ISSN: 2957-3874 (Print)
Journal of Al-Farabi for Humanity Sciences (JFHS)
مجلة الفارابي للعلوم الإنسانية التي تصدرها كلية الفارابي الجامعة

Key Elements of Forensic Linguistics: An Overview

Mohammed Aladdin Abdlhameed[*] ,

[*] Aliraqia University – College of Law and Political Sciences/ Iraq/ Baghdad

Author Email: mohammed.abdlhameed@aliraqia.edu.iq

Abstract

Forensic linguistics is the linguistic knowledge and techniques applied to the language of the law, the legal process, and the administration of justice. We can distinguish forensic linguistics from the academic area, as a particular aspect of applied linguistics, and from the professional field, as the provision of language-related services for law practice. In recent years, this discipline has seen a multiplication of publications, conferences, and courses and the formation of specialized professional networks, associations, and working parties. This research is intended to offer an overview of the interlocked domains of interest shared by many practitioners within the forensic linguistics community.

Technological improvements, especially in machine learning, artificial intelligence, and currently associated corpora, are improving forensic linguistic techniques. This research aims to review the current directions in forensic linguistics and associate each with the current advanced methods involving these technologies.

نظرة عامة في العناصر الأساسية لعلم اللغويات الجنائية

المستخلص

علم اللغويات الجنائية هو التقنيات اللغوية والمعرفة المطبقة على لغة القانون والعملية القانونية وإدارة العدالة. يمكننا التمييز علم اللغويات الجنائية أكاديميا على اته جانب معين من علم اللغة التطبيقي. اما على الصعيد المهني فيقوم بتوفير خدمات متعلقة باللغة في ممارسة القانون. في السنوات الأخيرة، شهدنا تزايد في المنشورات والمؤتمرات والدورات وتشكيل شبكات مهنية متخصصة وجمعيات ومجموعات عمل في هذا المجال. يهدف هذا البحث إلى تقديم نظرة عامة على المجالات المتشابكة ذات الاهتمام المشترك بين العديد من الممارسين لهذا المجال داخل مجتمع اللغويات الجنائية.

إن التطور التكنولوجي، وخاصة في التعلم الآلي والذكاء الاصطناعي والمكانز اللغوية الموجودة حاليًا، تعمل على تحسين تقنيات علم اللغويات الجنائية. يهدف هذا البحث إلى مراجعة التوجهات الحالية في علم اللغويات الجنائية وربط كل منها بالطرق المتقدمة الحالية التي تنطوي على هذه التقنيات.

1. Introduction

1.1. Forensic linguistics: an overview

Forensic linguistics holds a key position at the intersection of language and law since it is designated to assist the law in matters where linguistic evidence is involved. When working with legal procedures, linguists look into language features that might assist in determining truth or studying deception. This includes studying questioned authorship, interpretation, and translation; contrasts between idiolects, dialects, and languages; the language of the law and law enforcement, including the discourse of national security; and language and the judiciary, language, and corrections.

Forensic linguistics has vast boundaries and is a field where many other areas of linguistics meet and are passed on to other specialists, such as forensic psychologists and police officers, to assist them in their studies and duties.

1.2. Research Questions

The study is directed by the subsequent inquiries:

1. Are there any organized directions of this field of study that assist researchers and professionals in promoting this field?
2. Is there enough research in this field regarding the emergence of the current technologies, and how are they employed?
3. What implications does digital communication have for forensic linguistics?

1.3. Aims of the Research

This research has different aims:

1. To evaluate existing AI, corpora, and applications and their efficacy in legal contexts.
2. To investigate digital advances in the language of the written law. Hence, examine how they are in line with current challenges.
3. To explore language use in forensic and judicial processes and identify best practices in the field, enhancing the accuracy and reliability of linguistic analyses in legal settings.

1.4. Importance of the Study

This research has some important aspects sorted as follows:

1- This research helps to describe how forensic linguistics can be applied to contemporary legal issues.

2- The research seeks to enlist the successful approaches for analysis by assessing AI tools, machine learning, and corpora.

3- The research reviews the contemporary significant sociolinguistic issues in forensic linguistics.
4- The research will provide practical insights for researchers and legal practitioners, which help them to employ linguistic analysis more effectively.

1.5. Methodology

1. An in-depth literature review on forensic linguistics and current technological developments is conducted in order to examine recent works that address almost all the topics within this field.
2. Review the effectiveness of various AI apps and technologies. This requires classifying and comparing different AI models.
3. The study concluded with several suggestions for future directions in forensic linguistics technology.

2. Literature review

2.1. Forensic Linguistics: Definition and Scope

Forensic linguistics is an interdisciplinary field. It "is a branch of applied linguistics relating to the law and legal processes" (Perkins & Grant, 2013, p. 174)[1]. It employs language analysis methods to aid in criminal justice and legal investigations. Moreover, it utilizes linguistic theories and methodology to analyze language in court documents, crime-related communications, and testimonies (Perkins & Grant, 2013)[2].

The use of forensic linguistics in the legal system is critical. Language analysis serves as vital evidence when traditional forensic evidence is limited. It examines written and spoken language and consists of subdisciplines such as authorship attribution, phonetic analysis, discourse analysis, and sociolinguistic profiling. This field is varied and constantly evolving in accordance with developments in both legal and technological domains (Perkins & Grant, 2013)[3].

2.2. Chronological Evolution

Forensic linguistics is a recent discipline. The officially recognized case can be traced to Jan Svartvik's 1968 examination of the Evans Statement. He performed research to demonstrate that law enforcement agencies had altered the records, leading to the erroneous conviction and

[1] Perkins, R., & Grant, T. (2013). Forensic Linguistics. In J. A. Siegel, P. J. Sauko, & M. M. Houck (Eds.), *Encyclopedia of Forensic Science* (2nd ed., pp. 174–177). Elsevier Ltd.

[2] Perkins, R., & Grant, T. (2013). Forensic Linguistics. In J. A. Siegel, P. J. Sauko, & M. M. Houck (Eds.), *Encyclopedia of Forensic Science* (2nd ed., pp. 174–177). Elsevier Ltd.

[3] Perkins, R., & Grant, T. (2013). Forensic Linguistics. In J. A. Siegel, P. J. Sauko, & M. M. Houck (Eds.), *Encyclopedia of Forensic Science* (2nd ed., pp. 174–177). Elsevier Ltd.

execution of an innocent individual. (Svartvik, 1968)[4]. Most researchers consider Svartvik's research as "the first typical work related to forensic linguistics." (Khoyi & Behnam, 2014, p.439)[5]. It highlighted the reliability of linguistic analysis in legal contexts, stimulating interest in forensic applications.

In the 1980s and 1990s, forensic linguistics experienced considerable expansion, especially in authorship analysis. The 1991 Unabomber case illustrates how experts recognized Theodore Kaczynski by his unique linguistic style in his writings, aiding in his identification and capture. These cases revealed that language "fingerprints" may be used in criminal investigations (Kreuz, 2023)[6].

The discipline has expanded into various specialized sub-disciplines. For example, forensic phonetics, which is "often distinguished as a separate domain" (Crystal, 2008, p. 194)[7], examines vocal patterns for identification. Sociolinguistic profiling is "whereby the linguist offers assistance in the investigation of what sort of person produced a text" (MacLeod, 2021, p.159)[8]. These subfields provide distinct insights into forensic investigations. This helped enhance suspect profiling and contextual analysis.

Forensic linguistics is increasingly vital to criminal investigations and the judicial system. The International Association of Forensic Linguists advocates for the academic interests of the discipline, publishes a peer-reviewed journal, and frequently organizes international conferences. The progression of computational technology offers forensic linguistics both prospects and ethical dilemmas, especially with AI applications, privacy, and data security (Derin & Hamuddin, 2019)[9].

---

[4] Svartvik, J. L. (1968). The Evans statements : a case for forensic linguistics. In *Acta Universitatis Gothoburgensis eBooks*. https://ci.nii.ac.jp/ncid/BA11447154

[5] Khoyi, A. M., & Behnam, B. (2014). Legal Discourse: Analysis of education and criminal convictions in Iranian courts. *Asian Journal of Education and e-Learning*, *2*(6). http://ajouronline.com/index.php?journal=AJEEL&page=article&op=view&path%5B%5D=1859&path%5B%5D=1031

[6] Kreuz, R. (2023). *Linguistic fingerprints: How Language Creates and Reveals Identity*. Rowman & Littlefield.

[7] Crystal, D. (2008). *A Dictionary of Linguistics and Phonetics* (6th ed.). John Wiley & Sons. https://doi.org/10.1002/9781444302776

[8] MacLeod, N. (2021). Assuming identities online: Authorship synthesis in undercover investigations. In M. Coulthard, A. May, & R. Sousa-Silva (Eds.), *The Routledge Handbook of Forensic Linguistics* (2nd ed., pp. 159–173). Routledge. http://nrl.northumbria.ac.uk/41825/

[9] Derin, T., Evizareza, Deliani, S., & Hamuddin, B. (2019). *Exploring the Past, Present, and Future of Forensic Linguistics Study: A Brief Overview*. The First Conference of Indonesian Community for Forensic Linguistics (KLFI-1), Pekanbaru, Riau, Indonesia.

This research examines the various domains in relation to the applications and technologies in forensic linguistics.

2.3. Directions of Forensic Linguistics:

As a developing field, academics have classified forensic linguistics in several ways. Their use and focus on language in legal situations are often the basis for these categories. The classifications of legal language, courtroom discourse, linguistic evidence, and pragmatics underscore the multidisciplinary essence of forensic linguistics and its vital function within the judicial system. (McMenamin, 2002)[10] argued that:

> "Forensic linguistics is a well established area of applied linguistics. However, when a field such as forensic linguistics goes through the process of defining itself, there are certain to be instances of ambiguity related to what is or is not part of the discipline"(McMenamin, 2002, p. 68)[11]

Major contributors like (Tiersma, 1999)[12], (Coulthard & Johnson, 2010)[13], (McMenamin, 2002)[14], and others have laid the foundations for three directions, enabling the field to evolve into an essential component of modern legal practice. These directions can be an umbrella for the more specified classifications. Nevertheless, this discipline is currently being redefined, especially due to the current emergence of artificial intelligence, machine learning, and natural language processing. Accordingly, this paper sorts these directions of study referring to their bidirectional relation between law and linguistics. These directions will be detailed and selective advancements in AI, machine learning, corpora, and datasets will be reviewed. Mainly, three major border directions can be categorized:

- The Written Language of the Law
- Language Use in Forensic and Judicial Processes
- Linguistic Evidence

These directions are based on the fields of study and their contributors which will be discussed in detail in this paper as follows:

2.3.1. The Written Language of the Law

Building, interpreting, and applying the language of written legal documents such as legislation, contracts, and laws are the primary focuses of this direction. Legal linguistics is an

---

[10] McMenamin, G. R. (2002). *Forensic Linguistics: Advances in forensic stylistics*. CRC Press. https://ci.nii.ac.jp/ncid/BA75691949

[11] McMenamin, G. R. (2002). *Forensic Linguistics: Advances in forensic stylistics*. CRC Press. https://ci.nii.ac.jp/ncid/BA75691949

[12] Tiersma, P. M. (1999). *Legal language*. University of Chicago Press.

[13] Coulthard, M., & Johnson, A. (2010). *The Routledge Handbook of Forensic Linguistics*. Routledge.

[14] Coulthard, M., & Johnson, A. (2010). *The Routledge Handbook of Forensic Linguistics*. Routledge.

academic discipline that studies the pragmatics, syntax, and semantics of the law (Coulthard & Johnson, 2010)[15]. Solan (1993)[16] has done substantial research on the topic of how judges and solicitors read legislation. He delves into the ways that legalese influences judges' decision-making process, *The Language of Judges*. By contrast, Tiersma's (1999)[17] book *Legal Language* is a classic in the field of legal language and its influence on both interpretation and practice. Despite their emphasis on this area in their writings, it was only after scholars and practitioners of law began to delve more into the complexities of legal texts, lawmaking, and interpretation that it was officially acknowledged as a distinct discipline.

Different subfields fall within the Written Language of the Law as follows:

1. Statutory Interpretation

Statute interpretation is a critical aspect in the legal system. It entails reading statutes and determining their meanings and proper application. It is essential that laws use clear language. When laws are either poorly worded or not explicit enough, the courts may reach different conclusions. Many possible judicial conclusions may emerge from this.

