Semantic Variation Across Registers: A Corpus-Based Study of IT, Judiciary and Medical Discourses

Faria Batool[1], Maidah Noor Qamar[2] and Saba Nousheen[3]

https://doi.org/10.62345/jads.2025.14.1.103

Abstract

The study refers to the varying role of language in constructing meanings differently across different registers. Semantic diversity refers to manifold meanings and multiple interpretations of language within different domains. This study reveals the presence of semantic diversity within the fields of Information Technology (IT), Judiciary, and Medicine, highlighting the divergent meanings, terminologies, and linguistic features. Each of these domains has established its conventions, specialized vocabulary, and contextual nuances. The analysis explores significant diversity encountered within these registers, and how professionals in these domains understand and convey information, while also elucidating the chances of miscommunication. Through qualitative and quantitative analysis, the study aims to explore the impact of semantic diversity on communication between experts and non-experts. The findings have implications of this semantic diversity for pedagogical approaches, Natural Language Processing (NLP) and promoting effective language understanding. This paper presents the objectives, targets, and methodologies for an inclusive analysis of the linguistic variabilities and features by focusing on the registers of Medicine, IT, and Judiciary.

Keywords: Natural Language Processing (NLP), Multilingual Computing, Medicine Terminologies, Legal Lexicon, Semantic Diversity, Linguistic Flexibility, Contextual Semantics.

Introduction

Language is a means of communication and a vigorous tool for expressing ideas. It is dynamic in its application. The interpretation and connotation of language over and over again differ based on the context in which it is used. The application of language is understood not only by the structures but also by its useability within a specific situation. The language register is one of the primary factors that influences the variation in different registers. A language register shows the degree of formality or informality of language which is often formed by the way of communication, conventions, and purpose of communication. The registers of these fields can be full of specialized terminologies that are hard to understand for those unfamiliar with the register's vocabulary. Semantic diversity refers to words or phrases that have differences in meaning, interpretation, and usage of words across different fields concentrating on their different contexts or registers.

This study focuses on the adaptability of language and its capacity to transfer nuanced and context-dependent meanings as a 'social semiotic system', by showing different registers in which a word

[1]*MPhil Scholar, Department of Languages, The University of Chenab, Gujrat.*
[2]*MPhil Scholar, Department of Languages, The University of Chenab, Gujrat.*
[3]*Lecturer in English, Department of Languages, The University of Chenab, Gujrat.*
Corresponding Author Email: sabanousheen.uchenab@gmail.com

Compliance with ethical standards: There are no conflicts of interest (financial or non-financial). This study did not receive any funding.

serves in different semantic ways (Halliday, 1978). The central point of this research is to observe the semantic diversity in language registers such as IT, Medicine, and Judiciary to establish better understanding of word choice and sentence structure. Different fields have built their specialized registers and jargon that tend to have more specific, well-defined meanings and identities in that specific community. It is also observed that these registers include some common terminologies which carry different semantic values. This diversity in different registers maximizes ambiguity for non-learned speakers.

The analysis explores significant diversity encountered within the registers of IT, Medicine, and Judiciary, and how professionals in these domains understand and convey information. For example, the word "charge" signifies the prices for some services in Medicine, it refers to adding power to the battery of a device to store energy in the field of IT, and it suggests the statement by police to accuse someone in the discipline of the judiciary. This study of semantic diversity also highlights the productivity of human language and emphasizes the importance of linguistic flexibility in aiding effective communication in diverse settings. Keeping in view the linguistic flexibility it sometimes causes ambiguity in an interdisciplinary context.

Objectives of the Study

The primary objectives of this research are to enrich the understanding of semantic diversity and its implications across registers of IT, Medicine, and Judiciary.

