

---

**| RESEARCH ARTICLE**

## Meaning-text Theory for Cross-linguistic Comparisons: A Study on Typological Investigations

**Ly Ngoc Toan, Ph.D**

*Senior University of Labor Social Affairs (Campus II), 1018 To Ky Street, Tan Hiep Chanh Ward, District 12, Ho Chi Minh City – Vietnam*

**Corresponding Author:** Ly Ngoc Toan, **E-mail:** [toanln@ldxh.edu.vn](mailto:toanln@ldxh.edu.vn)

---

**| ABSTRACT**

This study explores the Meaning-Text Theory (MTT henceforth) and its application in cross-linguistic comparisons for typological investigations. MTT models the correspondence between meanings and their textual expressions, adopting a stratificational approach with semantic, syntactic, morphological, and phonological representation levels. A qualitative approach analyzes and compares the semantic, syntactic, and morphological structures across typologically diverse languages using the formal representations of MTT, including semantic networks, syntactic trees, lexical functions, and paraphrasing rules. MTT effectively facilitates cross-linguistic comparisons and typological investigations by capturing language-specific patterns and highlighting similarities and differences across languages at various linguistic representation levels. The findings contribute to understanding language typology and variation. MTT provides a principled approach to modeling and analyzing linguistic data across diverse languages, applicable in areas like natural language processing, machine translation, language learning, and documentation.

**| KEYWORDS**

Meaning-Text Theory, Cross-linguistic Comparisons, Typological Investigations

**| ARTICLE INFORMATION**

**ACCEPTED:** 02 September 2021

**PUBLISHED:** 30 November 2021

**DOI:** 10.32996/ijels.2021.3.11.7x

---

### 1. Introduction

This section discusses the application of Meaning-Text Theory (MTT) to cross-linguistic comparisons and typological investigations across diverse languages. MTT models the correspondence between meanings and their textual expressions using a stratificational approach with semantic, syntactic, morphological, and phonological representation levels (Mel'čuk 1988; Milićević 2006; Mel'čuk & Milićević 2020).

While previous MTT research has focused on specific phenomena or constructions within languages, this study aims to conduct a comprehensive cross-linguistic analysis comparing the lexicalization and structural organization of entire meaning spaces across typologically diverse languages like English, French, and Vietnamese. The findings reveal significant cross-linguistic patterns and universals in semantic representation, such as the use of semantic primitives/molecules (e.g., "go", "cause", "human"), hierarchical organization of semantic networks, and the representation of causality, temporality, and participant roles (agent, patient, instrument). For example, the notion of causality can be modeled in all three languages using MTT representations like "the rain caused the floods" or "le vent a cassé la branche" (the wind broke the branch) (Žolkovskij & Mel'čuk 1967; Apresjan 2000).

However, language-specific idiosyncrasies and divergences were also observed. In lexicalization patterns, while all three languages employ compounding and light verb constructions, Vietnamese productively uses verb serialization (e.g., "đi mua" - go buy), unlike English and French (Milićević 2007; Polguère 2014). Structural divergences include the topic-prominent nature of Vietnamese,

allowing flexible word order based on topicalization, versus the subject-prominence of English and French (Mednik 1992; Kahane 2003). Additionally, meaning-form mismatches like lexical gaps, polysemy divergences, structural mismatches, and pragmatic meaning gaps highlight the complex interplay between linguistic form and meaning across languages.

The study contributes to refining typological classifications by confirming and elaborating on the analytic, topic-prominent nature of Vietnamese, with widespread null arguments and productive serial verb constructions. It also sheds light on linguistic universals, such as the cross-linguistic patterns in semantic representation and the diversity of lexicalization strategies employed by different languages (Comrie 1989; Croft 2002). Future research directions could involve expanding the language sample, examining specific semantic domains (e.g., emotions, natural events) in more depth, and integrating cross-linguistic data into computational models and natural language processing applications.

Overall, by systematically comparing meaning-text correspondences across languages using the formal representations of MTT, this study enhances our understanding of language typology, universals, and the intricate ways languages encode and express meaning (Mel'čuk & Milićević 2020).

## **2. Literature review**

### **2.1. Historical background**

Three main stages in the development of Meaning-Text Theory may be identified. Zholkovsky and Melchuk's groundbreaking work on semantic and syntactic analysis based on dependency connections marked the start of the first phase in the 1960s. Their study served as the basis for the first development of the meaning-text model, which sought to establish a relationship between a language's meanings and their textual representations.

The theoretical framework of MTT and modeling technique underwent a methodical development throughout the second phase, which spanned the 1970s and 1980s. Semantics, grammar, paraphrase, and other modules of the meaning-text paradigm were constructed during this time by scholars. Prominent theoretical publications on MTT established its tenets and formalisms, including lexical functions, syntactic trees, and semantic networks. As stated in the book "An Advanced Introduction to Semantics" by Mel'čuk and Milićević (2020), MTT was first practiced by philosophers and logicians, who to this day continue to be interested in fundamental questions of semantics, such as the nature of linguistic meaning and its links with thought, meaning expressibility and meaning representation.

In the third phase, which spans the 1990s to the present, MTT is used to analyze linguistic phenomena in particular languages. Scholars have employed the theoretical and practical resources offered by MTT to create thorough meaning-text models for a variety of languages. During this period, MTT has also seen a great deal of actual implementation in fields including machine translation, natural language processing, dictionary building, and language instruction. The authors also observe, "Linguists' interest in linguistic meaning has rekindled, leading them to focus on lexical semantics—the study of word meaning. This has thus provided fresh momentum for the study of LUs in a language's semantic and syntactic domain, known as lexicology.