One famous example of how confusing language maybe is the 1984 case of Chevron U.S.A. Inc. v. Natural Resources Defense Council, Inc. It laid up a two-pronged process for courts to follow when examining agency interpretations of the law. This case shows how different government agencies and courts might reach different conclusions when using vague legislative language. The enforcement of environmental standards might be affected by this (Gries, 2021)[18].

Solicitors use a variety of tools and information to figure out how to interpret the law. When building corpora, it is possible to utilize readily available sources like Westlaw. The comprehensive collection of laws, court decisions, and legal studies helps lawyers to grasp and use complex legal language. Platforms include the Oyez Project (www.oyez.org) and the US Supreme Court (www.supremecourtus.gov) started granting access to audio recordings of oral arguments put on before the US Supreme Court in 1955. The materials cited by Sprowl (1981)[19] offer forensic linguists studying legal language great possibilities (Coulthard & Johnson, 2010)[20].

---

[15] Coulthard, M., & Johnson, A. (2010). *The Routledge Handbook of Forensic Linguistics*. Routledge.

[16] Solan, L. M. (1993). *The language of judges*. https://doi.org/10.7208/chicago/9780226767895.001.0001

[17] Tiersma, P. M. (1999). *Legal language*. University of Chicago Press.

[18] Gries, S. Th. (2021). Corpora and legal interpretation: Corpus approaches to ordinary meaning in legal interpretation. In M. Coulthard, A. May, & R. Sousa-Silva (Eds.), *The Routledge Handbook of Forensic Linguistics* (2nd ed., pp. 628–643). Routledge.

[19] Sprowl, J. A. (1981). WESTLAW vs LEXIS: computer-assisted legal research comes of age. *Program Electronic Library and Information Systems*, *15*(3), 132–141. https://doi.org/10.1108/eb046827

[20] Coulthard, M., & Johnson, A. (2010). *The Routledge Handbook of Forensic Linguistics*. Routledge.

2. Contract Language

Contracts are essential legal instruments that specify "the rights and duties of the parties concerned" (InfoToHow, 2024, para. 16)[21]. Contracts must have exact wording to avoid disputes and costly litigation caused by poorly worded agreements. Common problems with contract draughting include unclear or conflicting language, vague or unclear terms, and ambiguous terminology (Cao, 2010)[22].

The importance of precise contractual language is highlighted when courts interpret ambiguous terms. The California Supreme Court highlighted the importance of considering the broader context and the parties' intentions when interpreting ambiguous language in Pacific Gas and Electric Co. v. G.W. Thomas Drayage & Rigging Co. (1968). Tiersma argued that "it is nonsense to try to decide whether there is an ambiguity based purely on examination of the text itself …… A California judge still needs to decide, after examining the extrinsic evidence that is being offered, that the text has an ambiguity." (Tiersma, 2010, p. 127)[23]. This case highlights the importance of clear contractual language, as ambiguity may result in expensive disputes and extended litigation (Tiersma, 2010)[24].

Legal professionals frequently employ specialized tools and applications to improve contract drafting and tackle these challenges. "Today's well-established applications include the following: "ActiveDocs Opus, ContractExpress DealBuilder, Exari, HotDocs, Innova, Leaflet, MacPac, Pathagoras, Rapidocs, TheFormTool, and XpressDox" (Lauritsen, 2021, p. 76)[25]. These tools reduce misinterpretation. They provide designs and support for drafting exact and clear contracts. These tools offer uniform language and guidelines. They assist in creating legally valid contracts that reflect the parties' goals (Lauritsen, 2021[26]; Stygall, 2010)[27]. "LexPredict Legal Dataset, which is available at GitHub, and covers courts across the US, Canada, Australia, and

[21] InfoToHow. (2024, January 22). *Essential Elements of a Mortgage Deed: Unveiling Its Components and Protections*. Info to How. Retrieved October 14, 2024, from https://infotohow.com/finance/essential-elements-of-a-mortgage-deed/

[22] Cao, D. (2010). Legal translation: Translating legal language. In M. Coulthard & A. Johnson (Eds.), *The Routledge Handbook of Forensic Linguistics* (1st ed., pp. 78–91). Routledge.

[23] Tiersma, P. M. (2010). *Parchment, paper, pixels: Law and the Technologies of Communication*. University of Chicago Press.

[24] Tiersma, P. M. (2010). *Parchment, paper, pixels: Law and the Technologies of Communication*. University of Chicago Press.

[25] Lauritsen, M. (2021). Document Automation. In D. M. Katz, R. Dolin, & M. J. Bommarito (Eds.), *Legal Informatics*. Cambridge University Press.

[26] Lauritsen, M. (2021). Document Automation. In D. M. Katz, R. Dolin, & M. J. Bommarito (Eds.), *Legal Informatics*. Cambridge University Press.

[27] Stygall, G. (2010). Legal writing: complexity Complex documents/average and not-so-average readers. In M. Coulthard, A. May, & R. Sousa-Silva (Eds.), *The Routledge Handbook of Forensic Linguistics* (2nd ed., pp. 32–47). Routledge. https://doi.org/10.4324/9780203855607-14

Germany."(Bommarito et al., 2021, p. 221)[28]. It analyzes contracts, agreements, and legal documents using machine learning. It can identify contextual patterns (Bommarito et al., 2021)[29]. These developments in technology provide insights that improve the accuracy of drafting.

3. Legal Definitions

The accuracy of legal definitions is crucial. It is a crucial element in influencing the interpretation and implementation of the law. Ambiguities in definitions can provide issues in legal procedures, especially when a single phrase may possess several meanings in diverse circumstances. "Other legal languages likewise have a great deal of technical vocabulary. What can be frustrating for those doing comparative analysis, as well as translators and interpreters, is that legal terminology is extremely parochial." (Tiersma, 2008, p. 16)[30] For instance, the term "reasonable" can have different implications depending on the area of law (contract law, tort law, or criminal law) leading to potential confusion and differing interpretations (Mattila, 2006)[31].

Legal dictionaries, such as Black's Law Dictionary online, serve as vital English legal lexicon for legal professionals seeking clarity on the meanings and implications of legal terms (Tiersma, 2008)[32]. It is typically regarded as the most authoritative American legal dictionary (Solan, 2006)[33]. These dictionaries have uniform definitions. They help to ensure consistent legal language application and interpretation. Moreover, law case study clarifies understanding by showing how different contexts have been interpreted by courts on specific legal language.

In this regard, current technologies are rather important. Built on IBM's Watson, ROSS Intelligence "attempted to create a generative AI legal search engine to make it easier for users to get direct quotes from judicial opinion" (KISBYE, 2024, p. 8)[34]. It can surface pertinent cases and

[28] Bommarito, M. J., II, Katz, D. M., & Detterman, E. M. (2021). LexNLP: Natural language processing and information extraction for legal and regulatory texts. In *Edward Elgar Publishing eBooks*. https://doi.org/10.4337/9781788972826.00017

[29] Bommarito, M. J., II, Katz, D. M., & Detterman, E. M. (2021). LexNLP: Natural language processing and information extraction for legal and regulatory texts. In *Edward Elgar Publishing eBooks*. https://doi.org/10.4337/9781788972826.00017

[30] Tiersma, P. (2008). The nature of legal language. In J. Gibbons & M. Teresa Turell (Eds.), *Dimensions of Forensic Linguistics* (pp. 7–25). John Benjamins B.V.

[31] Mattila, H. (2006). Legal language: history. In *Elsevier eBooks* (pp. 8–13). https://doi.org/10.1016/b0-08-044854-2/04504-1

[32] Mattila, H. (2006). Legal language: history. In *Elsevier eBooks* (pp. 8–13). https://doi.org/10.1016/b0-08-044854-2/04504-1

[33] Solan, L. (2006). Definition/Rules in legal language. In K. Brown (Ed.), *Encyclopedia of Language & Linguistics* (2nd ed., pp. 403–409). Elsevier. https://doi.org/10.1016/b0-08-044854-2/00692-1

[34] KISBYE, E. (2024). Man vs. Machine: AI's Impact on Intellectual Property and Copyright Law. *TRINITY COLLEGE UNDERGRADUATE LAW JOURNAL*, *1*.

statutes (Sobowale, 2016)[35] and examine legal concerns in straightforward terms. It uses AI to predict how certain terms are treated in legal cases, helping to understand legal terminology in a way that aligns with court interpretations. Some corpora include a large volume of legal texts, statutes, and judicial opinions. These corpora facilitate the examination of the evolution of legal language interpretation over time. A corpus aids legal scholars and practitioners in analyzing the application of specific language in rulings, providing insights that extend beyond dictionary definitions. The predominant corpora in this domain include the LawNet Corpus (Li & Li, 2021)[36] and the Corpus of US Supreme Court Opinions (SCOTUS Corpus) (Yuan & Liberman, 2008)[37].

4. Plain Language Movement

The plain language movement "propagates a simpler and more reader-friendly legal language" (Hiltunen, 2012, p. 50)[38]. It aims to make legal texts easier for the general public to understand. Legal terminology and intricate language can obstruct understanding, confusing and misinterpreting persons attempting to navigate the legal system. Consequently, legal professionals and organizations have increasingly acknowledged the necessity for clear and succinct communication (Hiltunen, 2012)[39].

Instruments like the Gunning Fog Index Calculator are employed to evaluate the readability of legal texts, assisting legal practitioners in creating materials that are more comprehensible for laypersons. This approach enhances public trust in legal institutions. It also encourages educated citizen engagement in legal procedures (Timaná et al., 2020)[40].

The U.S. federal government has promoted the use of simple language in official documents, acknowledging its capacity to enhance public comprehension of legal rights and obligations. The Plain Writing Act of 2010 is the first explicit plain language mandate to be legislated at the federal

[35] Sobowale, J. (2016). How artificial intelligence is transforming the legal profession. *ABA Journal*, *1*(1).

[36] Li, X., & Li, J. (2021). Law - net: A New Method for Legal Text Mining. *Scientific Journal of Technology*, *3*(7). https://api.semanticscholar.org/CorpusID:251244662

[37] Yuan, J., & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *The Journal of the Acoustical Society of America*, *123*(5_Supplement), 3878. https://doi.org/10.1121/1.2935783

[38] Hiltunen, R. (2012). Th e Grammar and Structure of Legal Texts. In P. M. Tiersma & L. Solan (Eds.), *The Oxford Handbook of Language and Law* (pp. 39–51). Oxford University Press.

[39] Hiltunen, R. (2012). Th e Grammar and Structure of Legal Texts. In P. M. Tiersma & L. Solan (Eds.), *The Oxford Handbook of Language and Law* (pp. 39–51). Oxford University Press.

[40] Timaná, L. C. R., Lozano, D. F. S., & García, J. F. C. (2020). Software to determine the readability of written documents by implementing a variation of the gunning fog index using the Google Linguistic Corpus. In *Communications in computer and information science* (pp. 409–420). https://doi.org/10.1007/978-3-030-42517-3_31

level in the United States. It requires federal agencies to utilize straightforward communication in their publications (Stabler, 2014)[41].

5. Translation of Legal Texts

Translating legal texts presents a unique challenge. Legal vocabulary sometimes lacks direct counterparts in other languages. This can result in possible misinterpretations within multilingual legal frameworks. The precise translation is crucial for successfully communicating legal rights and obligations, particularly in multilingual cultures (Powell, 2010)[42].

Applications such as SDL Trados Studio aid legal practitioners in overcoming linguistic obstacles. It can "translate legal documents with complex legal terminology."(PUTRI, 2024, p.2)[43]. The translation process is efficiently aided by its terminology management and translation memory functions (Muravev, 2020)[44]. These tools assist translators in ensuring precision in legal terms and preserve the source documents' purpose and content.

The accuracy of interpretation greatly affects court proceedings, highlighting the essential need for professional language services to guarantee equal access to justice (Angermeyer, 2008)[45].

6. Jury Instructions

Jury instructions are "the legal texts produced by judicial committees and delivered by judges"(Heffer, 2008, p.47)[46]. These instructions are often presented in a template format, but they do not come from legal documents. Their effectiveness in understanding legal principles during case deliberations is defined. "Judges need to instruct juries on the technicalities of deliberation (selecting a foreman, unanimous verdict, retirement). These are often set out in template form but do not derive from legal texts"(Heffer, 2008, p.51)[47]. These instructions must be clear and precise.

