

PREFACE

Corpus linguistics can be generally defined as a methodology that uses computer resources to gain insights into the nature, structure and usage of language by looking at (relatively) large amounts of naturally occurring language. It entails the quantitative analysis of linguistic units and processes across regional, historical settings, by exploring language patterns and variation. Corpora are defined as collections of spoken and written texts compiled using specific criteria for the study of a particular language use or level. Corpus linguistics findings are thus profitable in many research areas of applied linguistics: lexicology, grammar, pragmatics, discourse analysis, language acquisition, sociolinguistics, terminology, translation studies, computational linguistics, natural language processing, etc.

This book presents the outcome of a collaborative terminological research project focused on terminological extraction and management within a multilingual framework, specifically in the automotive technical field. The project *Terminological Extraction and Management Using a Multilingual Corpus of Specialized text (technical domain)* was conducted at the Logos Foreign Language Center, National University of Science and Technology POLITEHNICA Bucharest.

The primary aim of the study was to develop a multilingual terminological database encompassing automotive-related terms, such as components, processes, operations, mechanics, automation, and sustainable development, in four languages: English, French, German, and Romanian. The database is meant to serve students, translators, technical writers and professionals involved in technical documentation or language mediation in automotive sector.

The researchers adopted a corpus-based approach (Tognini-Bonelli, 2001), building and analysing a multilingual specialized corpus composed of technical texts (academic or institutional documents related to sustainable transport and engineering, automotive manuals, technical specifications, promotion materials) to extract and manage terminology. Semi-automatic extraction tools were used to identify candidate terms, then a manual validation was performed to ensure accuracy across different languages. During extraction, key criteria were considered: term frequency, contextual usage, morphosyntactic patterns, cross-linguistic equivalence. This combination of quantitative and qualitative analysis helped to filter out non

relevant units and focus on domain-specific, context-relevant terminology. The terms selected were then validated with experts in the field.

The dynamics of terminology generates phenomena of denominative, polysemic, or conceptual variation that cannot be described without considering factors beyond the linguistic level. Terminologies are so closely interrelated that one cannot consider the terminology used in a single field; instead, one must follow the evolution and migration of the terminological unit in question in order to account for its full complexity and designation potential.

The study also addressed such challenges related to terminological variation, both intralingual (within the same language) and interlingual (across languages) which can complicate the standardization and translation of technical terms. This variation arises due to differences between conceptual frameworks between languages, presence of synonyms or near-synonyms, inconsistent use of terms in original documentation, coexistence of standardized and non-standardized terminology. Terminological variation complicates the process of establishing stable equivalence and highlights the need for rigorous documentation, especially in high-stakes contexts like automotive engineering and safety.

The project resulted in the creation of a contextualized multilingual glossary enriched with definitions from various dictionaries, terminological data bases or scientific works, contextual examples, cross-language equivalents, and source references. Such tools serve multiple purposes across fields like translation, technical documentation, education:

- they facilitate understanding of complex concepts by offering clear, context-relevant definitions;
- they maintain coherent documentation in different languages within multinational organisations;
- they ensure specific terms are translated consistently and accurately across languages, especially in fields where precise terminology is crucial;
- if introduced in translation systems, they improve accuracy and contextual relevance in automatic translation, by reducing ambiguity in machine processing of language data.

Considering the wide range of applications and communication stakes which terminology study and management entail, the research activities were based on studies in the field of lexicology, terminology, translation but also computational linguistics. The bibliography section at the end of this book lists the main works consulted.

The authors

ENGLISH