### **2.2. Meaning- Text Theory**

A thorough framework for creating formal models of natural languages is called Meaning-Text Theory. It is distinguished by two qualities. First and foremost, MTT is synthesis-oriented – that is, it aims at speech production (rather than speech understanding); as a result, MTT concentrates on the description of how meaning is expressed by the corresponding texts (Mel'čuk and Milićević, 2020: 8). Additionally, it is dependency-based – that is, all semantic and syntactic representations it uses are conceived in terms of dependency relations.

Using a stratificational approach to language modeling is a fundamental component of MTT. Mel'čuk and Milićević (2020: 12) posit that a stratificational model presupposes several levels of linguistic representation, and its rules are organized in a modular fashion: each representation level reflects a specific aspect of the organization of a verbal message, and the rules of the same nature are grouped into sets of manageable size, called modules, which operate between representations of adjacent levels. With representations for semantic, syntactic, morphological, and phonological elements, this multi-level design enables linguists to approach the challenges of meaning mapping to texts in an organized manner.

The layered design of MTT models offers a number of advantages. The method allows the linguist to proceed step by step and take the difficulties inherent to text synthesis (and analysis) one at a time. The general meaning or meanings conveyed by one or more possibly synonymous phrases are represented at the semantic level. Word relationships that are hierarchical, linear word order and inflection, and sound patterns are all taken into consideration at successively lower levels. MTT is ideally suited for applications such as natural language processing and documenting understudied languages because of its defined stratificational technique.

Most importantly, consistent with its synthesis-oriented orientation, MTT takes a speaker-oriented viewpoint with an emphasis on language creation. MTT. Mel'čuk and Milićević (2020: 12) add that the stratificational architecture and analytical approach of MTT are influenced by the synthesis orientation of the company. Through the process of modeling meaning-to-text mapping in natural languages, MTT offers a systematic framework for describing the meaning-to-text mapping from sounds to semantics.

### 2.3. Typological approach

Typological investigations aim to uncover cross-linguistic patterns and identify both similarities and differences among languages. This approach is concerned with the systematic comparison of linguistic structures and the classification of languages based on their shared and divergent properties. A key aspect of typological research is the identification of linguistic universals, which are patterns or tendencies that hold across a diverse range of languages, as well as language-specific idiosyncrasies. Comrie (1989 :1) says that the study of language universals allows us to separate the idiosyncrasies of individual languages from those properties which are common to all, or at least a huge range of languages.

When using the typological technique, a representative sample of languages from different language families and geographical areas are usually analyzed. A variety of linguistic aspects are studied by researchers, including syntactic structures, semantic patterns, morphological processes, and phonological inventories. Typologists compare these characteristics between languages in an effort to find typological generalizations and implicational universals, which represent the relationships between various linguistic characteristics (Žolkovskij & Mel'čuk 1967; Apresjan 2000; Kahane 2003).

The formal representations and modeling approaches offered by Meaning-Text Theory may be utilized by the typological approach when applying it to cross-linguistic comparisons. With its several levels of representation (semantic, syntactic, morphological, and phonological), the stratificational design of MTT makes it easier to compare linguistic structures systematically between languages. Typologists examine the relationships between meanings and their textual manifestations at every level in order to find patterns unique to a language as well as similarities and differences between languages (Milićević 2007; Polguère 2014).

This strategy is in line with the goals of the current research, which are as follows: We aim to uncover both universal tendencies and language-specific idiosyncrasies in the lexicalization and representation of semantic information in order to further our understanding of linguistic variation and the nature of the meaning-text interaction. Through the utilization of the formal framework of MTT and a methodical cross-linguistic examination of meaning-text correspondences, scholars may make valuable contributions to our comprehension of language typology and the multiplicity of language expression. According to William Croft (2002:1), the study of linguistic typology is one of the major ways, along with the study of language universals, of capturing generalizations across languages.

## 3. Methodology

### 3.1. Research methods

The current study uses the formal framework of Meaning-Text Theory to do typological studies and cross-linguistic comparisons utilizing a qualitative method. The following crucial elements are included in the research methods:

**Language sample selection:** To facilitate study, the English, French and Vietnamese languages representing various language families, geographical locations, and typological characteristics will be chosen. A wide variety of linguistic diversity, including both well- and understudied languages, will be represented in the selection process.

**Semantic domain identification:** To facilitate cross-linguistic comparison, particular semantic domains or meaning spaces will be identified. Concepts from a variety of subjects, such as emotions, natural events, social interactions, or cognitive processes, may be included in these categories.

**MTT representation modeling:** The formal representations of Meaning-Text Theory, such as syntactic trees, lexical functions, semantic networks, and paraphrasing rules, will be used to model the chosen semantic domains for each language in the sample.

**Cross-linguistic analysis and comparison:** A methodical analysis and comparison of the morphological, syntactic, and semantic structures within the language sample will be conducted. This entails spotting potentially universal linguistic patterns, emphasizing parallels and divergences in meaning-text correspondences, and finding language-specific patterns.

### 3.2. Data collection

English, French, and Vietnamese are the three languages that will provide data for this cross-linguistic Meaning-Text Theory study. These languages represent different typologies and families of languages. Its Indo-European origins and analytical characteristics are covered by English data. Resources on French describe its fusional nature and characteristics of the Romance language. Vietnamese offers a linguistic approach that is both analytical and Austroasiatic.

Semantic dictionaries, grammars, translations, and natural language samples from academic and subject-matter experts are some of the data sources used for these three languages. We will make use of trustworthy web resources, documents from language experts, and published works by linguists. The Indo-European and analytical traits are captured in the English dataset. French data accentuates the fusional aspects of its Romance languages. Analytic typology and the language roots of Austroasia are mentioned in relation to Vietnamese. This dataset uses modeling techniques and representations from Meaning-Text Theory to balance language variability and provide reliable cross-linguistic comparisons.