---

[41] Stabler, R. (2014). "What We've Got Here is Failure to Communicate": The Plain Writing Act of 2010. *SSRN Electronic Journal*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2574207

[42] Powell, M. B., Wright, R., & Hughes-Scholes, C. H. (2010). Contrasting the perceptions of child testimony experts, prosecutors and police officers regarding individual child abuse interviews. *Psychiatry Psychology and Law*, *18*(1), 33–43. https://doi.org/10.1080/13218710903566995

[43] PUTRI, S. H. (2024). *An Evaluation of Usability Computer-Aided Translation Tools Among Bispro Students*. Jakarta State Polytechnic.

[44] Muravev, Y. (2020). Machine translation and legal tech in legal translation training. *Proceedings of the International Scientific Conference - Digital Transformation on Manufacturing, Infrastructure and Service*, 1–7. https://doi.org/10.1145/3446434.3446553

[45] Angermeyer, P. S. (2008). Creating monolingualism in the multilingual courtroom. *Sociolinguistic Studies*, *2*(3), 385–404. https://doi.org/10.1558/sols.v2i3.385

[46] Heffer, C. (2008). The language and communication of jury instruction. In J. Gibbons & M. Teresa Turell (Eds.), *Dimensions of Forensic Linguistics* (pp. 47–65). https://doi.org/10.1075/aals.5.05hef

[47] Heffer, C. (2008). The language and communication of jury instruction. In J. Gibbons & M. Teresa Turell (Eds.), *Dimensions of Forensic Linguistics* (pp. 47–65). https://doi.org/10.1075/aals.5.05hef

Inaccurate instructions can lead to ambiguity and misinterpretation (Dumas, 2000)[48]. Research indicates that juries often struggle to understand complex legal terminology and concepts. Consequently, solicitors maintain a 'presumption of comprehension' concerning jury instructions (Tiersma, 2001)[49].

AI-driven tools, machine learning, and specialized corpora are vital in analyzing and composing jury instructions. These tools help produce and evaluate jury instructions by improving clarity, homogeneity, and legal norm compliance.

Tools like LEGAL-BERT can analyze linguistic patterns. Platforms like ROSS Intelligence provide legal insights (KISBYE, 2024)[50]. AI tools like SHAP help classify or rewrite unclear instructions and detect implicit biases (Lundberg & Lee, 2017)[51]. Specialized corpora, such as Cornell's Legal Information Institute database or Corpus of Contemporary American English (henceforth: COCA), can contextualize jury instructions within legal and general language usage (Mubin et al., 2024)[52]. AI-powered technologies such as LexisNexis Context and Westlaw Edge examine judicial interpretations of specific terminology (Ahlbrand, 2020)[53].

7. Patent Law and Intellectual Property

Patent law and intellectual property have complexity depending on precise language. Careful formulation of patent claims helps to protect inventors' rights and allow appropriate interpretations. Linguistic complexity arises in characterizing the degree of protection granted by patents, usually leading to disputes over infringement.

Tools like Innography help to examine patent claims and provide information that helps to resolve conflicts and improve the caliber of patent drafting. (Yang & Yu, 2019)[54]. The significance of linguistic clarity in patent language is emphasized in instances where courts are required to interpret ambiguous claims. In Markman v. Westview Instruments, Inc. (1996), the U.S. Supreme

[48] Dumas, B. K. (2000). U.S. Pattern Jury Instructions: Problems and proposals. *International Journal of Speech, Language and the Law*, *7*(1), 49–71.

[49] Tiersma, P. (2001). The Rocky Road to Legal Reform: Improving the language of jury instructions. *Brooklyn Law Review*, *66*(4), 1081. https://brooklynworks.brooklaw.edu/cgi/viewcontent.cgi?article=1709&context=blr

[50] KISBYE, E. (2024). Man vs. Machine: AI's Impact on Intellectual Property and Copyright Law. *TRINITY COLLEGE UNDERGRADUATE LAW JOURNAL*, *1*.

[51] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, arXiv:1705.07874.

[52] Mubin, R., Ozoda, M., Mashhura, R., & Ozoda, H. (2024). THE USAGE OF COCA AND INVESTIGATIONS OF VOCABULARY FORMING. *Ta'lim Innovatsiyasi Va Integratsiyasi*, *21*(3), 88–92.

[53] Ahlbrand, A. A. (2020). Analyzing Analytics: Litigation Analytics in Bloomberg Law, Westlaw Edge, and Lexis Advance. *42 the CRIV Sheet 9 (Feb. 2020)*, *2*.
https://www.repository.law.indiana.edu/cgi/viewcontent.cgi?article=3954&context=facpub

[54] Yang, X., & Yu, X. (2019). *Identifying Patent Risks in Technological Competition: A Patent Analysis of Artificial Intelligence Industry*. 8th International Conference on Industrial Technology and Management (ICITM), Cambridge, United Kingdom of Great Britain and Northern Ireland. https://doi.org/10.1109/icitm.2019.8710719

Court highlighted the imperative for precise and unequivocal patent claims, accentuating the critical importance of language in assessing the validity and enforceability of patents (Pieper, 1998)[55].

2.3.2. Language Use in Forensic and Judicial Processes

Examining the utility of legal language in judicial, law enforcement, and courtroom discussions, forensic linguistics focusses on the application of language within forensic and legal environments. This sector addresses the challenges of law interpretation, comprehension of courtroom discussion, and the provision of clear, easily available communication for every person engaged in judicial procedures (Ali & Algane, 2013)[56].

Contributions to courtroom discourse and police interrogations have helped establishing this domain as a vital aspect of forensic linguistics. There is great advancement of discourse analysis techniques. Implementations in legal environments underscored the significance of examining spoken language within legal frameworks.

Principal applications employed in this domain encompass the examination of jury instructions, witness testimonies, confessions, and judicial rulings. Progress in artificial intelligence and corpus linguistics is improving this domain by increasing clarity and diminishing ambiguity.

This field focuses on various important subfields as follows:

1. Courtroom Discourse

Courtroom discourse is a "public, face-to-face verbal interaction between individuals in a forensic setting, aimed to adjudicate a particular dispute and arrive at a determination by the application of the law to the facts, be it a criminal, civil, military or ecclesiastical matter."(Chaemsaithong, 2021, p.212)[57] The interaction is both oral and written among judges, attorneys, witnesses, defendants, and jurors. This interaction is governed by legal standards, established protocols, and the necessity for clarity and precision. The primary characteristics of courtroom discourse refer to question-and-answer formats. This discourse is crucial for effective

[55] B Pieper, D. (1998). The Appropriate Judicial Actor for Patent Interpretation: A Commentary on the Supreme Court's Decision in Markman v. Westview Instruments, Inc. *51 Ark. L. Rev.*, 159.

[56] Ali, S. A., & Algane, M. (2013). The Role of Forensic Translation in Courtrooms Contexts. *Arab World English Journal*, *4*.

[57] Chaemsaithong, K. (2021). Advances in studies of the historical courtroom: (Con)Textual, ideational and interpersonal dimensions. In M. Coulthard, A. May, & R. Sousa-Silva (Eds.), *The Routledge Handbook of Forensic Linguistics* (2nd ed., pp. 211–227). Routledge.

argumentation, evidence presentation, and verdict delivery, all of which mirror broader societal values and justice systems (Gibbons, 2003)[58].

CAQDAS (Computer-Assisted Qualitative Data Analysis Software) is a prominent discourse analysis technology. It is used to analyze courtroom conversations and identify linguistic patterns. "Mapping tools can provide visual ways of making connections or viewing patterns and relationships in the data at an abstract level"(Silver & Lewins, 2014, p.56)[59] Analyzing the language witnesses use could also help one understand their reliability and the caliber of their testimony.

2. Police Interviews and Interrogations

Language structure and function play a vital role in police interviews and interrogations. Question phrasing and language choices can significantly influence the information witnesses and suspects provide. Effective communication is essential in getting accurate information during interviews, as emphasized by the PEACE paradigm (Preparation and Planning, Engage and Explain, Account, Closure, and Evaluate) (Clarke & Milne, 2001)[60].

Software programs for interrogation analysis are extensively utilized. Diverse software applications offer extensive solutions. i2 Analyst's Notebook facilitates the identification of patterns in interrogation data (Zaiets, 2015)[61], whereas Praat examines speech for phonetic indicators, which are beneficial for evaluating voice stress (Magdin et al., 2019)[62]. Speech-to-text technologies such as Google Cloud Speech-to-Text, in conjunction with NLTK or spaCy, facilitate language pattern and sentiment analysis in transcribed interviews (Apturkar et al., 2020)[63]. Dragon Law Enforcement provides precise legal transcription services for law enforcement agencies (Moe,

[58] Gibbons, J. (2003). *Forensic Linguistics: An Introduction to language in the justice system*. Blackwell. https://ci.nii.ac.jp/ncid/BA60279023

[59] Silver, C., & Lewins, A. (2014). *Using Software in Qualitative Research: A Step-by-Step Guide*. https://doi.org/10.4135/9781473906907

[60] Clarke, C., & Milne, R. J. (2001). *National evaluation of the PEACE investigative interviewing course*. Interdepartmental and Cross-Faculty Research Groups and Centres (until 2017), Services and Administration (until 2017). Retrieved October 8, 2024, from https://eprints.kingston.ac.uk/31400/

[61] Zaiets, O. (2015). Application Software IBM I2 Analyst's Notebook in Law Enforcement in Ukraine for Pretrial Investigation of Criminal Offenses. *European Reforms Bulletin*, *1*, 69–73.

[62] Magdin, M., Sulka, T., Tomanová, J., & Vozár, M. (2019). Voice analysis using PRAAT software and classification of user emotional state. *International Journal of Interactive Multimedia and Artificial Intelligence*, *5*(6), 33. https://doi.org/10.9781/ijimai.2019.03.004

[63] Apturkar, A., Iliev, A. I., Anand, A., Oli, A., Siddenki, S. R., & Meka, V. R. (2020). Sentiment Analysis of Speech with Application to Various Languages. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, *10*, 103–118. https://doi.org/10.55630/dipp.2020.10.6

2023)[64]. Collectively, these instruments augment the comprehension and evaluation of language and conduct within forensic settings.

3. Language in Judicial Decisions

The legal terms employed in judicial rulings are essential. They influence the comprehension of legal doctrines and establish precedents for subsequent cases. Judges must articulate intricate legal principles simply and convincingly to guarantee that their decisions are comprehended and adhered to.

Resources such as LexisNexis to access judicial rulings are usually utilized by legal professionals. These resources allow them to examine linguistic patterns and interpretative methodologies used by judges (Hall & Windett, 2013)[65]. Legal professionals can understand judicial thinking and anticipate future court interpretations of similar cases by analyzing the language used in court opinions.

4. Legal Pragmatics

Legal pragmatics involves the analysis of how context influences the interpretation of legal terms. Understanding the pragmatic aspects of language is essential for properly interpreting legal texts and guaranteeing the efficacy of legal communications (Gao, 2010)[66].

In legal contexts, pragmatic and discourse analysis can be carried out using particular software tools and corpora. Notable corpora include the British National Corpus (BNC) Legal Subcorpus for British legal discourse (Bilić & Gaspar, 2018)[67] and the SCOTUS Corpus for U.S. Supreme Court opinions (Yuan & Liberman, 2008)[68]. AntConc is a software tool that enables the detection of pragmatic signals. LIWC provides a psycholinguistic analysis of tone and stance (Boyd et al., 2022)[69]. These developments improve the understanding of the pragmatics of legal language.

---

[64] Moe, M. K. (2023). *Post-processing automatic speech recognition transcriptions: A study for investigative interviews* [Master thesis]. NTNU.

[65] Hall, M. E. K., & Windett, J. H. (2013). New data on state supreme court cases. *State Politics & Policy Quarterly*, *13*(4), 427–445. https://doi.org/10.1177/1532440013497971

[66] Gao, J. (2010). Review and prospects of the research of forensic linguistics in China. *Asian Social Science*, *6*(10). https://doi.org/10.5539/ass.v6n10p127

[67] Bilić, M., & Gaspar, A. (2018). Extraction of Phrasal Verbs from the Comparable English Corpus of Legal Texts. *International Journal of English Language & Translation Studies*, *6*(2), 184–194.

[68] Yuan, J., & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *The Journal of the Acoustical Society of America*, *123*(5_Supplement), 3878. https://doi.org/10.1121/1.2935783

[69] Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin*. https://www.liwc.app

5. Multilingual Courtrooms

The increasing diversity of countries directly affects the unique problems that multilingual courtrooms provide for legal practitioners. When many languages are involved, judicial procedures could be more complex and call for strategies for good communication to ensure understanding amongst all parties (Angermeyer, 2008)[70].

