IJFL (International Journal of Forensic Linguistic)
Vol. 5, No. 2 October 2025, Page 43-55

P-ISSN: 2723-1542
E-ISSN: 2723-5254
Available Online at https://pusatpublikasi.com/index.php/ijfl

Criminal Investigation Techniques: The Role of Forensic Linguistics in Digital Evidence

Yarman Zalukhu

Master of Linguistics, Universitas Warmadewa, Denpasar, Indonesia
yarmanzalukhu@gmail.com

How to cite (in APA style):

Zalukhu, Yarman. (2025). Criminal Investigation Techniques: The Role of Forensic Linguistics in Digital Evidence. *IJFL (International Journal of Forensic Linguistic)*, 5(2), 43-55.

Abstract-The rapid advancement of digital technology presents both opportunities and challenges for law enforcement, particularly through the increasing use of digital evidence in criminal investigations. The complexity of digital data, ranging from text messages and emails to social media posts and encrypted communications, requires adaptive and multidisciplinary analytical methods. Forensic linguistics has emerged as an interdisciplinary field that bridges language studies and law, offering significant contributions in author identification, communication pattern analysis, lie detection, and the interpretation of implicit meanings in digital texts. This study aims to provide a comprehensive synthesis of the role of forensic linguistics in managing digital evidence through a Systematic Literature Review (SLR) approach. Relevant literature was collected from reputable academic databases using targeted keywords and selected based on inclusion and exclusion criteria. The data were analyzed thematically to identify methodological trends, major findings, and ongoing challenges. The results reveal two dominant approaches in digital forensic linguistics: a qualitative approach grounded in discourse analysis, and a quantitative approach utilizing computational and artificial intelligence techniques that demonstrate high analytical precision. Case studies indicate that stylistic markers, persuasive strategies, and microlinguistic features assist in reconstructing criminal behavior patterns, while stylometry, n-grams, and machine learning algorithms achieve over 90% accuracy in authorship attribution and criminal vocabulary detection. This synthesis suggests that the strength of digital forensic linguistics lies in integrating qualitative interpretation with computational modeling. Developing hybrid frameworks that combine linguistic insight and algorithmic precision is essential for improving investigative effectiveness, evidentiary validity, and methodological consistency in digital-era criminal investigations.

Keywords: Forensic Linguistics, Digital Evidence, Criminal Investigation, Authorship Attribution, Language Analysis.

I. INTRODUCTION

The digital revolution that has taken place in the last two decades has fundamentally changed the way humans communicate, work, and interact. This transformation has a positive impact on improving global connectivity, accelerating information distribution, and bringing efficiency in various aspects of life. However, on the other hand, the development of digital technology also poses serious challenges in the field of security and law enforcement. One of the most crucial aspects is the increasing use of digital evidence in criminal investigations. Digital evidence is no longer limited to simple electronic documents, but extends to include instant messaging, emails, voice recordings, videos, social media posts, online forums, and even conversations through encryption-protected applications. The complexity of this form and diversity of media demands a more adaptive, multidisciplinary, and technology-based approach to investigation.

In the context of modern criminal law, digital evidence has a strategic position as one of the main evidentiary instruments in court. However, the unique characteristics of digital proof make them vulnerable to manipulation, forgery, or erasure. The fact that digital data can be replicated without a clear trace raises a debate regarding its validity, authenticity, and integrity as evidence. Therefore, law enforcement officials need special investigative techniques that are able not only to secure data, but also to interpret it correctly according to linguistic, social, and legal contexts.

This is where forensic linguistics plays an important role. Forensic linguistics, as an interdisciplinary field that combines linguistics with law, focuses on the analysis of language relevant to criminal investigations and judicial processes. In the digital realm, forensic linguistics is applied for various purposes, ranging from author identification, communication pattern analysis, lie detection, interpretation of the meaning of speech, to the disclosure of hidden intentions from a text. For example, studies in the field of stylometry show that a person's writing style is relatively consistent so it can be used to identify individuals in criminal cases. Pragmatic and semantic analysis of e-mail messages can also help uncover non-explicit intent, which is often an important clue in the investigation of cases of online fraud, hate speech, or internet-based terrorism.