**3.3. Data analysis**

A methodical cross-linguistic investigation of the semantic, syntactic, and morphological structures in English, French, and Vietnamese - languages that reflect various language families and typological profiles - will be the focus of the data analysis. Each language in the sample will have the chosen semantic domains represented using the formal modeling tools of Meaning-Text Theory, such as syntactic trees, lexical functions, and paraphrasing rules. These MTT representations in the three languages will be carefully analyzed in order to identify any common trends as well as variances and similarities in the way meanings are mapped onto textual utterances. Language-specific quirks in the lexicalization and structural arrangement of the semantic spaces under study will receive special consideration. The study aims to clarify linguistic diversity and the complex relationship between meaning and text within the framework of various linguistic systems by utilizing this methodical approach based on the multilevel architecture of MTT.

**Table 1. Typological characteristics of MTT representations**

Languages	Language Family	Typological Characteristics	MTT Representations Used
English	Indo-European	Analytic	Semantic networks, Syntactic trees, Lexical functions, Paraphrasing rules
French	Romance (Indo-European)	Fusional	Semantic networks, Syntactic trees, Lexical functions, Paraphrasing rules
Vietnamese	Austroasiatic	Analytic	Semantic networks, Syntactic trees, Lexical functions, Paraphrasing rules

**4. Finding and discussion**

**4.1. Cross-linguistic Patterns and Universals**

4.1.1. Semantic Representation Similarities

Using the Meaning-Text Theor framework as a model, the study identifies numerous significant parallels in the semantic representation of the English, French, and Vietnamese languages. These cross-linguistic parallels imply that universal patterns exist in the meaning encoding of different languages, beyond typological and genetic distinctions. Table 2 summarizes the semantic representation similarities of the surveyed language.

**Table 2. Semantic representation similarities across languages**

Similarity Feature	English	French	Vietnamese
Use of semantic primitives/molecules	✓	✓	✓
Hierarchical organization of semantic networks	✓	✓	✓
Representation of causality	✓	✓	✓
Representation of temporality	✓	✓	✓
Representation of participant roles (agent, patient, instrument)	✓	✓	✓

The utilization of semantic primitives, sometimes known as molecules, is one notable commonality among the three languages. Semantic primitives, such as "go", "cause", and "human", are assembled to generate more complicated meanings, such as someone causing an accident or a person going to school. Another commonality among these semantic primitives is their hierarchical organization into semantic networks, which enables languages to systematically represent semantic relations like meronymy (wheel is a meronym of car), hyponymy (dog is a hyponym of animal), and other dependencies.

Moreover, there are similarities in the semantic representations of causality and temporality in English, French, and Vietnamese. Cause links between events or circumstances may be more easily modeled because to the formal representations of MTT, which include lexical functions and semantic networks. Examples of such constructs include the rain caused the floods and le vent a cassé la branche (the wind snapped the branch). Parallel to this, both languages share the representation of temporal information such as aspectual differences (đang đọc sách - is reading a book), length (elle a couru pendant une heure - she ran for an hour), and event ordering (I studied before taking the exam). One especially striking parallel is how participant roles—agents, patients, instruments, and other theme roles - are portrayed. As demonstrated by instances such as John (agent) shattered the vase (patient) with a hammer (instrument) or Tôi (agent) đã nhận được một món quà (patient) (I received a gift), the semantic representations in all three languages have ways to describe these roles. This makes it possible to compare how various languages allocate and encode these responsibilities inside their grammatical structures across languages.

Although these parallels suggest general patterns in semantic representation, it is important to note that languages might differ in the particular ways in which these semantic characteristics are realized and lexicalized. To show how ideas like causality are conveyed lexically or grammatically, for example, the research may include samples from each language, revealing both common trends and quirks unique to that language. Moreover, we may examine the subtleties and possible variations in how the semantic representations in English, French, and Vietnamese encode different semantic domains or phenomena in order to further enhance the study. Examining how different domains—like emotions, natural occurrences, or cognitive processes - are represented—like joy, frustration, or earthquake, tsunami, or think, understand, might provide important new insights on the variety of meaning-text correspondence between languages.

The study, taken as a whole, highlights significant cross-linguistic parallels in semantic representation, as encapsulated by the MTT framework. These parallels include the usage of semantic primitives, hierarchical structure, causality, timing, and participant roles. To get a better grasp of language typology and the complex links between meaning and text, a more thorough examination bolstered by concrete examples and cross-linguistic comparisons might further clarify the common patterns and language-specific quirks.

#### 4.1.2. Syntactic Structure Correspondences

Examining the syntactic structures of English, French, and Vietnamese discloses fascinating patterns of similarity and difference across the three languages. These languages have essential similarities in their basic structural arrangement, even though they belong to separate language families and have diverse typological profiles.