Different useful tools are used in multilingual courts. The Court Interpreters Database (CID) and the European Parliament Interpreting Corpus (EPIC) (Russo et al., 2012[71]; Núñez, 2017)[72] are two databases that provide useful information on real-time interpretation and speech transitions. The Multilingual Corpus of Survey Questionnaires (MCSQ) can be "helpful for training translators in the highly technical field of questionnaire translation, by providing examples of phrases, lexicon, response options, etc."(Zavala-Rojas et al., 2022, p.76)[73] Translational tone can be better evaluated using speech and language technology (SLT) tools (Schwab et al., 2014)[74].

6. Emergency Calls

"The delivery of police, fire, or emergency medical service depends on a complex communications system."(Whalen, 1995, p.1)[75]. Certain questions and responses in emergency calls "can have a significant detrimental effect on witness memory, as well as the potential to mislead a police investigation or render witness evidence inadmissible in court"(Gabbert et al., 2015, p.109)[76]. Effective emergency communication is essential for rescuers to handle situations promptly and accurately.

---

[70] Angermeyer, P. S. (2008). Creating monolingualism in the multilingual courtroom. *Sociolinguistic Studies*, *2*(3), 385–404. https://doi.org/10.1558/sols.v2i3.385

[71] Russo, M., Bendazzoli, C., Sandrelli, A., & Spinolo, N. (2012). The European Parliament Interpreting Corpus (EPIC): implementation and developments. In *Peter Lang eBooks* (Vol. 147, pp. 53–90). https://cris.unibo.it/handle/11585/121334

[72] Núñez, G. G. (2017). Law and translation at the US–Mexico border: Translation policy in a diglossic setting. In G. G. Núñez & R. Meylaerts (Eds.), *Translation and Public Policy* (1st ed., pp. 152–170). Routledge.

[73] Zavala-Rojas, D., Sorato, D., Hareide, L., & Hofland, K. (2022). The Multilingual Corpus of Survey Questionnaires: A tool for refining survey translation. *Meta Journal Des Traducteurs*, *67*(1), 71–93. https://doi.org/10.7202/1092191ar

[74] Schwab, D., Tchechmedjiev, A., Goulian, J., & Sérasset, G. (2014). Comparisons of relatedness measures through a word sense disambiguation task. *Text, Speech and Language Technology*, 221–243. https://doi.org/10.1007/978-3-319-08043-7_13

[75] Whalen, J. (1995). Expert systems versus systems for experts: computer-aided dispatch as a support system in real-world environments. In *Cambridge University Press eBooks* (pp. 161–183). https://dl.acm.org/citation.cfm?id=214811.214832

[76] Gabbert, F., Hope, L., Carter, E., Boon, R., & Fisher, R. (2015). The role of initial witness accounts within the investigative process. *Communication in Investigative and Legal Contexts: Integrated Approaches From Forensic Psychology, Linguistics and Law Enforcement*, 107–131.

Analytical tools are very useful in emergency calls. Computer-aided dispatch (CAD) systems aid in "automated call-taking, case number assignment, address verification, status monitoring of units in the field, dispatch recommendations, and information management"(Whalen, 1995, p.2)[77]. It can also aid law enforcement and emergency workers in determining the nature and severity of calls (Whalen, 1995)[78]. IBM Watson Tone Analyzer assesses callers' emotional states and "calculates the percentage of sadness, joy, fear, disguest, and anger emotions in the text."(VanBuskirk & Letson, 2020, p.4)[79]. Programs like HateSonar identify hostile speech patterns, while Perspective API "supports six production attributes namely; Toxicity, Severe_Toxicity, Identity Attack, Insult, Profanity and Threat"(Kwarteng et al., 2022, p.165)[80]. Databases like the 911 Call Database provide transcripts for linguistic research. Transcription and annotation tools such as Express Scribe (Moe, 2023)[81] can help analyze these databases. Using these tools and capabilities is vital in improving emergency call handling.

7. Linguistic Profiling in Court

Linguistic profiling is the examination of a person's linguistic characteristics. It "can be used as a tool in courts. As many courts have already recognized, testimony based on linguistic profiling is helpful in establishing the issue of identity." (MacNeal et al., 2019)[82] Dialect, vocabulary, and grammatical constructions can reveal the educational level and intentions of speakers (Bosco et al., 2021)[83].

Instruments for profiling and analyzing linguistic characteristics, such as Forensic Voice Comparison Software which "after in-depth analysis of the different voice patterns involved– is to report strength of evidence, i.e. whether the analyzed evidence supports the same-speaker

[77] Whalen, J. (1995). Expert systems versus systems for experts: computer-aided dispatch as a support system in real-world environments. In *Cambridge University Press eBooks* (pp. 161–183). https://dl.acm.org/citation.cfm?id=214811.214832

[78] Whalen, J. (1995). Expert systems versus systems for experts: computer-aided dispatch as a support system in real-world environments. In *Cambridge University Press eBooks* (pp. 161–183). https://dl.acm.org/citation.cfm?id=214811.214832

[79] VanBuskirk, K., & Letson, M. R. (2020). An IBM Watson tone analysis of selected judicial decisions. *The Scribes Journal of Legal Writing*, *19*, 25–48.

[80] Kwarteng, J., Perfumi, S. C., Farrell, T., Third, A., & Fernandez, M. (2022). Misogynoir: challenges in detecting intersectional hate. *Social Network Analysis and Mining*, *12*(1). https://doi.org/10.1007/s13278-022-00993-7

[81] Moe, M. K. (2023). *Post-processing automatic speech recognition transcriptions: A study for investigative interviews* [Master thesis]. NTNU.

[82] MacNeal, A., Fiallo, K., Jones, A., Jones, S., Laureano, S., Monjarrez, M., & Xu, Y. (2019). Sounding black: the legal implications of linguistic profiling. *Northeastern University Working Papers in Linguistics*, *4*.

[83] Lo Bosco, G., Pilato, G., & Schicchi, D. (2021). DeepEva: A deep neural network architecture for assessing sentence complexity in Italian and English languages. *Array*, *12*, 100097. https://doi.org/10.1016/j.array.2021.100097

hypothesis"(Jessen, 2018, p.219)[84]. Experts can provide significant insights into the person's credibility of testimony.

8. Interpreter Use in Legal Settings

Translators and interpreters have a significant role in legal contexts. Many cases involve non-native language speakers, which in turn requires skilled bilingual professionals. "The job of the interpreter is possibly the most difficult and demanding of all the officers of the court" (Tallentire, 2009, p.151)[85]. Translators and interpreters communicate between legal practitioners and individuals who are not proficient in the court's language (Leung, 2008)[86].

In legal contexts, prominent technologies encompass Zoom Interpretation for remote multilingual assistance. Other translation assistants are "Google Translate, Microsoft Translator, and DeepL represent some of the foremost translation services available online."(Lu et al., 2023, p.10)[87]. KUDO AI is "a computer-assisted interpreting tool specifically designed for the integration in [remote simultaneous interpreting] RSI scenarios. Interpreter Assist comprises two main features: an automatic glossary creation tool and a real-time suggestion system."(Fantinuoli et al., 2022, p.1)[88]. OpenAI's Whisper provides sophisticated multilingual automatic speech recognition for courtroom contexts. It "is an advanced ASR system developed by OpenAI, boasting high accuracy in transcribing audio into text. Trained on an extensive 680 000 h of multilingual and multitask data collected from the web"(Graham & Roll, 2024, p.2)[89]. These innovations enhance precision, effectiveness, and adherence to ethical standards in legal interpretation.

9. Child Testimonies

For forensic linguistics, analyzing young testimonies presents different challenges. Witnesses must have skills in encoding, storing, retrieving memories, and speaking them in a foreign context.

---

[84] Jessen, M. (2018). Forensic voice comparison. In *Handbook of Communication in the Legal Sphere* (pp. 219–255). De Gruyter Mouton. https://doi.org/10.1515/9781614514664-012

[85] Tallentire, G. (2009). The Hong Kong (police) magistrate in the 1880s and 1990s. In G. Bickley (Ed.), *A Magistrate's Court in Nineteenth Century Hong Kong: Court in Time* (pp. 133–143).

[86] Leung, E. S. M. (2008). Interpreting for the minority, interpreting for the power. In J. Gibbons & M. Teresa Turell (Eds.), *Dimensions of Forensic Linguistics* (pp. 197–211). John Benjamins B.V.

[87] Iu, K. Y., Tu, L., & Liu, Y. (2024). AI-Powered Legal Translation: Can ChatGPT Facilitate the Development of Chinese Common Law? *SSRN*. https://doi.org/10.2139/ssrn.4920531

[88] Fantinuoli, C., Marchesini, G., Landan, D., & Horak, L. (2022). KUDO Interpreter Assist: Automated real-time support for remote interpretation. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2201.01800

[89] Graham, C., & Roll, N. (2024). Evaluating OpenAI's Whisper ASR: Performance analysis across diverse accents and speaker traits. *JASA Express Letters*, *4*(2). https://doi.org/10.1121/10.0024876

Unfortunately, children are still learning these skills. A gap exists between legal requirements and young children's capacities. (SAYWITZ, 2002)[90]

Software solutions help gather and evaluate child testimony, arming professionals with tools to handle the nuances of child language development. Avaya OneCloud CPaaS, which records audio and video and uses emotion detection to help interviewers assess comfort levels, is one AI tool for child witness interviews (Carias, 2023)[91].

2.3.3. Linguistic Evidence

This area pertains to the application of linguistic methodologies to furnish evidence in legal proceedings. This encompasses the analysis of both written and oral materials. The analysis can disclose trends, authorship, or intent that may function as evidence (Fowler et al., 2015)[92]. Methods such as authorship attribution, discourse analysis, and voice recognition are utilized to discern essential traits and irregularities in language usage. They can elucidate concealed significances, evaluate credibility, and discern potential threats. Linguistic evidence is essential for ensuring equitable and precise legal results (Olsson, 2009)[93].

The main subfields of this category can be enumerated as follows:

1. Authorship Attribution

Authorship attribution is a very important field in forensic linguistics. It entails identifying the probable author of a contested work (Hilton & Holmes, 1993)[94]. This procedure depends on the examination of linguistic characteristics, such as vocabulary, grammar, and stylistic indicators, "Individual differences in writing style are related to individual choices of alternative forms" (McMenamin, 2010, p. 492)[95].

---

[90] SAYWITZ, K. (2002). Developmental Underpinnings of Children's Testimony. In H. L. Westcott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's Testimony* (pp. 3–19). John Wiley & Sons, Ltd.

[91] Carias, J. M. A. (2023). Desarrollo de aplicación móvil y desarrollo de aplicación de comercio electrónico. *CRAI*, https://repositorio.unitec.edu/xmlui/handle/123456789/10422.

[92] Fowler, Y., Vaughan, M., & Wheatcroft, J. (2016). The Interpreter-Mediated Police Interview. In G. Oxburgh, T. Myklebust, T. Grant, & R. J. Milne (Eds.), *Communication in investigative and legal contexts* (1st ed., pp. 315–333). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118769133

[93] Olsson, J. (2009). *WordCrime: Solving crime through forensic linguistics*. A&C Black. http://ci.nii.ac.jp/ncid/BA91502286

[94] Hilton, M. L., & Holmes, D. I. (1993). An assessment of cumulative sum charts for authorship attribution. *Literary and Linguistic Computing*, *8*(2), 73–80. https://doi.org/10.1093/llc/8.2.73

[95] McMenamin, G. R. (2010). Forensic stylistics: Theory and practice of forensic stylistics. In M. Coulthard & A. Johnson (Eds.), *The Routledge Handbook of Forensic Linguistics* (1st ed., pp. 487–507). Routledge.

JGAAP is one tool that aids with forensic authorship analysis. It offers categorization and style algorithms (Juola & Vescovi, 2010)[96]. "it could as easily be used to classify documents by age, authorial gender, genre, or many other categories" (Juola, 2009, p. 4)[97]. Custom model construction is made easier by Weka and Python-based stylometry tools (Gupta et al., 2022)[98]. For author-specific linguistic analysis, sophisticated NLP models like BERT can be trained (Fabien et al., 2020)[99]. For stylometric analysis, large text samples are available in databases such as the British National Corpus (Aston & Burnard, 1998)[100]. Accurate authorship analysis is made easier by these resources.