- To analyze Lexical Variance: Examine different terminologies used within IT, Medicine, and Judiciary and identify distinctiveness and similarities between them. And to identify how the meaning of words, phrases, and technical terms differ across IT, medical, and judiciary contexts.
- Link Communication Gaps: To understand difficulties and provide solutions in communication between professionals and non-professionals in these fields.
- Discover Specialized Vocabulary and Link Communication Gap: To study how semantic diversity effects understanding and translation across language and spot challenges to provide solutions in fields of IT, medical, and judiciary.
- To explore contextual Usage: Investigate how the semantic meanings change with the change of context, audience and purpose and to examine how cultural and contextual factors influence semantic diversity within each register.
- Pedagogic and Training Connections: Develop strategies and enhance semantics analysis to teach domain-specific semantics effectively to professionals and non-professionals and study how semantic diversity impacts understanding and learning in specialized domains

Research Questions

- What are the most common terms used across different registers?
- How does semantic diversity impact understanding and learning in the fields of IT, medicine, and judiciary?

Target of the Researcher

This study targets several groups:

- Teachers and Trainers: It includes those who are interested in learning, teaching and improving linguistics and public communication regarding technical, legal and medical issues.
- Professionals of IT, Judiciary, and Medicine: It involves the experts in these fields who want to understand the nuances in their fields.

- Effective Communication: The registers of these fields can be full of specialized terminologies that are hard to understand to those unfamiliar with the register's vocabulary. The research targets to bridge effective communication between experts and non-experts.

Significance of the Study

This study is significant because of very many reasons. First of all, it enhances cross-register communication, which is useful for professionals working in interdisciplinary settings. Secondly, it will provide an insight into the use of language that varies across registers which will further help individuals to prevent miscommunications and will promote effective collaborations. Thirdly, this study contributes to the fields of linguistics and cognitive studies within professional settings. Moreover, this study can guide curriculum development, ensuring that learners grasp the linguistic nuances of their respective fields. Finally, this study highlights pedagogic applications also because the professionals and the students of the target registers can benefit from understanding register-specific language.

Literature Review

Language, as a complex and dynamic tool for communication, is not uniform in its application. The meaning and interpretation of words often vary based on the context in which they are used, and one of the primary factors influencing this variation is the language register. A language register refers to the level of formality or informality of language, often shaped by the social context, purpose of communication, audience, and the relationship between speakers. These registers ranging from formal to informal, technical to casual have a profound impact on the meaning of words and expressions, leading to what linguists refer to as semantic diversity. (Halliday, 1978). The independent words available within the registers of other fields have variable semantic diversity which can be better understood by the individuals belonging to it.

Semantic diversity refers to the differences in meaning that a word or phrase can take depending on its usage across different contexts or registers. The way a term is understood in a scientific journal, for example, can differ significantly from how it is interpreted in everyday conversation or within specific professional fields. This phenomenon highlights the adaptability of language and its ability to convey nuanced, context-dependent meanings, while also reflecting the cultural, social, and historical influences that shape communication. When it comes to meaning it depends upon the display of the linguistic structures also out of which the interpretations are drawn.

Understanding semantic diversity in language registers is crucial for comprehending how individuals adjust their speech or writing to suit various settings. For instance, in a formal register such as academic, legal, or diplomatic communication precision, clarity, and adherence to established norms often dictate word choice and sentence structure. Words in these registers tend to have more specific, well-defined meanings, minimizing ambiguity. In contrast, informal registers like casual conversation, social media discourse, or colloquial speech are more flexible and allow for greater use of slang, idioms, and figurative expressions, leading to a broader spectrum of interpretations (Chomsky, 2002).

Furthermore, the relationship between semantic diversity and language registers is also influenced by power dynamics, social status, and cultural context. For example, terms of respect for authority, such as titles or honorifics, vary widely across cultures and are dependent on the register in which they are used. The meaning of a phrase like "I'll see you later" might shift dramatically depending on whether it is said in a formal business meeting or a casual conversation with a friend.

At the core of understanding semantic diversity across different language registers is an appreciation for the role of context be it situational, cultural, or relational. Each register operates within its own set of conventions and expectations, influencing the interpretation of words, phrases, and even sentence structures. By examining the semantic nuances of language across registers, we can gain deeper insight into how meaning is constructed in communication, how speakers navigate social dynamics, and how language evolves to meet the demands of varied communicative contexts (Trudgill, 2000).

This exploration of semantic diversity not only highlights the richness of human language but also underscores the importance of linguistic flexibility in facilitating effective communication across diverse settings. Whether in a professional setting, during an academic lecture, or in everyday social interactions, language registers shape how we express ourselves and understand others. Through this lens, semantic diversity serves as both a reflection of the versatility of language and an essential component of its ongoing evolution (Lakoff, 1987).