Furthermore, the development of Natural Language Processing (NLP) and artificial intelligence (AI) has further strengthened the role of forensic linguistics in the digital realm. Various text processing algorithms are now used to analyze big data that emerges from online communication. This allows investigators to find hidden patterns, analyze criminal communication networks, and systematically track propaganda or disinformation. However, the application of such technology also faces major challenges, such as the problem of algorithmic bias, the limitations of the language data corpus, and the ethical and privacy aspects inherent in the use of digital evidence.

Despite its enormous potential, the study of forensic linguistics in the context of digital evidence is still fragmented. Some studies place more emphasis on developing computational models for author identification, while others highlight the legal and regulatory aspects of the use of language as evidence. There are also studies that discuss methodological issues, such as how to determine scientific standards in linguistic analysis so that the results can be accepted in court. This fragmentation makes it difficult for academics, legal practitioners, and investigators to get a comprehensive picture of the development, effectiveness, and limitations of forensic linguistics in digital evidence-based investigations.

In addition, there is still a research gap in the integration between linguistic theory and criminal investigation practice. For example, a number of cybercrime cases in different countries show that linguistic analysis is often used on an ad hoc basis, without clear standard standards. This condition has implications for the court's difficulty in accepting the results of the analysis as valid evidence. In fact, if methodological standards and best practices can be built through research synthesis, then forensic linguistics will be increasingly recognized and widely used in the criminal justice system.

With this background, a comprehensive study is needed to collect, analyze, and evaluate previous studies in a systematic manner. Systematic Literature Review (SLR) is an appropriate method for this purpose as it allows researchers to sift through relevant literature, assess the quality of the methodology, and present a synthesis that can be accounted for. Through the SLR approach, this article aims to provide a comprehensive overview of how forensic linguistics is applied in digital evidence-based criminal investigations, evolving research trends, dominant methodologies, challenges faced, and prospects for its future development.

Thus, this research is expected to be able to contribute not only to the academic level, but also to law enforcement practice. Academically, this SLR will enrich the literature by presenting a more complete synthesis of knowledge. Practically, the results of this study can be a reference for law enforcement officials, investigators, and policymakers in developing criminal investigation strategies that are more effective, accurate, and adaptive to the dynamics of digital crime.

II. METHODS

This study uses the Systematic Literature Review (SLR) approach with the aim of collecting, analyzing, and synthesizing previous research related to the role of forensic linguistics in digital evidence-based criminal investigations. This approach was chosen because it is able to provide a comprehensive picture of scientific developments that are often still separate, as well as identify research gaps that can be the basis for future studies.

The initial stage was carried out through a literature search in various reputable scientific databases, such as Scopus, Web of Science, ScienceDirect, SpringerLink, Taylor & Francis Online, and IEEE Xplore. To enrich the coverage, additional searches are also conducted through Google Scholar and university repositories that provide open access. Article search uses a combination of keywords such as forensic linguistics, digital evidence, criminal investigation, author identification, cybercrime, and linguistic analysis.

Furthermore, literature screening was carried out based on inclusion and exclusion criteria. The selected articles are publications in the last ten years (2015–2025), in English or Indonesian, published in a scientific journal or conference proceedings that have gone through a peer review process, and directly discuss the application of forensic linguistics in the context of digital evidence or criminal investigation. Articles that are opinionated, editorial, or irrelevant to the focus of the study are excluded from the analysis process.

The next stage is the selection of relevant articles. Articles that pass the initial stage are thoroughly examined (full-text review) to ensure that they fit the focus of the research. Each article is analyzed based on bibliographic information, research objectives, methods used, types of digital evidence analyzed, key findings, and limitations.

The collected data is then analyzed through a thematic approach to find key patterns, categories, and themes. This analysis helps identify the contribution of forensic linguistics in criminal investigations, emerging methodological trends, as well as challenges faced in practice. The analysis process is carried out carefully by comparing and grouping the results of the research so as to produce a complete picture, not just a separate description.

The last stage is the synthesis of the study results, which is summarizing and integrating findings from various studies to provide a deeper understanding. This synthesis is expected to answer research questions, provide direction for the development of more effective investigative strategies, and highlight areas that still need further research. In this way, the results of this SLR can contribute academically and practically to strengthening the role of forensic linguistics in the handling of digital evidence in the realm of criminal investigation.