2. Voice Identification and Forensic Phonetics

Forensic phonetics is mostly concerned on the identification process by means of voice features. By means of spectrographic analysis (Umiyati, 2020)[101], forensic linguists can detect speakers by means of their sonic traits.

Pitch, formants, and intensity can all be examined in a comprehensive speech analysis using forensic phonetics techniques such as Praat (Magdin et al., 2019)[102]. Nuance provides AI-powered voice recognition for speaker verification (Hasan et al., 2021)[103], whereas OpenSMILE extracts

[96] Juola, P., & Vescovi, D. (2010). Empirical evaluation of authorship obfuscation using JGAAP. In *Proceedings of the 3rd ACM workshop on Artificial Intelligence and Security* (Vol. 8, pp. 14–18). https://doi.org/10.1145/1866423.1866427

[97] Juola, P. (2009). JGAAP: A System for Comparative Evaluation of Authorship Attribution. *Journal of the Chicago Colloquium on Digital Humanities and Computer Science*, *1*(1). https://doi.org/10.6082/m1n29v4z

[98] Gupta, S., Patra, T. K., & Chaudhuri, C. (2022). Role of Machine Learning in Authorship Attribution with Select Stylometric Features. In *Lecture notes in networks and systems* (pp. 920–932). https://doi.org/10.1007/978-3-030-96308-8_86

[99] Fabien, M., Villatoro-Tello, E., Motlicek, P., & Parida, S. (2020, December 1). *BertAA : BERT fine-tuning for Authorship Attribution*. ACL Anthology. Retrieved November 7, 2024, from https://aclanthology.org/2020.icon-main.16

[100] Aston, G., & Burnard, L. (1998). *The BNC Handbook: Exploring the British National Corpus with SARA*. http://corpus.leeds.ac.uk/teaching/aston-burnard-bnc.pdf

[101] Umiyati, M. (2020). A Literature review of Forensic Linguistics. *International Journal of Forensic Linguistics*, *1*(1), 23–29.

[102] Magdin, M., Sulka, T., Tomanová, J., & Vozár, M. (2019). Voice analysis using PRAAT software and classification of user emotional state. *International Journal of Interactive Multimedia and Artificial Intelligence*, *5*(6), 33. https://doi.org/10.9781/ijimai.2019.03.004

[103] Hasan, I., Rizvi, S., Jain, S., & Huria, S. (2021). The AI enabled Chatbot Framework for Intelligent Citizen-Government Interaction for Delivery of Services. *International Conference on Computing for Sustainable Global Development*, 601–606. https://doi.org/10.1109/indiacom51348.2021.00106

audio features for profiling (Kumar et al., 2022)[104]. Notable datasets include the VoxCeleb Dataset, which contains diverse speech samples for accent analysis (Nagrani et al., 2019)[105].

3. Threat Assessment and Analysis

Evaluation of threats expressed in oral or written language depends on linguistic study. By improving their awareness of the linguistic elements in threatening communications, law enforcement agencies can assess the degree of threats and create suitable reactions (Meloy & Hoffmann, 2021)[106].

Advances in AI systems, exemplified by Affectiva's Emotion AI and Replika's chatbot technology, analyze vocal emotions to gauge threat intensity (Dew, 2023)[107]. LIWC detects aggressive language and psychological markers in a text (Boyd et al., 2022)[108]. To identify toxicity in written messages "toxic content classifiers for Jigsaw's Perspective API" (Lees et al., 2022, p. 3198)[109] is a new framework for accurate toxic content detection. Important datasets include GDELT, which tracks language trends related to conflicts worldwide and helps with proactive threat identification (Qiao et al., 2017)[110], and Darknet Corpora, which analyses cyber-threat language (Ke et al., 2023)[111].

4. Linguistic Deception Detection

The analysis of language for the purpose of detecting deception has become a focal point in both legal and psychological studies. Key linguistic variables facilitate the analyst's acquisition of deeper insights into the language elements of deception and fraud. Research indicates that

---

[104] Kumar, D., Patil, P. K. V., Agarwal, A., & Prasanna, S. R. M. (2022). Fake speech detection using OpenSMILE features. In *Lecture notes in computer science* (pp. 404–415). https://doi.org/10.1007/978-3-031-20980-2_35

[105] Nagrani, A., Chung, J. S., Xie, W., & Zisserman, A. (2019). Voxceleb: Large-scale speaker verification in the wild. *Computer Speech & Language*, *60*, 101027. https://doi.org/10.1016/j.csl.2019.101027

[106] Meloy, J. R., & Hoffmann, J. (2014). *International Handbook of Threat Assessment*. Oxford University Press, USA.

[107] Dew, R. (2023). *The empathetic algorithm: Leveraging AI for Next-Level CX* (1st ed.). CapFeather Global.

[108] Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin*. https://www.liwc.app

[109] Lees, A., Tran, V. Q., Tay, Y., Sorensen, J., Gupta, J., Metzler, D., & Vasserman, L. (2022). A new generation of perspective API: efficient multilingual character-level transformers. *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/3534678.3539147

[110] Qiao, F., Li, P., Zhang, X., Ding, Z., Cheng, J., & Wang, H. (2017). Predicting Social Unrest Events with Hidden Markov Models Using GDELT. *Discrete Dynamics in Nature and Society*, *2017*, 1–13. https://doi.org/10.1155/2017/8180272

[111] Ke, L., Xiao, P., Chen, X., Yu, S., Chen, X., & Wang, H. (2023). A novel cross-domain adaptation framework for unsupervised criminal jargon detection via pre-trained contextual embedding of darknet corpus. *Expert Systems With Applications*, *242*, 122715. https://doi.org/10.1016/j.eswa.2023.122715

deceptive statements frequently display specific linguistic characteristics, including a higher prevalence of passive voice, a reduced use of first-person pronouns, and increased complexity (Eggington, 2008)[112].

LIWC (Linguistic Inquiry and Word Count) "can be used both to clarify the current knowledge about linguistic deception indicators and to look for new features that distinguish true statements from false statements" (Sarzynska-Wawer et al., 2023)[113]. VeriPol analyses police statements for indicators of deception (Frant, 2020)[114]. NLP models such as BERT are capable of being trained to identify nuanced patterns of deceit (Fabien et al., 2020)[115]. Important datasets include the CSC Deception Corpus, which contains samples of deceptive speech (Enos et al., 2007)[116]. Evaluating the language in statements allows experts to assess the credibility of witnesses and suspects, thereby enhancing the investigative process.

5. Plagiarism and Copyright Infringement

A crucial component of forensic linguistics is plagiarism. "In recent years high- profile cases involving politicians and famous journalists have shown that plagiarism is not a 'mere' academic problem" (Sousa-Silva, 2021, p. 576)[117]. When it comes to identifying plagiarism and copyright violations in written content, linguistic analysis is crucial. Similarities in language, structure, and style can be found by comparing texts.

Tools like Turnitin and CopyLeaks use AI to detect duplicated or rephrased content, even across multiple languages (Uzun, 2023)[118]. Grammarly's Plagiarism Checker scans billions of sources for content similarity (Eyong, 2022)[119]. Important datasets that support precise

---

[112] Eggington, W. (2008). Deception and fraud. In J. Gibbons & M. Teresa Turell (Eds.), *Dimensions of Forensic Linguistics* (pp. 249–264). https://doi.org/10.1075/aals.5.17egg

[113] Sarzynska-Wawer, J., Pawlak, A., Szymanowska, J., Hanusz, K., & Wawer, A. (2023). Truth or lie: Exploring the language of deception. *PLoS ONE*, *18*(2), e0281179. https://doi.org/10.1371/journal.pone.0281179

[114] Frant, A. E. (2020). Detecting Lies by Using Natural Language Processing. *Analele Stiintifice Ale Universitatii Alexandru Ioan Cuza Din Iasi Stiinte Juridice*, *66*, 293–302.

[115] Fabien, M., Villatoro-Tello, E., Motlicek, P., & Parida, S. (2020, December 1). *BertAA : BERT fine-tuning for Authorship Attribution*. ACL Anthology. Retrieved November 7, 2024, from https://aclanthology.org/2020.icon-main.16

[116] Enos, F., Shriberg, E., Graciarena, M., Hirschberg, J., & Stolcke, A. (2007). Detecting deception using critical segments. *Interspeech 2022*. https://doi.org/10.21437/interspeech.2007-619

[117] Sousa- Silva, R. (2021). Plagiarism: Evidence- based detection and analysis in forensic contexts. In M. Coulthard, A. May, & R. Sousa-Silva (Eds.), *The Routledge Handbook of Forensic Linguistics* (2nd ed., pp. 576–592). Routledge. https://doi.org/10.4324/9780203855607

[118] Uzun, L. (2023). ChatGPT and academic integrity concerns: Detecting artificial intelligence generated content. *Language Education and Technology*, *3*(1), 45–54.

[119] EYONG, E. I. (2022). Grammarly Plagiarism Checker in Promoting Quality. *Journal of Research in Multidisciplinary Methods and Applications*, *1*(4), 01220104003.

infringement analysis in legal and educational contexts include the PAN Plagiarism Corpus for real-world plagiarism cases (Asghari et al., 2015)[120], "The OSP has used a large amount of computing and human resources to collect" (Sekiya et al., 2022)[121], and Text Alignment Data from PAN competitions (Vani & Gupta, 2017)[122].

6. Disputed Texts

Contested texts frequently provide difficulties in legal contexts. Stratman (2015)[123] conducted a test "to reduce the possibility of expert hindsight bias contaminating the court's view regarding how the disputed text would be interpreted" (Stratman, 2015, p. 227)[124]. Ambiguities or inconsistencies in language may result in disputes among parties. The linguistic analysis seeks to elucidate meanings and evaluate the intent behind the employed language.

There are several methods for identifying copyright violations and plagiarism. "Various scientific journals employ different plagiarism checking systems, for example, iThenticate, eTBLAST, SPlaT, CrossCheck, Turnitin, andWriteCheck"(Polyanin & Shingareva, 2021, p. 4)[125]. AntConc looks at text patterns (Boyd et al., 2022)[126]. While Lexalytics does thorough text analytics (Halper et al., 2013)[127], snippets highlight code plagiarism. Important datasets such the Gutenberg

[120] Asghari, H., Khoshnava, K., Fatemi, O., & Faili, H. (2015). Developing Bilingual Plagiarism Detection Corpus using Sentence Aligned Parallel Corpus Notebook for PAN at CLEF 2015. *CLEF (Working Notes)*. http://ceur-ws.org/Vol-1391/148-CR.pdf

[121] Sekiya, T., Matsuda, Y., & Yamaguchi, K. (2022). Improvements of a hybrid syllabus search tool by syllabus-related heuristics. *2021 IEEE Frontiers in Education Conference (FIE)*, 1–8. https://doi.org/10.1109/fie56618.2022.9962619

[122] VANI, K., & GUPTA, D. (2017). IDENTIFYING DOCUMENT-LEVEL TEXT PLAGIARISM: A TWO-PHASE APPROACH. *Journal of Engineering Science and Technology*, *12*(12), 3226–3250.

[123] Stratman, J. (2016). *A forensic linguistic approach to legal disclosures*. Routledge. https://doi.org/10.4324/9781315687421

[124] Stratman, J. (2016). *A forensic linguistic approach to legal disclosures*. Routledge. https://doi.org/10.4324/9781315687421

[125] Polyanin, A. D., & Shingareva, I. K. (2021). The similarity index of scientific publications with equations and formulas, identification of self-plagiarism, and testing of the iThenticate system. *arXiv*, arXiv:2201.09062.

[126] Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin*. https://www.liwc.app

[127] Halper, F., Kaufman, M., & Kirsh, D. (2013). Text Analytics: The Hurwitz Victory Index Report. *Hurwitz and Associates*, 1–21.

Project for literary analysis (Jalobeanu, 2014)[128] and PAN Text Alignment for training on pairs of plagiarized texts (Asghari et al., 2015)[129] help to enable comprehensive forensic examination.

7. Cybercrime and Digital Linguistic Evidence

The emergence of digital communication has presented novel issues for forensic linguistics, especially concerning cybercrime. Examining linguistic data from digital platforms can yield significant insights into criminal behavior and intentions (Williams et al., 2017)[130].