Methodology

The study points out how language varies semantically in different social contexts. The analysis utilizes a mixed-methods approach and blends both qualitative and quantitative methods to investigate semantic ambiguities within medicine, IT, and Judiciary. As per the requirements of the research, the reason to adopt the mixed method approach is to provide not only the number of terminologies used across the registers but also their semantic values. So, the identification of the semantic qualities is done at a qualitative level. This framework provides a comprehensive lens through which the semantic complexities of domain-specific language can be examined**.** **(**McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge University Press).

- To achieve the research objectives, a mixed-method approach will be employed:

Data Collection

- Literature Review: Review existing academic literature on semantics registers of IT, Medicine, and Judiciary to identify key components of respective fields.
- Corpus analysis: Corpus of texts from each register, which may include medical journals, legal documents, IT manuals, and reports will be studied to analyze semantic diversity.

Sample Selection

Convenient sampling techniques are adopted to gather the data. The reason to adopt this technique is to utilize ease of access and availability. Furthermore, the objective of the research is to only highlight and bring forth the exhibition of language on perspectives of sematic diversity that is why this sample technique is well suited.

Tools and Techniques

The following tools have been used to support data analysis:

Corpus software: We have used Corpus analysis software such as BNC (British National Corpora) and AntConc to identify the pattern and frequency of a word.

Annotation tools of semantic: We have employed semantic annotation tools, such as FrameNet and WordNet to find the annotations and relationships between terms, concepts, and ideas.

NLP Library: we have used tool like spaCy, NLTK, and TextBlob to analysis the corpus and identify the meaning, connotation and usage of word in context.

Data Analysis

To analyze the data, this study will employ these methods:

- *Mixed-method approach:* Mixed method will be used in statistical way to analyze the frequency and usage of a word within the corpora, and apply discourse analysis to examine how language is used in various contexts across these three registers.
- *Comparative analysis:* To draw comparisons and contrasts between three registers by using criteria such as semantic similarities and differences.
- *Word selection:* We have selected 15 words that are commonly used in the three registers with different meanings, usage, and context.

Quantitative Analysis

Table 1: Semantic Similarities and dissimilarities in register of IT, Medicine and Judiciary

Semantic purpose	Registers Information Technology & Medicine	Registers Judiciary & Medicine	Registers Information Technology and Judiciary
Semantic Similarities	20%	20%	0%
Semantic Dissimilarities	66.666 %	60%	0%

Dissimilarities between IT and Judiciary Registers

The study reveals a significant similarity of approximately 86.666% between the register of IT and Judiciary. This suggests that the language which is used in these two domains is quite distinctive and contrastive.