III. RESULT AND DISCUSSION

Below is the manuscript of the article under analysis, supplemented with detailed elaborations of the outcomes;

Table 1. Selected Empirical and Theoretical Studies on Language, Cybercrime, and Authorship Attribution

No.	Title	Author / Year	Focus & Relevance to the Topic	Notes Advantages / Limitations	Results and Findings
1.	Electronic crimes from a forensic linguistic point of view	Huda Salah Rasheed (2024)	Discusses how forensic linguistics is used to analyze electronic crime through text communication (social media, forums, messages). Suitable for general background and theory sections.	Advantages: up-to-date, geared towards digital proof; Limitations: there may not be much discussion of computational or ML techniques.	The article Huda Salah Rasheed (2024) highlights the role of forensic linguistics in the investigation of electronic crimes. The main finding of this study is that the majority of cybercrime relies on textual communication, so language analysis is one of the keys in uncovering perpetrators. Forensic linguistics has been proven to be helpful in identifying the author of the message, analyzing the communication style

					between the perpetrator and the victim, evaluating the risk of threats, and dismantling the recruitment or propaganda strategies of terrorist groups online. Several concrete cases are shown, such as murders that were revealed through the analysis of social media messages and cases of falsification of suicide messages. This study concludes that language functions not only as a means of communication, but also as the main evidence tool in uncovering the truth, so that the involvement of forensic linguists is equally important as investigators in the cybercrime justice process.
2.	Intelligent system for detection of cybercrime vocabulary on websites	Castillo-Zúñiga et al. (2024)	Combines NLP + machine learning to detect cybercrime vocabulary on websites. Relevant for the linguistic auto-detection technique of digital evidence.	Great as an engineering study; less focus on legal validity may be; The data may be web text, not all formal criminal contexts.	Research by Iván Castillo-Zúñiga et al. (2020) developed an intelligent system to detect cybercrime-related vocabulary on websites by utilizing Big Data Analytics, NLP, and machine learning. This system uses a web scraper to collect 1,326 cybercrime-related sites, then through the stages of data cleaning, tokenization, stopword removal, lemmatization, and the formation of a semantic ontology about cybercrime terms. The dataset was then analyzed using the Neural Network, Boosting, and Random Forest algorithms, with the best results from Random Forest

					achieving 97.64% accuracy through the LOOCV cross-validation test. In addition, the use of parallel processing with genetic strategies significantly improves time efficiency, with savings of 292% in corpus processing and 1220% in classification compared to sequential processes. These findings show that the integration of linguistic ontology and machine learning is very effective in automatically detecting cybercrime languages, and can be the basis for the development of digital threat detection systems in the future.
3.	Empirical evaluations of language-based author identification techniques	Carole E. Chaski (2001)	A classic study of author identification techniques based on various linguistic features: syntax, spelling errors, vocabulary density, etc. Relevant as a theoretical and technical basis.	Age enough, computational techniques may now be more advanced; but both for basic concepts and historical comparisons.	Carole E. Chaski's (2001/2007) article "Empirical Evaluations of Language-based Author Identification Techniques" shows that not all language-based authorship attribution techniques are reliable. Of the ten hypotheses tested on a sample of writings by four authors with similar sociodemographic backgrounds, only syntactic analysis and syntactically classified punctuation were proven to be able to distinguish and classify documents with a high degree of accuracy. Other techniques such as sentence complexity, vocabulary richness, readability, content analysis, and spelling/grammatical errors have high error

					rates or fail to replicate the results well. This study emphasizes the importance of empirical evaluation and measurement of error rates in forensic linguistics, while warning that methods that are subjective or unmeasurable should not be used as a basis for proving in court.
4.	Using word n-grams to identify authors and idiolects	David Wright (2017)	Use *the word n-grams* in a large corpus (Enron email) to identify anonymous authors. A highly relevant technique for digital evidence text analysis.	Advantages: big data, statistical techniques and corpus; Limitations: the context of English, may be less direct to the realm of law in Indonesia or general law.	David Wright's research shows that the use of the word n-grams can help identify a person's author and idiolects. From Enron's email corpus analysis, this method was able to guess anonymous text writers with an average accuracy of about 64%, even reaching more than 90% on a larger text sample. The effectiveness of n-gram lengths varies from author to author, and case studies show that certain n-grams are characteristic of individual idioms. These findings confirm that repetitive language patterns can be used as a basis for linguistic forensic analysis.
5.	Patterns of linguistics features in private chats of social media account leading someone to be a victim of a cyber crime	Teaching Pradika Ananta Tour (2024)	Analyze the language features in private chats related to cybercrime. Suitable for viewing digital evidence directly from social media/chat communications.	Advantages: local, social media context; Limitations: qualitative methods and samples may be limited; It is not as complete as automatic engineering.	Pradika Ananta Tur's (2019) teaching research revealed that linguistic features in private conversations on social media can be used by perpetrators to manipulate identities and commit cybercrimes. With a qualitative method through interviews and analysis of the victim's Facebook chat, this study found that the perpetrator