In forensic linguistics, FTK and Magnet AXIOM are crucial for analyzing language in digital devices (Stanković & Khan, 2023)[131]. "X1 Social Discovery target digital evidence examination on PC based systems" (Charalambous et al., 2015, p. 4)[132] can retrieve social media data. On the other hand, Maltego delineates communication patterns (Marx, 2014)[133]. Essential datasets comprise the Enron Email Dataset for the analysis of corporate language (Chapanond et al., 2005)[134], the British National Corpus as a reference for linguistic comparison (Aston & Burnard, 1998)[135], and SCOWL for the identification of distinctive vocabulary (Kim & Kim, 2015)[136], facilitating digital forensic inquiries.

8. Hate Speech and Defamation

---

[128] Jalobeanu, M. S. (2014). A 43 YEARS HISTORY, PASSING FROM THE GUTENBERG PROJECT INITIATIVE TO THE OPEN EDUCATIONAL RESOURCES MOVEMENT . *eLearning and Software for Education*. https://doi.org/10.12753/2066-026x-14-298

[129] Asghari, H., Khoshnava, K., Fatemi, O., & Faili, H. (2015). Developing Bilingual Plagiarism Detection Corpus using Sentence Aligned Parallel Corpus Notebook for PAN at CLEF 2015. *CLEF (Working Notes)*. http://ceur-ws.org/Vol-1391/148-CR.pdf

[130] Williams, K., Facciola, J. M., McCann, P., & Catanzaro, V. M. (2017). The Legal Technology Guidebook. In *Springer eBooks*. https://doi.org/10.1007/978-3-319-54523-3

[131] Stanković, M., & Khan, T. M. (2023). Check for Digital Forensics Tool Evaluation on Deleted Files. In *Digital Forensics and Cyber Crime: 13th EAI International Conference, ICDF2C 2022, Boston, MA, November 16-18, 2022* (Vol. 508). Springer Nature.

[132] Charalambous, E., Bratskas, R., Karkas, G., Anastasiades, A., & Koutras, N. (2015). An innovative Digital Forensic Tool assisting evidence analysis in Cyprus. In *Big Data, Knowledge and Control Systems Engineering*.

[133] Marx, M. (2024). *The extension and customization of maltego data mining environment into anti-phishing system* [B.Sc. Thesis]. Rhodes University.

[134] Chapanond, A., Krishnamoorthy, M. S., & Yener, B. (2005). Graph Theoretic and spectral analysis of Enron email data. *Computational and Mathematical Organization Theory*, *11*(3), 265–281. https://doi.org/10.1007/s10588-005-5381-4

[135] Aston, G., & Burnard, L. (1998). *The BNC Handbook: Exploring the British National Corpus with SARA*. http://corpus.leeds.ac.uk/teaching/aston-burnard-bnc.pdf

[136] Kim, S., & Kim, D. (2015). Automatic identifier inconsistency detection using code dictionary. *Empirical Software Engineering*, *21*(2), 565–604. https://doi.org/10.1007/s10664-015-9369-5

The linguistic analysis of hate speech and defamation plays a crucial role in addressing issues of discrimination and harm. Certain nations possess particular laws categorizing online hate speech as a criminal offence, while others have opted to establish civil redress mechanisms or contend that free speech rights should remain unimpeded on the Internet (Jakubowicz, 2006)[137]. Defamation rules stipulate that if false information is published, either in writing or orally, presented as fact rather than opinion, the author may face prosecution for defamation of character. (W. Shuy, 2015)[138] Understanding the language used in hateful or defamatory statements is essential for legal proceedings that seek to hold individuals accountable for their actions.

Applications for hate speech analysis in forensic linguistics, Hatebase provides a database of hate terms (Boyd, 2022)[139], while Perspective API scores text for toxicity and hate language (Lees et al., 2022)[140]. Key datasets include the Hate Speech and Offensive Language (Davidson et al., 2017)[141] for labeled hate content, the Twitter Hate Speech Dataset (Mulki et al., 2019)[142] for online hate language, are valuable analysis tools in comprehensive hate speech analysis.

9. Forensic Sociolinguistics

Forensic sociolinguistics investigates the intersection of language and social variables in legal settings (Friedland, 1989)[143]. Comprehending the effects of social elements on linguistic evidence enable the "judgments of witness credibility were made as a function of social psychological judgments about the witness" (Levi, 1990, p. 19)[144].

---

[137] Jakubowicz, A. (2006). Media and marginalized groups. In *Elsevier eBooks* (pp. 602–607). https://doi.org/10.1016/b0-08-044854-2/00708-2

[138] W. SHUY, R. (2015). Discourse analysis in the legal context. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), *The Handbook of Discourse Analysis* (pp. 437–452). John Wiley & Sons.

[139] Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin*. https://www.liwc.app

[140] Lees, A., Tran, V. Q., Tay, Y., Sorensen, J., Gupta, J., Metzler, D., & Vasserman, L. (2022). A new generation of perspective API: efficient multilingual character-level transformers. *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/3534678.3539147

[141] Davidson, T., Warmsley, D., Macy, M., & Weber, I. (2017). Automated hate speech detection and the problem of offensive language. *Proceedings of the International AAAI Conference on Web and Social Media*, *11*(1), 512–515. https://doi.org/10.1609/icwsm.v11i1.14955

[142] Mulki, H., Haddad, H., Ali, C. B., & Alshabani, H. (2019). L-HSAB: A Levantine Twitter Dataset for Hate Speech and Abusive Language. In *Proceedings of the Third Workshop on Abusive Language Online*. Association for Computational Linguistics. https://doi.org/10.18653/v1/w19-3512

[143] Friedland, S. I. (1989). On common sense and the evaluation of witness credibility. *Case Western Reserve Law Review*, *40*(1), 165. https://scholarlycommons.law.case.edu/caselrev/vol40/iss1/5/

[144] Levi, J. N. (1990). The study of language in the Judicial Process. In J. N. Levi & A. G. Walker (Eds.), *Language in the Judicial Process. Law, Society and Policy* (Vol. 5, pp. 3–35). Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-3719-3_1

The advancements in technology are enhancing sociolinguistic analysis in forensic linguistics. Atlas.ti and NVivo facilitate the coding and examination of social language patterns (Lewis, 2004)[145]. COCA provides comprehensive statistics on language usage, facilitating the examination of demographic-specific linguistic changes (Mubin et al., 2024)[146]. TalkBank compiles spoken language within legal and social frameworks, facilitating the examination of courtroom interactions and sociolinguistic characterization (MacWhinney, 2001)[147].

10. Threats in Digital Communication

The examination of internet dangers has become progressively vital as digital communication evolves. The array of issues, challenges, and threats linked to the digitalization of modern society and the integration of artificial intelligence technologies and neural networks into present socio-political interactions is considerable (Volodenkov & Fedorchenko, 2021)[148].

Various methodologies can be conducted to perform digital threat analysis. Perspective API scores text for toxicity and hate language (Lees et al., 2022). It can detect threats and violence. The Cogito Intelligence API identifies attitudes and hazards in communication (Zucco et al., 2021)[149]. Guardian AI surveils social media for hazardous words (Nakai & Oroy, 2024)[150]. OLID was thoroughly tested and evaluated by Rosenthal et al. "The taxonomy proposed in OLID makes it possible to represent different kinds of offensive content (Rosenthal et al., 2020, p. 915)[151]. On the other hand, TRAC concentrates on detecting threatening language patterns on social media platforms (Kumar et al., 2021)[152].

---

[145] Lewis, R. B. (2004). NVivo 2.0 and ATLAS. ti 5.0: A comparative review of two popular qualitative data-analysis programs. *Field Methods*, *16*(4), 439–464.

[146] Mubin, R., Ozoda, M., Mashhura, R., & Ozoda, H. (2024). THE USAGE OF COCA AND INVESTIGATIONS OF VOCABULARY FORMING. *Ta'lim Innovatsiyasi Va Integratsiyasi*, *21*(3), 88–92.

[147] MacWhinney, B. (2001). New developments in CHILDES. In A. Do, L. Domínguez, & A. Johansen (Eds.), *the 25th annual Boston University Conference on Language Development* (pp. 458–468).

[148] Volodenkov, S., & Fedorchenko, S. (2021). Subjectness of digital communication in the context of the technological evolution of the contemporary society: threats, challenges, and risks. *Przegląd Strategiczny*, *14*, 437–456. https://doi.org/10.14746/ps.2021.1.25

[149] Zucco, C., Calabrese, B., Agapito, G., Guzzi, P. H., & Cannataro, M. (2019). Sentiment analysis for mining texts and social networks data: Methods and tools. *Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery*, *10*(1). https://doi.org/10.1002/widm.1333

[150] Nakai, T., & Oroy, K. (2024). The Digital Guardian: Leveraging AI to Combat Cyber Threats. *EasyChair Preprint*, 13330.

[151] Rosenthal, S., Atanasova, P., Karadzhov, G., Zampieri, M., & Nakov, P. (2021). SOLID: A Large-Scale Semi-Supervised Dataset for Offensive Language Identification. *arXiv*, arXiv:2004.14454. https://doi.org/10.18653/v1/2021.findings-acl.80

[152] Kumar, R., Lahiri, B., & Ojha, A. K. (2021). Aggressive and Offensive language identification in Hindi, Bangla, and English: A Comparative study. *SN Computer Science*, *2*(1). https://doi.org/10.1007/s42979-020-00414-6

3. Conclusion

In conclusion, forensic linguistics serves as an interdisciplinary field that emphasizes the importance of language in the legal system. This study has highlighted the important function of language in shaping legal results by exploring statutory language, language employed in forensic and judicial proceedings, and the evaluation of linguistic evidence.

Forensic language analysis has consequences outside legal interpretation. They shape court system behavior. They also increase the effectiveness of decision-making and communication. The continuous development of technology will improve the integration of complex linguistic analytic techniques.

Moreover, the complexity shown by a globally integrated and varied world emphasizes the need for continuous study in forensic linguistics. Forensic linguistics will become much more important as societies negotiate bias, cybercrime, and the complexities of digital communication.

Future research should focus on the growing role of forensic linguistics in addressing contemporary challenges, particularly the increasing frequency of digital communication and the impact of social media on language use in court settings. Scholars and practitioners can improve the area of forensic linguistics by constantly investigating the junction of language and law, therefore preserving its relevance in the search for justice.

4. References

Ahlbrand, A. A. (2020). Analyzing Analytics: Litigation Analytics in Bloomberg Law, Westlaw Edge, and Lexis Advance. *42 the CRIV Sheet 9 (Feb. 2020)*, *2*. https://www.repository.law.indiana.edu/cgi/viewcontent.cgi?article=3954&context=facpub

Ali, S. A., & Algane, M. (2013). The Role of Forensic Translation in Courtrooms Contexts. *Arab World English Journal*, *4*.

Angermeyer, P. S. (2008). Creating monolingualism in the multilingual courtroom. *Sociolinguistic Studies*, *2*(3), 385–404. https://doi.org/10.1558/sols.v2i3.385

Apturkar, A., Iliev, A. I., Anand, A., Oli, A., Siddenki, S. R., & Meka, V. R. (2020). Sentiment Analysis of Speech with Application to Various Languages. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, *10*, 103–118. https://doi.org/10.55630/dipp.2020.10.6

Asghari, H., Khoshnava, K., Fatemi, O., & Faili, H. (2015). Developing Bilingual Plagiarism Detection Corpus using Sentence Aligned Parallel Corpus Notebook for PAN at CLEF 2015. *CLEF (Working Notes)*. http://ceur-ws.org/Vol-1391/148-CR.pdf

Aston, G., & Burnard, L. (1998). *The BNC Handbook: Exploring the British National Corpus with SARA*. http://corpus.leeds.ac.uk/teaching/aston-burnard-bnc.pdf

B Pieper, D. (1998). The Appropriate Judicial Actor for Patent Interpretation: A Commentary on the Supreme Court's Decision in Markman v. Westview Instruments, Inc. *51 Ark. L. Rev.*, 159.

Bilić, M., & Gaspar, A. (2018). Extraction of Phrasal Verbs from the Comparable English Corpus of Legal Texts. *International Journal of English Language & Translation Studies*, *6*(2), 184–194.

Bommarito, M. J., II, Katz, D. M., & Detterman, E. M. (2021). LexNLP: Natural language processing and information extraction for legal and regulatory texts. In *Edward Elgar Publishing eBooks*. https://doi.org/10.4337/9781788972826.00017

Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin*. https://www.liwc.app

Cao, D. (2010). Legal translation: Translating legal language. In M. Coulthard & A. Johnson (Eds.), *The Routledge Handbook of Forensic Linguistics* (1st ed., pp. 78–91). Routledge.