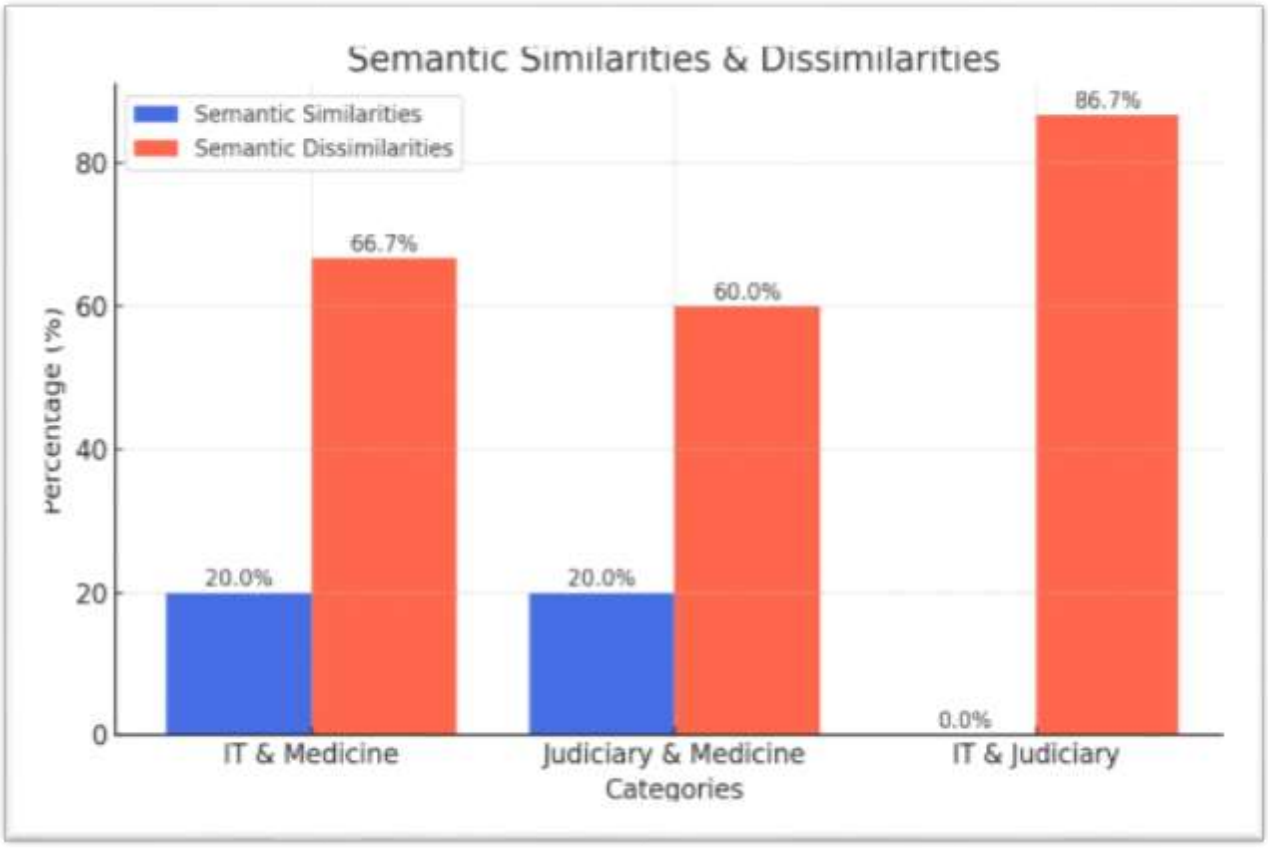
Similarities between Judiciary/Medicine and IT/Medicine Registers

The data shows that the comparison between the registers of Judiciary/Medicine and IT/Medicine is almost identical. This denotes that these two registers share some commonalities in between.

Lack of similarities between IT and Judiciary Registers

In contrast, the registers of IT and Judiciary unveil hardly any similarities which indicates that the language used in these two domains is quite specialized and distinct in its nature.

Figure 1: Semantic Similarities and dissimilarities



Semantic Diversity in Information Technology, IT, and Judiciary

Semantic Diversity in IT Register

Information Technology is highly technical and dynamic with an emphasis on functionality and precision. The IT sector register frequently advances which leads the constant evolution of technical terminologies. IT language includes programming languages (e.g., Java, Python, C+++), system terminologies, cybersecurity terms, and hardware descriptions to indicate instructions within the field. All these terminologies have their own syntax, semantic meanings and interpretations.

The terminologies in cybersecurity have their specific meanings which are found to be different in other registers (e.g., “firewall”, “cell”, “encryption”). In the register of IT “Cloud Computing” terminologies include abbreviations such as "SaaS" (Software as a Service), “PaaS” (Platform as a service) and "IaaS" (Infrastructure as a Service) have distinct meanings and interpretations that are unique to IT register. These widely used terms have nuances and connotations that are specific to the cloud computing domain. These terminologies may have distinct or industry-specific interpretation depending on the context and application which is crucial to understand for effective communication and collaboration among professionals as well as non-professionals in the field. Collaboration between IT specialists and professionals from other fields e.g., law and healthcare require precise translations of technical jargon.