					imitated the victim's communication style through three main aspects: addressing, tone (tonal style/intonation in text), and spelling and punctuation patterns. For example, the use of special greetings such as mas, mbak, or familiar nicknames is the key to convincing the interlocutor that the message is indeed from the original account owner. In addition, the use of typical suffixes such as "-e" or "-je" in Javanese, as well as the habit of abbreviations or abbreviations of words, were also successfully imitated by the perpetrators, thereby reducing the suspicion of the victim and his friends. The results of this study confirm that linguistic patterns that are often considered trivial can actually be a security gap that allows victims' accounts to be hacked and used for fraud.
6.	The form and meaning of language motifs in phishing crimes: A forensic linguistic study	Zahy R. Ariyanto & Laili Etika Rahmawati (2025)	Focus on language motifs in phishing: form, persuasion, manipulation. Especially relevant is the "modus operandi" and analysis of digital criminal discourse.	Good because it's very up-to-date and specific; But the topic is phishing alone, perhaps lacking generalization to all types of digital evidence.	Research by Zahy Riswahyudha Ariyanto & Laili Etika Rahmawati (2025) found that phishing uses persuasive and manipulative language patterns to deceive victims. From the data analyzed, three main forms of phishing were identified, namely deceptive phishing (10 cases), APK phishing (4 cases), and smishing or SMS phishing (7 cases). The linguistic strategies used by the perpetrators include imitation of

					official authorities (for example, banks, government agencies, or companies), the use of formal and technical language, the threat of account blocking, and the lure of rewards. These messages are designed to create an illusion of trust and a sense of urgency, so victims are encouraged to immediately provide personal information or access malicious links without verification. These findings confirm that language is a key weapon in phishing, and the study of forensic linguistics can help identify these criminal communication patterns as well as improve people's digital literacy to prevent cybercrime.
7.	Authorship attribution, idiolectal style, and online identity: A specialised corpus of Najdi Arabic Tweets	Mashael AlAmr (2022)	Focus on anonymous writing on social media, using an idiolectal style for author identification. It is relevant if you want to study comparative languages or between cultural contexts.	Advantages: modern methods & social media data; Limitations: Arabic – may not be a good fit at all, depending on the language of your research.	This study examines the use of a combined approach between corpus linguistics, stylistics, computational linguistics, and computer-based communication to solve the problem of attribution of forensic authorship in Arabic. The main goal is to understand how real authorship can reveal theoretical concepts about idiolects. In terms of computer science, especially machine learning, the focus of the study is on the aspect of accuracy in determining authors. Corpus linguistics is a meeting point between these two fields, as both linguists and computer scientists

					alike utilize the corpus to test their methods. The difference is that linguists emphasize qualitative analysis with a stylistic point of view, so that the results can be accounted for in court as expert witnesses. On the other hand, computer scientists place more emphasis on statistical-based quantitative analysis that often produces "black box" tools, which are difficult to trace the reasoning process.
8.	Towards a Linguistic Stylometric Model for the Authorship Detection in Cybercrime Investigations	Abdulfattah Omar & Aldawsari Bader Deraan (latest year)	Combines morpho-lexical features and frequency of letter pairs for author identification in short text from Twitter. Highly relevant for anonymous digital text analysis/cybercrime.	Advantages: contemporary techniques, short/anonymous texts; Limitations: specific social media platforms, perhaps small datasets versus large scales.	Research by Abdulfattah Omar and Aldawsari Bader Deraan proposes a morpho-lexical stylometry model to detect the authorship of very short texts on social media, especially Twitter. By analyzing 12,240 tweets from 87 accounts related to the issue of the removal of Confederate monuments in the United States, the study combined the frequency of letter-pairs and lexical features using the Self-Organizing Map (SOM) as a classification method. The results showed that this model was able to achieve an accuracy of about 76% in identifying anonymous text authors. The accuracy rate decreased by 22% when using only typical words, and decreased by 26% when relying only on morphological patterns or letter combinations. Thus, the integration of different linguistic variables in a single