Carias, J. M. A. (2023). Desarrollo de aplicación móvil y desarrollo de aplicación de comercio electrónico. *CRAI*, https://repositorio.unitec.edu/xmlui/handle/123456789/10422.

Chaemsaithong, K. (2021). Advances in studies of the historical courtroom: (Con)Textual, ideational and interpersonal dimensions. In M. Coulthard, A. May, & R. Sousa-Silva (Eds.), *The Routledge Handbook of Forensic Linguistics* (2nd ed., pp. 211–227). Routledge.

Chapanond, A., Krishnamoorthy, M. S., & Yener, B. (2005). Graph Theoretic and spectral analysis of Enron email data. *Computational and Mathematical Organization Theory*, *11*(3), 265–281. https://doi.org/10.1007/s10588-005-5381-4

Charalambous, E., Bratskas, R., Karkas, G., Anastasiades, A., & Koutras, N. (2015). An innovative Digital Forensic Tool assisting evidence analysis in Cyprus. In *Big Data, Knowledge and Control Systems Engineering*.

Clarke, C., & Milne, R. J. (2001). *National evaluation of the PEACE investigative interviewing course*. Interdepartmental and Cross-Faculty Research Groups and Centres (until 2017), Services and Administration (until 2017). Retrieved October 8, 2024, from https://eprints.kingston.ac.uk/31400/

Coulthard, M., & Johnson, A. (2010). *The Routledge Handbook of Forensic Linguistics*. Routledge.

Crystal, D. (2008). *A Dictionary of Linguistics and Phonetics* (6th ed.). John Wiley & Sons. https://doi.org/10.1002/9781444302776

Davidson, T., Warmsley, D., Macy, M., & Weber, I. (2017). Automated hate speech detection and the problem of offensive language. *Proceedings of the International AAAI Conference on Web and Social Media*, *11*(1), 512–515. https://doi.org/10.1609/icwsm.v11i1.14955

Derin, T., Evizareza, Deliani, S., & Hamuddin, B. (2019). *Exploring the Past, Present, and Future of Forensic Linguistics Study: A Brief Overview*. The First Conference of Indonesian Community for Forensic Linguistics (KLFI-1), Pekanbaru, Riau, Indonesia.

Dew, R. (2023). *The empathetic algorithm: Leveraging AI for Next-Level CX* (1st ed.). CapFeather Global.

Dumas, B. K. (2000). U.S. Pattern Jury Instructions: Problems and proposals. *International Journal of Speech, Language and the Law*, *7*(1), 49–71.

Eggington, W. (2008). Deception and fraud. In J. Gibbons & M. Teresa Turell (Eds.), *Dimensions of Forensic Linguistics* (pp. 249–264). https://doi.org/10.1075/aals.5.17egg

Enos, F., Shriberg, E., Graciarena, M., Hirschberg, J., & Stolcke, A. (2007). Detecting deception using critical segments. *Interspeech 2022*. https://doi.org/10.21437/interspeech.2007-619

EYONG, E. I. (2022). Grammarly Plagiarism Checker in Promoting Quality. *Journal of Research in Multidisciplinary Methods and Applications*, *1*(4), 01220104003.

Fabien, M., Villatoro-Tello, E., Motlicek, P., & Parida, S. (2020, December 1). *BertAA : BERT fine-tuning for Authorship Attribution*. ACL Anthology. Retrieved November 7, 2024, from https://aclanthology.org/2020.icon-main.16

Fantinuoli, C., Marchesini, G., Landan, D., & Horak, L. (2022). KUDO Interpreter Assist: Automated real-time support for remote interpretation. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2201.01800

Fowler, Y., Vaughan, M., & Wheatcroft, J. (2016). The Interpreter-Mediated Police Interview. In G. Oxburgh, T. Myklebust, T. Grant, & R. J. Milne (Eds.), *Communication in investigative and legal contexts* (1st ed., pp. 315–333). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118769133

Frant, A. E. (2020). Detecting Lies by Using Natural Language Processing. *Analele Stiintifice Ale Universitatii Alexandru Ioan Cuza Din Iasi Stiinte Juridice*, *66*, 293–302.

Friedland, S. I. (1989). On common sense and the evaluation of witness credibility. *Case Western Reserve Law Review*, *40*(1), 165. https://scholarlycommons.law.case.edu/caselrev/vol40/iss1/5/

Gabbert, F., Hope, L., Carter, E., Boon, R., & Fisher, R. (2015). The role of initial witness accounts within the investigative process. *Communication in Investigative and Legal Contexts: Integrated Approaches From Forensic Psychology, Linguistics and Law Enforcement*, 107–131.

Gao, J. (2010). Review and prospects of the research of forensic linguistics in China. *Asian Social Science*, *6*(10). https://doi.org/10.5539/ass.v6n10p127

Gibbons, J. (2003). *Forensic Linguistics: An Introduction to language in the justice system*. Blackwell. https://ci.nii.ac.jp/ncid/BA60279023

Graham, C., & Roll, N. (2024). Evaluating OpenAI's Whisper ASR: Performance analysis across diverse accents and speaker traits. *JASA Express Letters*, *4*(2). https://doi.org/10.1121/10.0024876

Gries, S. Th. (2021). Corpora and legal interpretation: Corpus approaches to ordinary meaning in legal interpretation. In M. Coulthard, A. May, & R. Sousa-Silva (Eds.), *The Routledge Handbook of Forensic Linguistics* (2nd ed., pp. 628–643). Routledge.

Gupta, S., Patra, T. K., & Chaudhuri, C. (2022). Role of Machine Learning in Authorship Attribution with Select Stylometric Features. In *Lecture notes in networks and systems* (pp. 920–932). https://doi.org/10.1007/978-3-030-96308-8_86

Hall, M. E. K., & Windett, J. H. (2013). New data on state supreme court cases. *State Politics & Policy Quarterly*, *13*(4), 427–445. https://doi.org/10.1177/1532440013497971

Halper, F., Kaufman, M., & Kirsh, D. (2013). Text Analytics: The Hurwitz Victory Index Report. *Hurwitz and Associates*, 1–21.

Hasan, I., Rizvi, S., Jain, S., & Huria, S. (2021). The AI enabled Chatbot Framework for Intelligent Citizen-Government Interaction for Delivery of Services. *International Conference on Computing for Sustainable Global Development*, 601–606. https://doi.org/10.1109/indiacom51348.2021.00106

Heffer, C. (2008). The language and communication of jury instruction. In J. Gibbons & M. Teresa Turell (Eds.), *Dimensions of Forensic Linguistics* (pp. 47–65). https://doi.org/10.1075/aals.5.05hef

Hilton, M. L., & Holmes, D. I. (1993). An assessment of cumulative sum charts for authorship attribution. *Literary and Linguistic Computing*, *8*(2), 73–80. https://doi.org/10.1093/llc/8.2.73

Hiltunen, R. (2012). Th e Grammar and Structure of Legal Texts. In P. M. Tiersma & L. Solan (Eds.), *The Oxford Handbook of Language and Law* (pp. 39–51). Oxford University Press.

InfoToHow. (2024, January 22). *Essential Elements of a Mortgage Deed: Unveiling Its Components and Protections*. Info to How. Retrieved October 14, 2024, from https://infotohow.com/finance/essential-elements-of-a-mortgage-deed/

Iu, K. Y., Tu, L., & Liu, Y. (2024). AI-Powered Legal Translation: Can ChatGPT Facilitate the Development of Chinese Common Law? *SSRN*. https://doi.org/10.2139/ssrn.4920531

Jakubowicz, A. (2006). Media and marginalized groups. In *Elsevier eBooks* (pp. 602–607). https://doi.org/10.1016/b0-08-044854-2/00708-2

Jalobeanu, M. S. (2014). A 43 YEARS HISTORY, PASSING FROM THE GUTENBERG PROJECT INITIATIVE TO THE OPEN EDUCATIONAL RESOURCES MOVEMENT . *eLearning and Software for Education*. https://doi.org/10.12753/2066-026x-14-298

Jessen, M. (2018). Forensic voice comparison. In *Handbook of Communication in the Legal Sphere* (pp. 219–255). De Gruyter Mouton. https://doi.org/10.1515/9781614514664-012

Juola, P. (2009). JGAAP: A System for Comparative Evaluation of Authorship Attribution. *Journal of the Chicago Colloquium on Digital Humanities and Computer Science*, *1*(1). https://doi.org/10.6082/m1n29v4z

Juola, P., & Vescovi, D. (2010). Empirical evaluation of authorship obfuscation using JGAAP. In *Proceedings of the 3rd ACM workshop on Artificial Intelligence and Security* (Vol. 8, pp. 14–18). https://doi.org/10.1145/1866423.1866427

Ke, L., Xiao, P., Chen, X., Yu, S., Chen, X., & Wang, H. (2023). A novel cross-domain adaptation framework for unsupervised criminal jargon detection via pre-trained contextual embedding of darknet corpus. *Expert Systems With Applications*, *242*, 122715. https://doi.org/10.1016/j.eswa.2023.122715

Khoyi, A. M., & Behnam, B. (2014). Legal Discourse: Analysis of education and criminal convictions in Iranian courts. *Asian Journal of Education and e-Learning*, *2*(6). http://ajouronline.com/index.php?journal=AJEEL&page=article&op=view&path%5B%5D=1859&path%5B%5D=1031

Kim, S., & Kim, D. (2015). Automatic identifier inconsistency detection using code dictionary. *Empirical Software Engineering*, *21*(2), 565–604. https://doi.org/10.1007/s10664-015-9369-5

KISBYE, E. (2024). Man vs. Machine: AI's Impact on Intellectual Property and Copyright Law. *TRINITY COLLEGE UNDERGRADUATE LAW JOURNAL*, *1*.

Kreuz, R. (2023). *Linguistic fingerprints: How Language Creates and Reveals Identity*. Rowman & Littlefield.

Kumar, D., Patil, P. K. V., Agarwal, A., & Prasanna, S. R. M. (2022). Fake speech detection using OpenSMILE features. In *Lecture notes in computer science* (pp. 404–415). https://doi.org/10.1007/978-3-031-20980-2_35

Kumar, R., Lahiri, B., & Ojha, A. K. (2021). Aggressive and Offensive language identification in Hindi, Bangla, and English: A Comparative study. *SN Computer Science*, *2*(1). https://doi.org/10.1007/s42979-020-00414-6

Kwarteng, J., Perfumi, S. C., Farrell, T., Third, A., & Fernandez, M. (2022). Misogynoir: challenges in detecting intersectional hate. *Social Network Analysis and Mining*, *12*(1). https://doi.org/10.1007/s13278-022-00993-7

Lauritsen, M. (2021). Document Automation. In D. M. Katz, R. Dolin, & M. J. Bommarito (Eds.), *Legal Informatics*. Cambridge University Press.

Lees, A., Tran, V. Q., Tay, Y., Sorensen, J., Gupta, J., Metzler, D., & Vasserman, L. (2022). A new generation of perspective API: efficient multilingual character-level transformers. *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/3534678.3539147

Leung, E. S. M. (2008). Interpreting for the minority, interpreting for the power. In J. Gibbons & M. Teresa Turell (Eds.), *Dimensions of Forensic Linguistics* (pp. 197–211). John Benjamins B.V.

Levi, J. N. (1990). The study of language in the Judicial Process. In J. N. Levi & A. G. Walker (Eds.), *Language in the Judicial Process. Law, Society and Policy* (Vol. 5, pp. 3–35). Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-3719-3_1

Lewis, R. B. (2004). NVivo 2.0 and ATLAS. ti 5.0: A comparative review of two popular qualitative data-analysis programs. *Field Methods*, *16*(4), 439–464.

Li, X., & Li, J. (2021). Law - net: A New Method for Legal Text Mining. *Scientific Journal of Technology*, *3*(7). https://api.semanticscholar.org/CorpusID:251244662

Lo Bosco, G., Pilato, G., & Schicchi, D. (2021). DeepEva: A deep neural network architecture for assessing sentence complexity in Italian and English languages. *Array*, *12*, 100097. https://doi.org/10.1016/j.array.2021.100097

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, arXiv:1705.07874.