Semantic Diversity in the Register of Judiciary

The register of Judiciary is distinguished by the levels of formality, prescriptiveness, and context-dependent nature. The semantic diversity of judicial language is also noteworthy as the register of judiciary is shaped by the complex legal system, the need for precision and clarity. This prescriptive natured and highly contextual based register is firmly adhered to established conventions, protocols and terminologies. Legal discourse is highly structured and often based on Latin terminology such as 'habeas corpus', 'prima facie', and 'mens rea' and specific legal expressions like "motion", "appeal", and "indictment" which have distinct and non-interchangeable meanings. This register is also often characterized by precedent-based interpretation, complexity, as it employs long, syntactically complex sentences which can simultaneously increases difficulty for laypersons.

Semantic Diversity in Medical Register

The language used in medical register is one of the most complex and nuanced forms of language characterized by its descriptive and analytical nature and its terminologies are derived from a variety of languages, including Latin, Greek, French, and German roots, creating a distinct lexicon different from everyday speech. Words such as "cardiology", "neurology", and "hematology" derived from Greek, "coronary" and "pulmonary**"** originated from Latin are hard to be understood by non-experts.

This specialized register has its own technical vocabulary, intricate syntax, jargon and discourse structures and its primary focus is on treatment and diagnosis, which make it distinct from other forms of communication and registers. This register includes disease names, anatomical terms, pharmaceutical references and procedural descriptions (e.g., 'anatomical terms' include; femur, patella, myocardium, 'disease classifications' are diabetes mellitus, carcinoma, arrhythmia and 'pharmacological terms' include beta-blockers, NSAIDs, statins). There is a significant divide between professional medical terminology and patient-friendly communication (e.g., "myocardial infarction" is typically replaced by "heart attack") for easier understanding and this step cannot be taken unless semantic diversity is recognized.

Challenges of Semantic Diversity in IT, Medicine and Judiciary

The analysis reveals that some words appear to be similar semantically in all the three fields of IT, Medicine and Judiciary, such as "agent" which plays an active role to yield special effect and the word "phishing" which means 'scamming' in all three registers. Some words serve same meaning in only two registers like IT/ Medicine, IT/Judiciary, Judiciary/Medicine, such as the word "battery" means 'a storage device for electricity' in the register of IT but it refers to 'the act of doing harm to other person' in Medicine as well as Judiciary. The term "virus" in IT refers to malware, while in medicine, it refers to biological pathogens. But some words display significant semantic multiplicity in all the three registers such as, the word "screening" refers to 'medical test' in medicine, 'the process of filtering data' in IT, and 'removal of person from the criminal justice system' in judiciary and the word "charge" signifies 'the prices for some services' in Medicine, 'to add power to device's battery to store energy' in IT, and 'statement by police to accuse someone' in judiciary (Jurafsky & Martin, 2021).