					system has been shown to improve the performance of short-text classification. These findings confirm the potential of SOM-based morpho-lexical methods in forensic linguistic investigations, especially in detecting perpetrators of hate speech and cybercrime on social media.
9.	Author Identification from Literary Articles with Visual Features: A Case Study with Bangla Documents	Ankita Dhar et al. (2022)	Focus on literary documents with visual + text features for author identification. It's useful if you want to expand to evidence that isn't just text: it's also visual, layout, or document format.	Advantages: different methods (visual + text); Limitations: literary is not criminal communication; may not be direct to "forensic legal" digital evidence.	Research by Ankita Dhar et al. proposed a CNN-based author identification system by utilizing the visual features of Bangla-language literary texts. A total of 1200 articles from 50 authors were collected, then processed through tokenization, stopword removal, and extraction of statistical and textual features visualized in the form of graphs (line, pie, imagesc) and entered into the five-layer CNN network. The results of the experiment showed a maximum accuracy of 93.58%, far exceeding manual feature-based methods and traditional machine learning algorithms. The system was also tested on the English dataset (C50) with an accuracy of 93.52%, proving its cross-lingual nature. These findings confirm that the visual feature-based CNN approach can recognize writing style patterns more effectively than conventional techniques, and is relevant for literature applications, linguistic

					forensics, and plagiarism detection.
10.	Author Identification, Idiolect, and Linguistic Uniqueness	Malcolm Coulthard (2004)	The theory of idiolect and linguistic uniqueness as the basis for how the identity of the author can be expressed through unique features. It is suitable for theoretical foundations.	Advantages: in-depth theory; Limitations: a bit old and may not have used the latest technology; English language and academic context.	This article discusses the concept of idiolects and the uniqueness of individual speech that has long been discussed in linguistics, as well as how they can be used to answer questions related to attribution of written documents. The main focus is to assess the extent to which similarities between texts can indicate plagiarism or the same authorship. Two measurement methods were used, namely the proportion of the same vocabulary and the number and length of the phrase were identical. Through real-life case examples from both criminal courts and student plagiarism cases, this article shows the relevance of forensic linguistic analysis in distinguishing the authenticity of texts. In the final section, the authors also highlight the challenges of receiving this linguistic evidence in American courts as well as strategies for making the findings understandable to a lay audience.

The results of the literature search show that the application of forensic linguistics in digital evidence-based criminal investigations develops in two main approaches: (1) a qualitative linguistic approach based on interpretation of styles and discourse, and (2) a computational approach based on statistics and artificial intelligence. These two approaches do not stand separately, but complement each other in revealing the identity of the perpetrator, the mode of criminal communication, and building the validity of language evidence in the legal realm.

In the realm of qualitative approaches, studies such as Rasheed (2024) and Tur (2019) confirm that most cybercrimes still rely heavily on text-based interpersonal communication, either through social media, private messages, or online forums. Analysis of greeting structure, language style, and spelling habits has been proven to be able to reveal perpetrator-victim relationships, dismantle identity manipulation strategies, and identify hidden criminal intentions. The research of Ariyanto & Rahmawati (2025) strengthens these findings by showing that the mode of linguistic persuasion in phishing is deliberately designed to build a

sense of urgency and trust. This shows that language is not just a means of communication, but a psychological instrument in cybercrime.

On the other hand, the quantitative-computational approach offers accuracy and scalability through automated modeling. Wright's (2017) study proved that the word n-grams can be used to identify anonymous authors with an accuracy of up to 64–90%. However, Chaski's (2001/2007) research provides an important criticism that not all linguistic features are forensically valid; Only syntactic features and grammatically structured punctuation have proven consistent in distinguishing authors. This confirms that although technology produces accuracy figures, not all computational results can be legally accounted for.

More recent developments can be seen in the research of Omar & Deraan (2023) and Ankita et al. (2022) which integrates machine learning and deep learning in the detection of short text authorship. With the combination of letter-pair frequency, lexical features, and SOM and CNN models, the author's classification accuracy reached 76–93%, indicating that the integration of linguistic multi-features is much more effective than the use of a single type of feature. Castillo-Zúñiga's research (2024) even shows that the Big Data + Random Forest approach is able to detect cybercrime vocabulary with 97.64% accuracy, indicating that forensic linguistics is now entering the stage of language-based predictive policing.

