence should therefore be evaluated through a combination of lexical, conceptual, functional, collocational, contextual and institutional criteria. Lexical similarity is only one component of equivalence and cannot independently determine terminological accuracy.

The distinction between acceptable variation and terminological inconsistency is one of the main analytical results. Variation is acceptable when the underlying concept remains stable; inconsistency arises when alternative forms change, weaken or obscure the institutional concept.

The research also shows that frequency should not be treated as the decisive factor in selecting a preferred equivalent. A harmonized term must demonstrate conceptual accuracy, contextual stability, collocational compatibility and institutional acceptability.

Accordingly, terminological harmonization can be interpreted as a sequence consisting of term extraction, parallel alignment, equivalent identification, semantic verification, variation analysis, corpus validation, preferred-term selection and terminological codification.

The proposed framework is applicable not only to the six official UN languages but also to languages such as Uzbek, where institutional terminology may still display variation caused by competing translation traditions. A specialized English–Uzbek parallel corpus of UN documents could therefore provide an empirical foundation for developing standardized and harmonized terminology resources.

In this respect, the broader value of corpus-based harmonization lies in its ability to transform scattered translation variants into systematically evaluated terminological knowledge. The process does not simply reduce variation; it clarifies conceptual relationships, establishes terminological priorities and strengthens consistency in multilingual institutional communication.

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), – P. 223–243.
3. Bowker, L., & Pearson, J. (2002). *Working with Specialized Language: A Practical Guide to Using Corpora*. London/New York: Routledge.
4. Cabré, M. T. (1999). *Terminology: Theory, Methods and Applications*. Amsterdam/Philadelphia: John Benjamins.

5. Faber, P. (Ed.). (2012). *A Cognitive Linguistics View of Terminology and Specialized Language*. Berlin/Boston: De Gruyter Mouton.
6. ISO. (2007). *ISO 860:2007. Terminology Work — Harmonization of Concepts and Terms*. Geneva: International Organization for Standardization.
7. Laviosa, S. (2002). *Corpus-Based Translation Studies: Theory, Findings, Applications*. Amsterdam/New York: Rodopi.
8. McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
9. Pearson, J. (1998). *Terms in Context*. Amsterdam/Philadelphia: John Benjamins.
10. Sager, J. C. (1990). *A Practical Course in Terminology Processing*. Amsterdam/Philadelphia: John Benjamins.
11. Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
1. Temmerman, R. (2000). *Towards New Ways of Terminology Description: The Sociocognitive Approach*. Amsterdam/Philadelphia: John Benjamins.
2. United Nations. *UNTERM: United Nations Terminology Database*. Department for General Assembly and Conference Management.