Table 2: Word frequencies in register of IT, Judiciary and Medicine

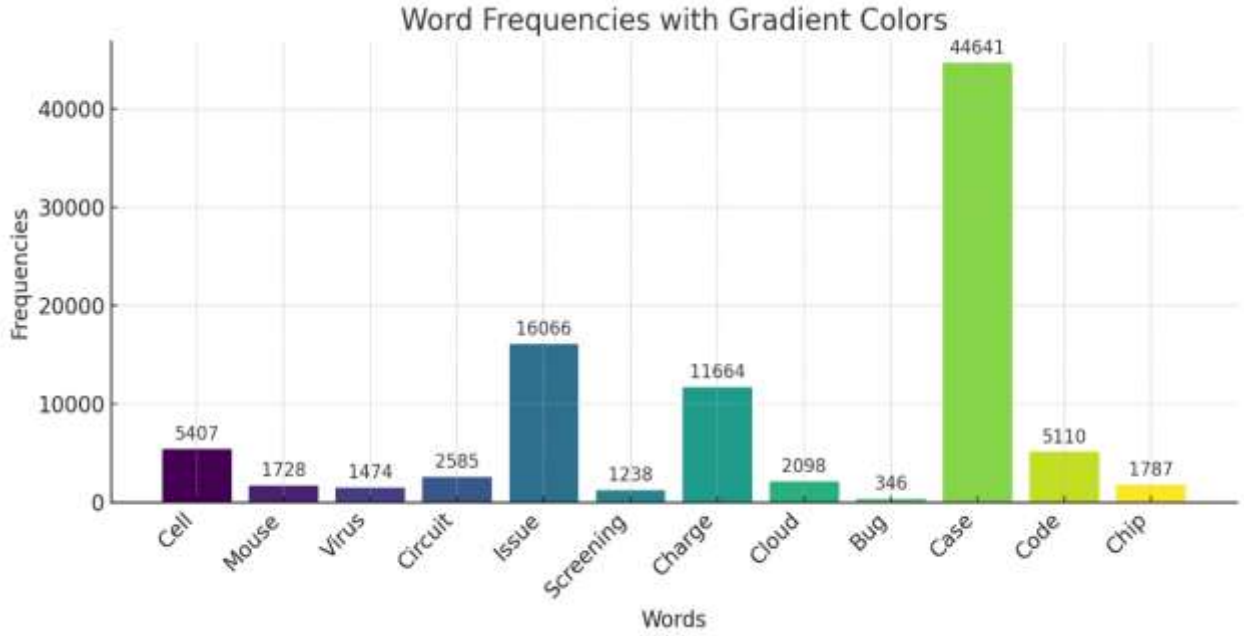
Words	Frequencies	IT	Judiciary	Medicine
Cell	5407	'a unit of data in spread sheet'	'a prison'	'smallest living organism'
Mouse	1728	'hand held device that controls a cursor on computer'		'an animals'
Virus	1474	'something that corrupts the computer files or documents'		'a microscopic infectious agent that causes disease in living organisms'
Circuit	2585	'a path that electric current flows through'	serve the meaning of a moveable court called as "circuit courts"	'a path that electric current flows through'
Issue	16066	'being problematic'	'to issue some document', like warrant	'being problematic'
Screening	1238	'the process of filtering data'	'removal of person from the criminal justice system'	'medical test'
Charge	11664	'to add power to device's battery to store energy'	'statement by police to accuse someone'	'the prices for some services'
Cloud	2098	'data storage'	'clouds on sky or weather normally'	'clouds on sky or weather normally'
Bug	346	'software error'		'an insect'
Case	44641	'a system scenario'	'a legal dispute'	'patient's condition'
Code	5110	'instruction written in computer'	'set of rules and guidelines'	'patient's respiratory arrest'
Chip	1787	'a flat piece of semi-conductor material'	'the child's protection in need'	'a disease'
Accessory	214	' a secondary object'	'aiding the criminal voluntarily'	'a secondary object'
Proxy	342	'a gateway between user and interne'.	'to make decision for another person'	'to make health care decision for another person'
Battery	1265	'storage device for electricity'	'act of doing harm to other person'	'harmful offense'

- IT field is highly dynamic, with frequent advancements leading to the constant evolution of technical terminologies. IT language includes programming languages, system terminologies, cybersecurity terms, and hardware descriptions, each with its own semantic interpretations. This register contains high usage of acronyms (e.g. AP, PC, CPU, GUI)
- In judiciary texts there is a significant use of Archaic Terminologies, terms such as habeas corpus and ex parte are frequently used in judicial settings. This register is found to be highly formalized, with terminologies rooted in Latin, historical English, and structured legal

frameworks. It emphasizes precision, precedent-based interpretation, and consistency in meaning over time, emphasizing modal verbs (e.g. shall, may, have).

- Medical language is a combination of Latin, Greek, and modern scientific terms. It includes disease names, anatomical terms, procedural descriptions, and pharmaceutical references.

Figure 2: Frequencies



Findings and Discussion

Semantic diversity across different registers reveals noteworthy variations in meaning, vocabulary and dialogue construction, such as IT, judiciary, and medical fields. This difference in meaning is influenced by the specialized nature of each domain. In IT, language is highly technical, abbreviated, and acronyms (e.g., CPU,PC, AI, AP). The register of IT often uses borrowed terminologies (e.g., "cloud," "charge," "bug"). IT discourse is dynamic and evolves with the advancement of technology. The register of judiciary, in contrast, has highly formalized and ancient language. It emphasizes on consistency, legal model, and accuracy. The terms such as "tort" (an act that causes harm to a person/property) "habeas corpus," (an independent court petition) and "jurisprudence" (the study, knowledge or science of law) have complex legal definitions and sentences to minimize ambiguity. Legal language includes Latin phrases and a constituent organization of meanings which are based on case law and legal interpretation which may lead to Ambiguity and a layperson often struggle to understand complex legal terminologies, affecting access to justice.