MacLeod, N. (2021). Assuming identities online: Authorship synthesis in undercover investigations. In M. Coulthard, A. May, & R. Sousa-Silva (Eds.), *The Routledge*

Handbook of Forensic Linguistics (2nd ed., pp. 159–173). Routledge. http://nrl.northumbria.ac.uk/41825/

MacNeal, A., Fiallo, K., Jones, A., Jones, S., Laureano, S., Monjarrez, M., & Xu, Y. (2019). Sounding black: the legal implications of linguistic profiling. *Northeastern University Working Papers in Linguistics*, *4*.

MacWhinney, B. (2001). New developments in CHILDES. In A. Do, L. Domínguez, & A. Johansen (Eds.), *the 25th annual Boston University Conference on Language Development* (pp. 458–468).

Magdin, M., Sulka, T., Tomanová, J., & Vozár, M. (2019). Voice analysis using PRAAT software and classification of user emotional state. *International Journal of Interactive Multimedia and Artificial Intelligence*, *5*(6), 33. https://doi.org/10.9781/ijimai.2019.03.004

Marx, M. (2024). *The extension and customization of maltego data mining environment into anti-phishing system* [B.Sc. Thesis]. Rhodes University.

Mattila, H. (2006). Legal language: history. In *Elsevier eBooks* (pp. 8–13). https://doi.org/10.1016/b0-08-044854-2/04504-1

McMenamin, G. R. (2002). *Forensic Linguistics: Advances in forensic stylistics*. CRC Press. https://ci.nii.ac.jp/ncid/BA75691949

McMenamin, G. R. (2010). Forensic stylistics: Theory and practice of forensic stylistics. In M. Coulthard & A. Johnson (Eds.), *The Routledge Handbook of Forensic Linguistics* (1st ed., pp. 487–507). Routledge.

Meloy, J. R., & Hoffmann, J. (2014). *International Handbook of Threat Assessment*. Oxford University Press, USA.

Moe, M. K. (2023). *Post-processing automatic speech recognition transcriptions: A study for investigative interviews* [Master thesis]. NTNU.

Mubin, R., Ozoda, M., Mashhura, R., & Ozoda, H. (2024). THE USAGE OF COCA AND INVESTIGATIONS OF VOCABULARY FORMING. *Ta'lim Innovatsiyasi Va Integratsiyasi*, *21*(3), 88–92.

Mulki, H., Haddad, H., Ali, C. B., & Alshabani, H. (2019). L-HSAB: A Levantine Twitter Dataset for Hate Speech and Abusive Language. In *Proceedings of the Third Workshop on Abusive Language Online*. Association for Computational Linguistics. https://doi.org/10.18653/v1/w19-3512

Muravev, Y. (2020). Machine translation and legal tech in legal translation training. *Proceedings of the International Scientific Conference - Digital Transformation on Manufacturing, Infrastructure and Service*, 1–7. https://doi.org/10.1145/3446434.3446553

Nagrani, A., Chung, J. S., Xie, W., & Zisserman, A. (2019). Voxceleb: Large-scale speaker verification in the wild. *Computer Speech & Language*, *60*, 101027. https://doi.org/10.1016/j.csl.2019.101027

Nakai, T., & Oroy, K. (2024). The Digital Guardian: Leveraging AI to Combat Cyber Threats. *EasyChair Preprint*, 13330.

Núñez, G. G. (2017). Law and translation at the US–Mexico border: Translation policy in a diglossic setting. In G. G. Núñez & R. Meylaerts (Eds.), *Translation and Public Policy* (1st ed., pp. 152–170). Routledge.

Olsson, J. (2009). *WordCrime: Solving crime through forensic linguistics*. A&C Black. http://ci.nii.ac.jp/ncid/BA91502286

Perkins, R., & Grant, T. (2013). Forensic Linguistics. In J. A. Siegel, P. J. Sauko, & M. M. Houck (Eds.), *Encyclopedia of Forensic Science* (2nd ed., pp. 174–177). Elsevier Ltd.

Polyanin, A. D., & Shingareva, I. K. (2021). The similarity index of scientific publications with equations and formulas, identification of self-plagiarism, and testing of the iThenticate system. *arXiv*, arXiv:2201.09062.

Powell, M. B., Wright, R., & Hughes-Scholes, C. H. (2010). Contrasting the perceptions of child testimony experts, prosecutors and police officers regarding individual child abuse interviews. *Psychiatry Psychology and Law*, *18*(1), 33–43. https://doi.org/10.1080/13218710903566995

PUTRI, S. H. (2024). *An Evaluation of Usability Computer-Aided Translation Tools Among Bispro Students*. Jakarta State Polytechnic.

Qiao, F., Li, P., Zhang, X., Ding, Z., Cheng, J., & Wang, H. (2017). Predicting Social Unrest Events with Hidden Markov Models Using GDELT. *Discrete Dynamics in Nature and Society*, *2017*, 1–13. https://doi.org/10.1155/2017/8180272

Rosenthal, S., Atanasova, P., Karadzhov, G., Zampieri, M., & Nakov, P. (2021). SOLID: A Large-Scale Semi-Supervised Dataset for Offensive Language Identification. *arXiv*, arXiv:2004.14454. https://doi.org/10.18653/v1/2021.findings-acl.80

Russo, M., Bendazzoli, C., Sandrelli, A., & Spinolo, N. (2012). The European Parliament Interpreting Corpus (EPIC): implementation and developments. In *Peter Lang eBooks* (Vol. 147, pp. 53–90). https://cris.unibo.it/handle/11585/121334

Sarzynska-Wawer, J., Pawlak, A., Szymanowska, J., Hanusz, K., & Wawer, A. (2023). Truth or lie: Exploring the language of deception. *PLoS ONE*, *18*(2), e0281179. https://doi.org/10.1371/journal.pone.0281179

SAYWITZ, K. (2002). Developmental Underpinnings of Children's Testimony. In H. L. Westcott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's Testimony* (pp. 3–19). John Wiley & Sons, Ltd.

Schwab, D., Tchechmedjiev, A., Goulian, J., & Sérasset, G. (2014). Comparisons of relatedness measures through a word sense disambiguation task. *Text, Speech and Language Technology*, 221–243. https://doi.org/10.1007/978-3-319-08043-7_13

Sekiya, T., Matsuda, Y., & Yamaguchi, K. (2022). Improvements of a hybrid syllabus search tool by syllabus-related heuristics. *2021 IEEE Frontiers in Education Conference (FIE)*, 1–8. https://doi.org/10.1109/fie56618.2022.9962619

Silver, C., & Lewins, A. (2014). *Using Software in Qualitative Research: A Step-by-Step Guide*. https://doi.org/10.4135/9781473906907

Sobowale, J. (2016). How artificial intelligence is transforming the legal profession. *ABA Journal*, *1*(1).

Solan, L. (2006). Definition/Rules in legal language. In K. Brown (Ed.), *Encyclopedia of Language & Linguistics* (2nd ed., pp. 403–409). Elsevier. https://doi.org/10.1016/b0-08-044854-2/00692-1

Solan, L. M. (1993). *The language of judges*. https://doi.org/10.7208/chicago/9780226767895.001.0001

Sousa- Silva, R. (2021). Plagiarism: Evidence- based detection and analysis in forensic contexts. In M. Coulthard, A. May, & R. Sousa-Silva (Eds.), *The Routledge Handbook of Forensic Linguistics* (2nd ed., pp. 576–592). Routledge. https://doi.org/10.4324/9780203855607

Sprowl, J. A. (1981). WESTLAW vs LEXIS: computer-assisted legal research comes of age. *Program Electronic Library and Information Systems*, *15*(3), 132–141. https://doi.org/10.1108/eb046827

Stabler, R. (2014). "What We've Got Here is Failure to Communicate": The Plain Writing Act of 2010. *SSRN Electronic Journal*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2574207

Stanković, M., & Khan, T. M. (2023). Check for Digital Forensics Tool Evaluation on Deleted Files. In *Digital Forensics and Cyber Crime: 13th EAI International Conference, ICDF2C 2022, Boston, MA, November 16-18, 2022* (Vol. 508). Springer Nature.

Stratman, J. (2016). *A forensic linguistic approach to legal disclosures*. Routledge. https://doi.org/10.4324/9781315687421

Stygall, G. (2010). Legal writing: complexity Complex documents/average and not-so-average readers. In M. Coulthard, A. May, & R. Sousa-Silva (Eds.), *The Routledge Handbook of Forensic Linguistics* (2nd ed., pp. 32–47). Routledge. https://doi.org/10.4324/9780203855607-14

Svartvik, J. L. (1968). The Evans statements : a case for forensic linguistics. In *Acta Universitatis Gothoburgensis eBooks*. https://ci.nii.ac.jp/ncid/BA11447154

Tallentire, G. (2009). The Hong Kong (police) magistrate in the 1880s and 1990s. In G. Bickley (Ed.), *A Magistrate's Court in Nineteenth Century Hong Kong: Court in Time* (pp. 133–143).

Tiersma, P. (2001). The Rocky Road to Legal Reform: Improving the language of jury instructions. *Brooklyn Law Review*, *66*(4), 1081. https://brooklynworks.brooklaw.edu/cgi/viewcontent.cgi?article=1709&context=blr

Tiersma, P. (2008). The nature of legal language. In J. Gibbons & M. Teresa Turell (Eds.), *Dimensions of Forensic Linguistics* (pp. 7–25). John Benjamins B.V.

Tiersma, P. M. (1999). *Legal language*. University of Chicago Press.

Tiersma, P. M. (2010). *Parchment, paper, pixels: Law and the Technologies of Communication*. University of Chicago Press.

Timaná, L. C. R., Lozano, D. F. S., & García, J. F. C. (2020). Software to determine the readability of written documents by implementing a variation of the gunning fog index using the Google Linguistic Corpus. In *Communications in computer and information science* (pp. 409–420). https://doi.org/10.1007/978-3-030-42517-3_31

Umiyati, M. (2020). A Literature review of Forensic Linguistics. *International Journal of Forensic Linguistics*, *1*(1), 23–29.

Uzun, L. (2023). ChatGPT and academic integrity concerns: Detecting artificial intelligence generated content. *Language Education and Technology*, *3*(1), 45–54.

VanBuskirk, K., & Letson, M. R. (2020). An IBM Watson tone analysis of selected judicial decisions. *The Scribes Journal of Legal Writing*, *19*, 25–48.

VANI, K., & GUPTA, D. (2017). IDENTIFYING DOCUMENT-LEVEL TEXT PLAGIARISM: A TWO-PHASE APPROACH. *Journal of Engineering Science and Technology*, *12*(12), 3226–3250.

Volodenkov, S., & Fedorchenko, S. (2021). Subjectness of digital communication in the context of the technological evolution of the contemporary society: threats, challenges, and risks. *Przegląd Strategiczny*, *14*, 437–456. https://doi.org/10.14746/ps.2021.1.25

W. SHUY, R. (2015). Discourse analysis in the legal context. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), *The Handbook of Discourse Analysis* (pp. 437–452). John Wiley & Sons.

Whalen, J. (1995). Expert systems versus systems for experts: computer-aided dispatch as a support system in real-world environments. In *Cambridge University Press eBooks* (pp. 161–183). https://dl.acm.org/citation.cfm?id=214811.214832

Williams, K., Facciola, J. M., McCann, P., & Catanzaro, V. M. (2017). The Legal Technology Guidebook. In *Springer eBooks*. https://doi.org/10.1007/978-3-319-54523-3

Yang, X., & Yu, X. (2019). *Identifying Patent Risks in Technological Competition: A Patent Analysis of Artificial Intelligence Industry*. 8th International Conference on Industrial Technology and Management (ICITM), Cambridge, United Kingdom of Great Britain and Northern Ireland. https://doi.org/10.1109/icitm.2019.8710719

Yuan, J., & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *The Journal of the Acoustical Society of America*, *123*(5_Supplement), 3878. https://doi.org/10.1121/1.2935783

Zaiets, O. (2015). Application Software IBM I2 Analyst's Notebook in Law Enforcement in Ukraine for Pretrial Investigation of Criminal Offenses. *European Reforms Bulletin*, *1*, 69–73.

Zavala-Rojas, D., Sorato, D., Hareide, L., & Hofland, K. (2022). The Multilingual Corpus of Survey Questionnaires: A tool for refining survey translation. *Meta Journal Des Traducteurs*, *67*(1), 71–93. https://doi.org/10.7202/1092191ar

Zucco, C., Calabrese, B., Agapito, G., Guzzi, P. H., & Cannataro, M. (2019). Sentiment analysis for mining texts and social networks data: Methods and tools. *Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery*, *10*(1). https://doi.org/10.1002/widm.1333