**Table 3. Syntactic structure correspondences across languages**

Syntactic Feature	English	French	Vietnamese
Basic Word Order (e.g. SVO, SOV)	SVO	SVO	SVO
Constituent Order	Noun Phrase precedes Verb Phrase	Noun Phrase precedes Verb Phrase	Noun Phrase precedes Verb Phrase
Similarities			
Dependency Structure	Head-initial for Verb Phrases	Head-initial for Verb Phrases	Head-initial for Verb Phrases
Parallels			
Syntactic Category	Noun, Verb, Adjective clearly distinct	Noun, Verb, Adjective clearly distinct	Flexible word categories
Representation			
Syntactic Alternation	Active/Passive, Dative alternation	Impersonal, Middle constructions	Topic/Subject prominence
Patterns			
Mapping of Semantic Roles to Syntax	Agents as subjects, Patients as objects	Agents as subjects, Patients as objects	Flexible mapping of roles

Regarding, their basic word order, which follows the Subject-Verb-Object (SVO) pattern, bears a remarkable similarity. Consider the following examples:

**English:** The student [S] reads [V] a book [O].

**French:** L'étudiant [S] lit [V] un livre [O].

**Vietnamese:** Sinh viên [S] đọc [V] một cuốn sách [O].

Despite the differences in language families and typological profiles between these languages, there is a fundamental correlation in the underlying structural organization shown in this shared SVO order. The three languages also show parallel component order, where the verb phrase comes before the noun phrase. This arrangement is demonstrated by:

**English:** The diligent student [NP] reads the book [VP].

**French:** L'étudiant studieux [NP] lit le livre [VP].

**Vietnamese:** Sinh viên chăm chỉ [NP] đọc cuốn sách [VP].

Additionally, for verb phrases, these languages have a head-initial dependence structure in which the verb functions as the head and comes before its complements and modifiers. The syntactic category representation and alternation patterns across the languages differ substantially, despite these commonalities. Vietnamese has more flexible word categories than English or French, which have firmly defined categories for nouns, verbs, and adjectives. For example, in Vietnamese, the same word can be used as an adjective, verb, or noun depending on the situation. Moreover, French has impersonal and intermediate constructions, whereas English uses dative and active/passive alternations. Conversely, Vietnamese demonstrates a topic/subject prominence pattern, in which pragmatic considerations rather than rigid syntactic constraints decide the grammatical subject.

When it comes to the mapping of semantic roles to syntax, patients and agents are often matched in English and French, as shown in:

**English:** John [agent, subject] broke the vase [patient, object].

**French:** Jean [agent, sujet] a cassé le vase [patient, objet].

Vietnamese, on the other hand, shows a more flexible mapping of semantic responsibilities, where syntactic placements are not the only determining factor.

All things considered, the three languages display peculiarities unique to each, such as differences in the way they encode syntactic categories, how they map semantic roles to syntactic structures, and how they organize words and constituents. These patterns advance our knowledge of language typology and cross-linguistic variation by highlighting the diversity and universality of linguistic structures.

4.1.3. Morphological Mapping Regularities

The mapping of semantic information onto morphological structures in the English, French, and Vietnamese languages exhibit notable parallels and contrasts, as seen in Table 4. Below is a breakdown of every feature, along with insights and examples:

Table 4. Morphological Mapping Regularities across languages

Morphological Feature	English	French	Vietnamese
Fusional/Agglutinative Typology	Fusional	Fusional	Agglutinative
Inflectional Categories (e.g., Case, Number, Gender)	Limited (Number for nouns, Tense for verbs)	Richer (Case, Gender, Number for nouns, Tense, Mood for verbs)	Limited (Classifiers for nouns, Aspect for verbs)
Productive Derivational Processes	Affixation, Compounding	Affixation, Compounding	Reduplication, Compounding
Morphological Exponence of Semantic Roles	Oblique cases for semantic roles	Prepositions for semantic roles	Word order for semantic roles
Morphological Causative/Applicative Formation	Limited (e.g., cause + Infinitive)	Productive (e.g., faire + Infinitive)	Productive (e.g., Causative serial verb constructions)

Fusional/agglutinative typology

Morphemes in English and French concurrently express several grammatical categories, a phenomenon known as fusional typology. For example, the French word *parles* (you talk) uses a single morpheme to represent person, number, and tense. In

contrast, Vietnamese follows an agglutinative typology, where each morpheme corresponds to a single grammatical category, as in *đang đọc* (present progressive aspect + read).

#### *Inflectional categories*

The inflectional system in French is more complex, designating case, gender, number, and tense and mood for verbs. The inflectional categories are more constrained in English and Vietnamese. Vietnamese employs aspect markers for verbs and classifiers for nouns, whereas English primarily inflects for number on nouns and tense on verbs.

#### *Productive derivational processes*

For word construction, all three languages use compounding and affixation. But reduplication is also used effectively in Vietnamese, as in *xanh xanh* (extremely green) and *đi đi* (to go about).

#### *Morphological exponence of semantic roles*

English uses oblique cases (like "to me") to indicate semantic roles such as beneficiaries or recipients. Prepositions such as *à moi* are used for the same purpose in French. Vietnamese, on the other hand, predominantly uses word order to encode semantic functions.

#### *Morphological causative/applicative formation*

English has limited morphological means for deriving causatives, often using periphrastic constructions like *cause + infinitive* (e.g., caused to fall). French and Vietnamese have more productive morphological processes for causatives. French uses the verb *faire + infinitive* (e.g., *faire tomber* - to make fall), while Vietnamese employs causative serial verb constructions (e.g., *làm rơi* - cause + fall).

These parallels and divergences highlight morphological mapping patterns that are peculiar to certain languages as well as general trends. For example, affixation and compounding appear to be cross-linguistic regularities in word creation. Nonetheless, these languages differ in terms of the particular inflectional categories and the morphological encoding of semantic functions, which reflects their typological divergences. Meanwhile, Vietnamese's agglutinative structure makes grammatical categories distinct and places greater emphasis on word order in relation to semantic functions. The fusional typology of French and English, on the other hand, permits more condensed morphological representations, albeit at the expense of possible opacity or ambiguity in some morphemes.

In short, these results show how meaning and form interact intricately in language and how morphological elements are used by various linguistic systems to map semantic information. The Meaning-Text Theory framework enables a methodical comparison that advances our comprehension of linguistic typology and the many ways in which languages organize and convey meaning.