The register of medicine has its own uniqueness in semantic framework. It is heavily influenced by Latin and Greek roots (e.g., "neurosurgery," "hematology", "cardiomyopathy,"). These terminologies are designed specifically for clarity, specificity, and diagnostic accuracy. Medical language maintains balance in technical precision with making it easy to understand by both professional and patients (e.g., "myocardial infarction" vs. "heart attack"). Unlike IT, where meanings change rapidly, or the judiciary, where definitions remain old-fashioned, medical

semantics grow by progressions in research. All these three registers exhibit variation in domain-specific semantic as well as they have distinction in their manner of change. The study explores that each register has its specialized terminology and adaptability to convey message equitably. The study also finds that how different registers reflect their different professional demands, terminologies and communicative goals and style (Coulthard & Johnson, 2007).

Impact on Cross-Field Communication

Semantic ambiguity can hinder cross-field communication, leading to miscommunication, errors, and misunderstandings. The recognition and minimization of semantic diversity is important for effective communication across different fields, such as law, medicine, and information technology (IT), which is crucial for positive coordination, collaboration and knowledge-sharing.

Miscommunication and Ambiguity

Miscommunication arises when the same word holds different meanings across registers. For instance, the term "sanction" has different meanings in law and medicine. In law, a sanction refers to a penalty or punishment, while in medicine, it refers to a measure taken to prevent the spread of disease. Such differences in terminology can lead to ambiguity and confusion, particularly in interdisciplinary collaborations.

Interdisciplinary Collaboration and Semantic Challenges

Interdisciplinary collaboration, such as medical-legal cases or cyber law, requires careful translation of terms to prevent ambiguity. Professionals from different fields must be aware of the nuances of terminology and be able to communicate effectively to achieve common goals. For example, in medical-legal cases, healthcare professionals and lawyers must collaborate to ensure that medical records and testimony are accurately interpreted and presented in court.

Influence of Semantic Shifts

Terms evolve due to technological and legal advancements, leading to semantic shifts. For instance, the legal definition of "privacy" has changed significantly in response to digital transformations. The rise of social media, big data, and cybersecurity threats has expanded the concept of privacy, requiring legal professionals to adapt and redefine their understanding of this term.

Role of NLP in Addressing Semantic Challenges

Artificial Intelligence (AI) and Natural Language Processing (NLP) tools, such as Google Translate and IBM Watson, assist in bridging semantic gaps by interpreting specialized terminologies. These tools can help professionals from different fields to communicate more effectively by providing real-time translations and interpretations of technical terms.

Educational and Training Implications

Institutions are integrating linguistic training in law, medicine, and IT to prepare professionals for interdisciplinary challenges. This training includes instruction on specialized terminology, semantic nuances, and communication strategies to facilitate effective collaboration across fields. By providing professionals with the necessary linguistic skills and knowledge, institutions can help mitigate the risks associated with semantic challenges and promote successful cross-field communication.

Solutions

Advancements in Terminology Standardization Cross-Disciplinary Communication

There is a growing recognition of the need for standardized glossaries and terminology across various fields, including Judiciary, Medicine, and Information Technology. This recognition has significantly directed the development of innovative solutions to facilitating cross-field communication in order to improve understanding, and enhance mutual collaboration among professionals (Baker, 2018).

Standardized Glossaries for Cross-Field Communication

In order to make communication effective it is essential to develop standardized glossaries or dictionaries across different fields. These glossaries may contain common vocabulary words and set of definitions to ensure better understanding and correct use of technical terms of people from diverse background such as the development of international legal glossaries has developed coordination and synchronization across jurisdictions and administration of justice.

The implementation of standardized medical terminologies, such as "SNOMED"(Systemized Nomenclature for Medicine-Clinical Terms) "CT" (Computerized Tomography) and "ICD" (International Classification of Diseases), have transformed the field of healthcare by providing a common set of definitions, enabling accurate and consistent communication across different jurisdictions and settings.