However, this study also found methodological and epistemological gaps. Qualitative research is rich in context, but it is often considered less objective and difficult to replicate. In contrast, computational research has high accuracy but is often considered a "black box" that does not provide a linguistic explanation for the results. It is at this point that idiolectal theories as discussed by Coulthard (2004) and AlAmr (2022) become an epistemological bridge: both show that individual language styles are both systematic and explainable, thus allowing for integration between statistical accuracy and qualitative explanations that are valid in court.

Overall, the results of this SLR show that the main strength of forensic linguistics on digital evidence lies in the combination of discourse interpretation, micropattern tracing, and automated statistical verification. Digital language is not only analyzed as text, but as psychological, social, and algorithmic traces that record the behavior of criminals. With the development of more standardized methodologies, forensic linguistics has the potential to become one of the main pillars of future cybercrime investigations, not just a complement, but a core evidentiary tool capable of uncovering perpetrators even if they hide behind the scenes and encryption.

IV. CONCLUSION

Based on the synthesis of various studies, it can be concluded that forensic linguistics in the realm of digital crime has evolved into a hybrid discipline that integrates qualitative analysis based on human language style with automated computation based on big data. Language is no longer seen only as a means of communication, but as an identity imprint (idiolectal), an instrument of psychological manipulation, as well as legal evidence. The qualitative approach shows that micro-characteristics such as greeting, spelling, text intonation, and persuasion patterns play an important role in the reconstruction of perpetrator-victim relations and crime modes. On the other hand, the computational approach proves that its accuracy can reach 90–97% when linguistic features are combined with statistical algorithms or deep learning. However, both have limitations when they stand alone. Therefore, the future of digital forensic linguistics lies in integrative models that are capable of explaining and calculating language styles, so that the results are technically accurate and legally valid.

V. REFERENCES

Ariyanto, Z. R., & Rahmawati, L. E. (2025). The form and meaning of language motifs in phishing crimes: A forensic linguistic study. LITE: Journal of Language, Literature, and Culture, 21(1), 55–70. https://publikasi.dinus.ac.id/index.php/lite/article/view/11463

Castillo-Zúñiga, I., Loya, V., Díaz, C., & Pérez, L. (2024). Intelligent system for detection of cybercrime vocabulary on websites. DYNA New Technologies, 9(2), 65–78. https://revista-dyna.com/dyna-newtech/article/view/92

Chaski, C. E. (2001). Empirical evaluations of language-based author identification techniques. Forensic Linguistics, 8(1), 1–65. https://journal.equinoxpub.com/IJSLL/article/view/9905

Coulthard, M. (2004). Author identification, idiolect, and linguistic uniqueness. Applied Linguistics, 25(4), 431–447. https://doi.org/10.1093/applin/25.4.431

Dhar, A., Roy, A., & Choudhury, P. (2022). Author identification from literary articles with visual features: A case study with Bangla documents. Future Internet, 14(10), 272. https://doi.org/10.3390/fi14100272

Omar, A., & Deraan, A. B. (2023). Towards a linguistic stylometric model for the authorship detection in cybercrime investigations. International Journal of English Linguistics, 13(3), 14–26. https://ccsenet.org/journal/index.php/ijel/article/view/0/40496

Rasheed, H. S. (2024). Electronic crimes from a forensic linguistic point of view. TWEJER: Journal of Literature,

Linguistics and Cultural Studies, 5(1), 88–103. https://journals.soran.edu.iq/index.php/Twejer/article/view/674

Tur, A. P. A. (2019). Patterns of linguistic features in private chats of social media accounts leading someone to be a victim of a cyber crime. LEXICA: Journal of Language, Literature and Its Teaching, 13(2), 127–137. https://jurnalnasional2.ump.ac.id/index.php/LEKSIKA/article/view/3858

Wright, D. (2017). Using word n-grams to identify authors and idiolects. International Journal of Corpus Linguistics, 22(2), 212–241. https://doi.org/10.1075/ijcl.22.2.03wri

AlAmr, M. (2022). Authorship attribution, idiolectal style, and online identity: A specialised corpus of Najdi Arabic tweets. International Journal of Speech, Language and the Law, 29(1), 1–25. https://journal.equinoxpub.com/IJSLL/article/view/27343