## **4.2. Language-Specific Idiosyncrasies**

### **4.2.1. Lexicalization Patterns**

This section aims to give a general overview of the ways in which these three languages use strategies like verb serialization, compounding, idiomatic expressions, polysemy/semantic extensions, light verb constructions, and lexical gaps to create new lexical items, as well as their similarities and differences. The way that the languages under investigation lexicalize - that is, encode meaning into words and expressions - shows some interesting patterns and quirks. Although there are certain similarities, every language has its own distinct features. A summary of some of the most significant lexicalization trends found in Vietnamese, French, and English is provided in Table 5.

**Table 5. Lexicalization patterns across languages**

Lexicalization Pattern	English	French	Vietnamese
Compounding (e.g. bookmark, fingerprint)	Common	Common	Common
Light verb constructions (e.g. take a walk, make a decision)	Productive	Productive	Productive
Verb serialization (e.g. go buy, come see)	Limited	Limited	Very productive
Noun incorporation (e.g. truck-driver, book-writing)	Limited, lexicalized cases	Limited, lexicalized cases	Not attested
Polysemy/Semantic Extensions (e.g. hand - body part/helper)	Very extensive	Extensive	Extensive
Idiomatic expressions (e.g. kick the bucket, costs an arm and a leg)	Very extensive inventory	Extensive inventory	Extensive inventory

Lexical gaps (e.g. no single word for expressing certain concepts)	Present for some concepts	Present for some concepts	Present for some concepts
--	---------------------------	---------------------------	---------------------------

All languages benefit from light verb formulations, which combine a semantically light verb with a noun to generate a predicate. Examples of verbs used in different languages are "take a walk" in English, "faire une promenade" in French, and "đi bộ" (go walk) in Vietnamese. Verb serialization, on the other hand, contrasts sharply when two or more verb roots are combined to indicate a single occurrence. Verb serialization is used sparingly in English and French, as in "go purchase" and "venir voir" (come see), while it is widely used in Vietnamese, as in "đi mua" (go buy) and "đến xem" (come see).

In these languages, noun incorporation - the insertion of a noun into a verb form - is less common. There are very few lexicalized situations in English and French, such as "truck-driver" and "livre-objet" (book-object), however Vietnamese does not follow this pattern. On the other hands, all three languages have large amounts of polysemy, which is the use of a single word in several related meanings. For example, "hand" in English may signify both a component of the body and an assistant; "tête" in French can imply both a part of the body and an intelligent person; and "chân" in Vietnamese can represent both the leg/foot or the base of an item. Regarding idiomatic expressions, these languages are also rife with idiomatic idioms, which are fixed phrases having nonliteral meanings. Examples are "kick the bucket" in English, 'coûter les yeux de la tête' (to cost the eyes of the head, meaning very expensive) in French, and "ăn cây nào rào cây ấy" (consume which tree, lean on that tree, meaning to bear the consequences of one's actions) in Vietnamese. Finally, to varied degrees all three languages have lexical gaps—a situation in which a language does not have a single term to convey a certain idea. For example, the notion of "awkward interpersonal silence" has no single word in English (though it may be represented through a phrase), but the concept of "privacy" has no single word in Vietnamese (though it can be circumlocuted).

These lexicalization patterns show how meaning is transferred onto lexical elements, showing both universal trends and language-specific quirks. While certain methods are common across languages, such as compounding, light verb formations, and polysemy, other strategies, such as verb serialization and noun inclusion, are not. Vietnamese, for instance, commonly uses verb serialization, as in "đi mua" (go purchase), which is indicative of the language's analytical typology and propensity to convey complicated events using a series of verbs. On the other hand, this pattern is less common in English and French, which are more synthetic languages where complicated events are usually described with a single verb or light verb formation. Conversely, idiomatic phrases highlight the diverse and frequently culturally particular ways that languages lexicalize certain ideas. In contrast to the English expression "cost an arm and a leg," the French expression "coûter les yeux de la tête" (to cost the eyes of the head) represents a distinct understanding of "expense." Language-specific peculiarities in lexicalization patterns reveal information about distinct worldviews and cognitive representations of reality, in addition to reflecting typological and historical differences between different languages.

To sum up, the examination of lexicalization patterns in English, French, and Vietnamese reveals the intricate relationship that exists between form and meaning as well as the variety and universality of linguistic encoding techniques. Certain patterns, like polysemy and compounding, appear to represent universal tendencies, while other patterns, like verb serialization and colloquial phrases, highlight the diverse range of cross-linguistic variance.

4.2.2. Structural Divergences

Table 6 illustrates the notable structural differences between English, French, and Vietnamese, which are indicative of their unique typological traits and quirks. Extensive analysis of every characteristic accompanied by illustrations clarifies the various methods in which these languages arrange and represent data.

Table 6. Structural divergences across languages

Structural Feature	English	French	Vietnamese
Topic Prominence	Topic-comment structure less prominent	Topic-comment structure less prominent	Topic-prominent structure common (e.g. Topic + Comment word order)
Null Arguments	Limited null arguments	Limited null arguments	Widespread null arguments (e.g. null subjects/objects)
Constituent Order in Subordinate Clauses	Strict subject-verb order	Some flexibility in ordering	Flexible constituent order
Recursion	Recursive structures common	Recursive structures common	Recursion more limited



Grammatical Relations	Strict marking	subject/object	Subject/object marked by case	Semantic roles mapped flexibly to syntax
Complex Predicates	Less common predicates	complex	Less common complex predicates	Productive serial verb constructions
Marking of Grammatical Categories	Overt inflections	marking via	Overt marking via inflections	Tendency towards analytic structures

### *Topic prominence*

A noteworthy distinction exists in the topic-prominent structure of Vietnamese, which contrasts with the subject-prominent patterns in English and French. Vietnamese permits a flexible construction of the topic-comment pair, in which the discourse context determines whether to highlight or topicalize components other than the subject. For example:

**Vietnamese:** Cuốn sách này, tôi đã đọc rồi.

*This book, I have read already.*

In this sentence, the focus shifts from the subject "tôi" (I) to the topicalized element "cuốn sách này" (this book). Vietnamese may emphasize or foreground various parts because to its flexibility, which is in line with the language's pragmatic orientation and information flow tactics.