Natural Language Processing (NLP) and Medical AI-Assisted Diagnosis Tools

The use of Natural Language Processing (NLP) tools has revolutionized the field of information technology by developing well-designed AI-powered chatbots (e.g., Alexa, Google Assistant, ChatGPT, Gemini) and language translation systems (e.g., iTranslate, Google Translator, DeepL, Memsource). These tools have played significant role in minimizing semantic gaps, providing effective and accurate collaboration, and knowledge-sharing in IT communications.

Simple Language Use in Legal Documents for Public Understanding

The use of plain and simple language in legal documents is important to enable individuals to develop clear and concise understanding, avoiding technical jargon and complex terminology. (Pinker, 1994).

Multilingual Medical Databases for Global Healthcare Communication and AI-Assisted Diagnosis Tools

The databases for Global Healthcare have provide a centralized repository of medical knowledge, enabling healthcare professionals to access and share medical information across different registers and languages. Medical field-assisted diagnosis tools can analyze complex medical data, identifying key concepts, definitions, facilitating accurate terminology usage and reduced medical errors.

Use of Visual Aids

Use visual aids, such as diagrams and flowcharts, to illustrate complex concepts to facilitate understanding and collaboration.

Table 3: Languages, challenges and solutions

Aspect	Register of IT	Register of Medicine	Register of Judiciary
Nature of Language	Technical and evolving	Scientific and structured	Formal and precedent-based
Challenges	Rapid changes, ambiguity	Complex terminologies, global variations	Legal ambiguity, translation
Solutions	NLP, AI chatbots, glossaries	Standardized medical terminologies, AI diagnosis	Legal AI, plain language laws

Conclusion

Semantic diversity in specialized registers like IT, the judiciary, and medicine plays a crucial role and reflects differences of knowledge, structures and communication styles. The domain-specific languages enhances precision within each field and presents challenges in interdisciplinary collaboration and public accessibility. Semantic ambiguity can hinder cross-field communication, leading to miscommunication, errors, and misunderstandings. This study focuses on the adaptability of language and its capacity to transfer nuanced and context-dependent meanings as a 'social semiotic system', by showing different registers in which a word serves in different semantic ways Advancements in standardized terminologies, simplicity of language and register specific tools can help mitigate semantic diversity issues. The research is significant because it has divisibility in its application not only in linguistics and language studies but also in curriculum designing and pedagogy. By examining the semantic nuances of language across registers, we can gain deeper insight into how meaning is constructed in communication, how speakers navigate social dynamics, and how language evolves to meet the demands of varied communicative contexts

Through this research we aim to uncover the complexities of language within these fields, fostering better understanding of language usage and removing barriers to communication (Pinker, 1994). Overall, the representation of the same terminologies in diverse platforms not only makes the language interlinked but also independent. For instance, it is true that different registers show the same type of words which reflect the unity in the display of the language but when it comes to the diversity it happens at the semantic level with the wide range of interpretations and referents. Concludingly, the comprehensibility of semantic diversity is essential for professional understanding across registers to avoid ambiguity.

References

- Baker, P. (2018). *Corpus Linguistics and Sociolinguistics: A Study of Variation and Change*. Routledge.
- Coulthard, M., & Johnson, A. (2007). *An Introduction to Forensic Linguistics: Language in Evidence*. Routledge
- Chomsky, N. (2002). *Syntactic Structures*. The Hague: Mouton.
- Crystal, D. (2003). *The Cambridge Encyclopedia of the English Language*. Cambridge University Press.
- Halliday, M.A.K. (1978). *Language as a Social Semiotic*. London: Edward Arnold.
- Hansen, B. (2018). Corpus linguistics and sociolinguistics: A study of variation and change in the modal systems of world Englishes, 82, Brill.

- Jurafsky, D., & Martin, J. H. (2021). *Speech and Language Processing*. Pearson
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. University of Chicago Press.
- McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge University Press.
- Pinker, S. (1994). *The Language Instinct*. Harper Perennial.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford University Press.
- Trudgill, P. (2000). *Sociolinguistics: An Introduction to Language and Society*. Penguin Books.
- Yule, G. (2016). *The Study of Language*. Cambridge University Press.