### *Null arguments*

Another significant difference is that null arguments, or omitted subjects or objects, are more common in Vietnamese than in English and French, where they are used far less frequently. For example:

**Vietnamese:** (Tôi) Đã đọc cuốn sách rồi.

*(I) have read the book already.*

Since the intended subject is obvious from the context, the subject "tôi" (I) need not be included. Vietnamese is characterized by a widespread null argument pattern that is part of its economy of expression. It reflects the language's analytical typology and absence of required subject-verb agreement.

### *Constituent order and recursion*

Vietnamese has a rather variable component order overall, especially in subordinate clauses, whereas French permits some flexibility and English adheres strictly to the subject-verb order in subordinate sentences. Furthermore, Vietnamese is less likely to have recursive structures than English or French. This is probably because Vietnamese is an isolating language that prefers simpler, more linear patterns. Recursive structures are linguistic units that are embedded within another of the same kind.

### *Grammatical relations and complex predicates*

French uses case marking, whereas English uses word order to precisely mark grammatical connections like subject and object. Vietnamese, on the other hand, frequently depends on pragmatic and contextual clues and transfers semantic functions more freely to syntactic structures. Furthermore, Vietnamese effectively use serial verb formulations to convey complicated occurrences and circumstances, whereas English and French have fewer complex predicates. Examples of these scenarios are:

**Vietnamese:** Tôi đi mua sách ở hiệu sách.

*I go buy book at bookstore.*

This verb structure, which reflects Vietnamese's analytic typology and information packing tactics, uses many verb roots to represent a single occurrence.

### *Marking of grammatical categories*

Vietnamese leans toward analytical structures, frequently employing distinct words or particles, whereas English and French clearly identify grammatical categories like tense and aspect through inflections. The underlying typological distinctions between these languages are reflected in this disparity.

In short, the fundamental typological differences between English, French, and Vietnamese provide the basis for these structural differences. While English and French, being more synthetic, rely more heavily on precise word order, overt morphological marking, and tougher limits on recursion, Vietnamese, because of its analytic character, allows greater freedom in word order, null arguments, and complicated predicate formulation. The many patterns that may be seen, including the relative emphasis on topics

in Vietnamese as opposed to subjects in English and French, or the widespread usage of serial verb formations in Vietnamese, demonstrate the variety of ways that languages can structure and encode meaning. Comprehending these distinctions is essential for analyzing language across languages and advances our knowledge of universals, language typology, and the astounding variety of human language systems.

#### 4.2.3. Meaning-Form Mismatches

Table 7 highlights peculiarities and differences in how meaning is encoded in English, French, and Vietnamese by displaying several kinds of meaning-form mismatches.

**Table 7. Meaning-Form Mismatches across languages**

Mismatch Type	English	French	Vietnamese
Lexical Gaps	e.g. No single word for "compadre" relationship	e.g. No single word for "mindfulness"	e.g. No single word for "privacy"
Polysemy Divergences	e.g. "Get" has many senses not mapped in French/Vietnamese	e.g. "Prendre" has senses not found in English/Vietnamese get/lay	e.g. "Lấy" has senses not matched in English/French
Structural Mismatches	e.g. Ditransitive alternation not present in French/Vietnamese	e.g. Impersonal constructions have no parallel in English/Vietnamese	e.g. Topic-prominence differs from subject-prominence in English/French
Pragmatic Meaning Gaps	e.g. No overt marking of respect/status in English	e.g. Tu/vous distinction not grammaticalized in English/Vietnamese	e.g. No overt plural marking based on animacy

#### *Lexical gaps*

In certain situations, a language is unable to convey a specific meaning using a single lexical unit. For example, the Spanish word "compadre" (co-parent) has no equivalent in English, "mindfulness" has no equivalent in French, and "privacy" has no direct counterpart in Vietnamese. To bridge these gaps and communicate these ideas, one must borrow or use circumlocution. For example, in Vietnamese, the concept of privacy is often expressed using a phrase like "không gian riêng tư" (private space) rather than a single word.

#### *Polysemy divergences*

A single term having several related senses is referred to as polysemy. Nonetheless, precise meanings of a polysemous word might differ throughout languages. For instance, the French verb "prendre" or the Vietnamese verb "lấy" are not mapped onto the English word "get," which contains meanings like "acquire" and "understand." On the other hand, "get" does not include the senses of these French and Vietnamese terms. Example: the Vietnamese "lấy" can mean "to take" or "to marry" (a spouse), a sense not shared by English "get" or French "prendre".

#### *Structural mismatches*

Various structural patterns may be used by various languages to represent the same meaning. French and Vietnamese lack the ditransitive alternation (e.g., "I gave him a book" vs. "I gave a book to him") that is present in English. The impersonal French phrases "Il faut partir" (meaning "One must leave"), on the other hand, have no exact translations in Vietnamese or English. For example, the Vietnamese language employs a topic-prominent structure, allowing flexible word order based on what is topicalized, unlike the strict subject-prominence in English and French.

#### *Pragmatic meaning gaps*

While pragmatic meanings are not grammaticalized in all languages, in some they are. For example, unlike the French tu/vous distinction, English does not have overt grammatical marking for respect/status levels. Unlike many other languages, Vietnamese does not explicitly denote plural based on animacy. For example, in Vietnamese, the same word "con" can refer to a single animal or multiple animals without plural marking, unlike in English where "dog" must be pluralized to "dogs".

In summary, these discrepancies highlight the intricate interplay between linguistic form and meaning in many languages. Some patterns point to universality, while others emphasize standards unique to a given language that have been influenced by

typological, historical, and cultural influences. Our comprehension of linguistic variety and the complex ways that languages encode meanings is deepened by studying such divergences. A methodical framework for capturing and analyzing these events across languages is offered by Meaning-Text Theory.

### 4.3. Implications for Linguistic Typology

This section looks at how the study's findings have affected our knowledge of language typology and universals. In particular, it talks about how the Meaning-Text Theory framework's cross-linguistic analysis reveals potentially universal patterns and offers new insights into the typological classifications of languages like English, French, and Vietnamese. It also makes recommendations for future research directions in this area.

#### 4.3.1. Contributions to typological classifications

The Meaning-Text Theory framework is used in this cross-linguistic study, which offers valuable insights on the typological classifications of Vietnamese, French, and English. The study provides fresh insights that deepen our comprehension of these languages while supporting earlier classifications.

The results confirm that English is one of the most analytical and subject-prominent languages in the Indo-European family. The inflexible subject-verb sequence and subject-prominence in English are confirmed by syntactic structure analysis. As an example, the phrase "The student reads a book" follows the standard SVO order, where the subject "the student" comes before the verb and noun. In addition, morphological analysis suggests that English has less morphological causals than French and Vietnamese. Rather of using specific morphological causatives, English frequently uses periphrastic formulations like "cause to fall".

Regarding French, the investigation confirms that French is a Romance language that is both fusional and subject-prominent. French is a fusional language, as evidenced by its complex inflectional system that encodes categories such as case, gender, and number on nouns and tense, mood on verbs (e.g., "parles" encodes person, number, and tense). French has some interesting morphological causatives constructed with "faire + infinitive" (e.g., "faire tomber" - to make fall), which are highlighted in this study. Furthermore, the results indicate a degree of latitude in the component order of French, departing from the rigorous subject-verb order restrictions found in English.

The study supports Vietnamese's analytical and topic-prominent typology, which is characteristic of Austroasiatic languages. Its topic-prominence is supported by the frequent use of null arguments and flexible topic-comment structures, which allow non-subjects to be topicalized (e.g., "Cuốn sách này, tôi đã đọc rồi" - This book, I have already read). Additionally, the study reveals how Vietnamese uses several verb roots to encode complicated events in highly productive serial verb constructions such as "đi mua" (go purchase), demonstrating the language's analytical character.

Through cross-linguistic comparisons utilizing the Meaning-Text Theory framework, these findings provide empirical data that advances our knowledge of language typologies. The study confirms broad categories and adds details particular to languages that help us understand their typological features better. For example, the discovery of Vietnamese's vast serial verb forms and French's productive morphological causatives enhances our knowledge of their respective morphosyntactic typologies.

All things considered, this study provides insightful information about the typological classifications of Vietnamese, French, and English. Through a methodical examination of semantic, syntactic, and morphological features in these varied languages, the research both confirms and refines existing typological classifications and reveals more subtle differences that lead to a more thorough description of language typologies.

#### 4.3.2. Insights into Language Universals

The Meaning-Text Theory framework's cross-linguistic examination offers insightful information about possible language universals, or patterns or trends that hold true across a wide variety of languages. Using a methodical modeling and comparison approach, the study uncovers some universal qualities that cut beyond typological and genetic barriers in meaning-text correspondences in English, French, and Vietnamese. The usage of semantic molecules or semantic primitives to create more complicated meanings is one notable commonality across these languages. A collection of fundamental semantic units, such as "go," "cause," "human," etc., are used in all three languages to convey complex ideas, such as "someone causing an accident," and are arranged hierarchically into semantic networks. It seems that this hierarchical structure of semantic representations is a ubiquitous feature that supports the methodical encoding of meanings. Additionally, the study found that the semantic frameworks of English, French, and Vietnamese all used similar approaches to express time, causation, and participant roles (actor, patient, and instrument). It appears that natural languages must be able to formally represent concepts of causality, temporal linkages, and thematic roles. This suggests that there may be universal principles in the way languages understand events and circumstances.

Lexical gaps, or situations in which a language lacks a single lexical unit to express a particular meaning, are another glaring universal characteristic. The research points out, for instance, that there is no term in English for the Spanish notion "compadre," there is no word in French for "mindfulness," and there is no word in Vietnamese for "privacy." Such lexical gaps between languages

imply that there are fundamental flaws in the way meaning and form are mapped, with certain concepts defying simple lexicalization across languages. Moreover, the results show that polysemy—a term with several related meanings—occurs often in different languages, even if the exact sense mappings may vary. This highlights the peculiarities unique to each language in the way that word meanings are extended through semantic linkages, while also pointing to a universal propensity for languages to do so.

In conclusion, the Meaning-Text Theory framework, with its formal representations and stratificational design, offers a methodical way to find both general patterns and linguistic quirks in the intricate mapping between meaning and text. The study sheds light on the universal cognitive and communicative skills that underpin the diversity of linguistic expression. These findings advance our knowledge of linguistic typology and the underlying basis of human language.

#### 4.3.3. Future Research Directions

A number of exciting new directions for linguistic typology and the study of meaning-text correspondences between languages have been made possible by this cross-linguistic exploration using the Meaning-Text Theory framework.

Increasing the number of genetically and typologically varied languages in the linguistic sample might be a beneficial area to pursue. A larger sample might provide more in-depth insights, even if the current study only looked at English, French, and Vietnamese, which represent analytical, fusional, and isolating kinds. For example, using polysynthetic languages such as Mohawk or Yup'ik might provide insight into the ways in which their highly synthetic nature affects the encoding of meaning. Tonal patterns in semantic representation may be clarified by investigating tonal languages such as Yoruba or Mandarin.

Furthermore, a concentrated study of particular semantic areas may yield insightful results. This study offered a broad picture across several domains; however, a more detailed investigation into a specific sector, such as emotion ideas or event structure, may reveal more detailed trends. One way to uncover interesting variations in event construal is to examine how languages encode the causative/inchoative alternation (e.g., "John broke the vase" vs. "The vase broke"). Computational implementations of MTT models across languages are a further approach. Large-scale cross-linguistic research might be facilitated and new applications in fields like machine translation could be made possible by creating computational grammars and semantic representations based on the idea. To quantify similarities and differences, one may, for example, computationally model the semantic networks and translate them to syntactic representations for other languages. Since genuine linguistic data should serve as the foundation for models, empirical studies are equally essential. By utilizing MTT's formalisms to gather and annotate cross-linguistic corpora, data-driven analysis of meaning-text correspondences may be possible. For example, one may reveal minute differences in the encoding patterns of different languages by examining a parallel corpus annotated with semantic networks, syntactic dependencies, and lexical functions.

The framework may also be expanded to include less-studied languages, which would aid in the documentation and preservation of those languages. The creation of MTT grammars for endangered languages may help to clarify their distinct meaning systems and offer a wealth of information for typological comparison. Finally, as meaning representations have ramifications outside of linguistics, it is important to establish multidisciplinary links. Research partnerships in disciplines such as philosophy, anthropology, and cognitive science may provide light on the connections between language, cognition, and culture.

Overall, this work has shown that MTT can support ethical cross-linguistic research, although there are still a lot of fascinating concerns. We anticipate that future studies exploring a wider range of languages, particular semantic phenomena, computational modeling, empirical analysis, language documentation, and interdisciplinary viewpoints will contribute to our knowledge of linguistic typology and the various ways meanings are encoded in languages around the globe.

## 5. Conclusion

The use of Meaning-Text Theory (MTT) for typological analysis and cross-linguistic comparisons has been investigated in this work. English, French, and Vietnamese are languages that represent different language families and typological profiles. We have systematically analyzed and compared the meaning-text correspondences across these languages by using the formal representations and modeling techniques of MTT, such as semantic networks, syntactic trees, lexical functions, and paraphrasing rules.

The results show that there are notable cross-linguistic parallels in the way that causality, temporality, and participant roles are represented in semantic networks, as well as in the way that semantic primitives are used and hierarchically arranged. Correspondences between syntactic structures were also noted, including dependency structures, component orders, and fundamental word orders. The study also found patterns in the morphological expression of semantic roles and causative/applicative creation, as well as morphological mapping regularities, such as the usage of affixation and compounding for word construction.

In spite of these similarities, the study also emphasized differences and quirks unique to each language. Languages differed in their lexicalization patterns, which included verb serialization, noun inclusion, and idiomatic phrases. There were clear structural differences in the marking of grammatical categories, topic prominence, null arguments, constituent order in subordinate sentences, and so forth. In addition, the research uncovered meaning-form incompatibilities, including pragmatic meaning gaps, polysemy divergences, lexical gaps, and structural mismatches.

**Limitations and Achievements:** The application of MTT to typological studies and cross-linguistic comparisons has been accomplished by this research. The research has advanced our knowledge of language typology, universals, and the complex interrelationship between meaning and linguistic form by using a methodical and ethical approach. The study's focus was restricted to three languages, though, and future investigations may include a larger variety of linguistic diversity in their language sample.

Expanding on the knowledge acquired from this investigation, several directions for further research might be pursued. Initially, broadening the scope of the research to encompass other semantic domains like emotions, natural occurrences, or cognitive processes may provide more understanding of how meaning is represented across languages. Second, adding a bigger and more varied language sample can make the typological generalizations stronger and reveal more patterns or anomalies. Third, interdisciplinary partnerships with disciplines such as computational linguistics, anthropology, and cognitive science may deepen our knowledge of the cultural and cognitive foundations of meaning-text correspondences. Lastly, examining MTT's practical applications in fields like language documentation, natural language processing, and language teaching might highlight the usefulness and significance of the theory.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

## Reference

- [1] Apresjan, J. D. (2000). *Systematic lexicography*. Oxford: Oxford University Press.
- [2] Comrie, B. (1989). *Language universals and linguistic typology: Syntax and morphology*. Chicago: University of Chicago press.
- [3] Croft, W. (2002). *Typology and universals*. Cambridge: Cambridge University Press.
- [5] Mel'čuk, I. A. (1988). *Dependency syntax: Theory and practice*. State University of New York Press.
- [6] Mel'čuk, I. A., & Miličević, J. (2020). *An advanced introduction to semantics: A meaning-text approach*. Cambridge University Press.
- [8] Miličević, J. (2006). A short guide to the Meaning-Text linguistic theory. *Journal of Koralex*, 8, 187-233.
- [9] Miličević, J. (2007). La paraphrase. *Modélisation de la paraphrase langagière [Paraphrasing: Modeling linguistic paraphrase]*. Peter Lang.
- [10] Polguère, A. (2014). From writing dictionaries to weaving lexical networks. *International Journal of Lexicography*, 27(4), 396-418.
- [11] Žolkovskij, A. K., & Mel'čuk, I. A. (1967). O semantic synthesis. *Problems of Cybernetics*, 19, 177-238.