

**An Assessment of methods and techniques used in Forensic  
Linguistics**

**Raed toghuj**

## ***Introduction***

Forensic linguistics is based on the discourse that occurs in court room, court room translation and interpretation, comprehensibility of texts and legal documents, comprising the police caution issued to criminals or suspects, and the utilisation of linguistic evidence in the courts proceedings. In the last few years, despite of the fact regarding advanced development of the state-of-the-art forensic linguistic methods and techniques such as speaker recognition system have depicted considerable progression in the last few years (Bali et al., 2020). Level of performance of these system do not yet appears to warrant massive introduction in comparison to applications with low risk (Marfianto and Riadi, 2018). Conditions and situations typical of forensic background for instance difference in equipment of recording and channels of transmission, the existence of background noise and variation due to communicative context differences endure to pose major challenge in the field of forensic linguistics (Baror et al., 2021). Consequently, the influence of technology of automatic speaker recognition and related technologies has been comparatively uncertain and practice of forensic speaker identification stays massively dominated by the utilisation of broad variety of hugely subjective processes (Wright, 2017).

Whereas current progresses in interpretation and analyses of the evidential worth of forensic indication and proofs evidently favour techniques that make it likely for fallouts to be articulated in perspective of a probability ratio, dissimilar to automatic processes, outdated methods and techniques in the arena of speaker identification do not usually encounter this necessity (Phakiti et al., 2018). Though, conclusions in binary yes or no-decision form or a qualified declaration of the likelihood of the hypothesis instead of the evidence are progressively criticised for being rationally defective. In contradiction of this background, the essential to put alternative authentication processes in place is developing as more extensively accepted (Baranov, 2017). Therefore, this research aims to examine and analyse

advanced methods and techniques used in forensic linguistics contributing to reduce highlighted issues in the field.

### ***Literature Review***

#### ***Forensic discourse analysis***

Forensic Linguistics tries to define and, where likely, elucidate the aspects that differentiate the linguistic used in lawful settings from the ordinary language. Also, discourse analysis is proficient of application in an extensive variety of contexts. Under discourse analysis, undercover recording is an instrument used by law enforcement to help examination by apprehending important evidence that rises in conversation (Baranov, 2017). Supposing that such recordings are of satisfactory quality to catch any suspect and collaborating witness or secret officer, it would appear that this kind of evidence would be very dependable. Though, Georgetown linguist Roger Shuy exemplifies ways what is heard is not unavoidably what was stated or meant, nor is it unavoidably even pertinent to the supposed criminal action. In [Texas v. T. Cullen Davis], the trial had reliant upon secret recordings of Davis and a liaising witness, David MCrory, who was hired supposedly by Davis to kill his wife and the magistrate ruling over their divorce (Filipovic and Gascón, 2018).

#### ***Authorship Analysis***

In cases where the documents' authorship is a set of documents or a single document is an argumentative, a linguist might be requested to discourse on the authorship of the conflicted or disputed writing (Heydon, 2019). This characteristically entails likening the conflicted sample with samples recognised to have ben documented and written by specific individual. There is one major approaches to authorship analysis: that is considered as linguistic. This includes linguistic approach that is taken into account as more conventional approach to the poof of language and it also includes a principled choice and selection of qualitative and quantitative analyses. Under this approach, one piece of writing is compared with another for ruling out or ruling in the common authorship (Filipovic and Gascón,

2018). Moreover, the notion of linguistic fingerprint is unique to every user of language, while it is assumed by Coulthard that individuals, with unique preferences and vocabularies for specific structures and combinations, characterise the style of user (Marfianto and Riadi, 2018). "Exclusive markers are tremendously sporadic, so authorship [ascription] necessitates the identification of a markers, each of which might be initiated in other writers.

### ***Sociolinguistic profiling***

This method under forensic linguistics discipline allows for providing information regarding the ways the functions of language in the background through classification and categorisation that finest define the range of ways language is operated in the legal context (Marfianto and Riadi, 2018). This method and technique under the domain of forensic linguistics, allows the forensic linguist to determine the information regarding authors who are anonymous or the secret origin and bases of the texts. The linguist by the use of this technique might not make psychological observations regarding the authors or their intentions, however, reliant on the features, in the text, they can also be able to illustrate the social origins of the authors or their background. The method has been used widely along with communications enhanced and mediated by computer and there has been several documented arenas of it being advantageous to the consequence of a case and the justice provision (Baror et al., 2021). Conclusions regarding the probable context of an anonymous author is likely to ever be sufficiently certain to offer proves and evidence for usage of courtroom, however, as evidence by means of prior case work, they could be utilised to determine the case logics, identifying any flaws or wrong information being provided by suspects and enhancing the case proceedings (Phakiti et al., 2018). This technique also allows the investigators to determine where the text has been written, the equipment used for its writing and any tempering that made to text-based evidences. This allows to smoothen the legal hearings processes and evidence collection (Baranov, 2017).

### ***Forensic phonetics***

Wright, (2017) argues that forensic phonetics is the implementation of methods, theories, knowledge of usual phonetic to practical jobs and tasks that rise from context of police of work or the presentation of proofs and evidence in the magistrates, as well as the development of novel, particularly forensic-phonetic theories, knowledge and methods. The most central feature of forensic phonetic is identification of speaker which is also recognised as speaker recognition they might have particular meanings in specific lawful systems or scientific backgrounds (Wright, 2017). Another feature is the task of assessing the linguistic context of speech passages where indelibility is robustly decreased for behavioural and technical causes. Identification of speakers, which has been classified as the most major feature of forensic phonetics comprising comparisons and profiling of voices (Baror et al., 2021).

Under voice comparisons, there is a presence of speech recordings of an unknown speaker who could be related to the crime. For instance, this could be a kidnapper asking ransom on the phone. Furthermore, someone is suspected to be similar with the unidentified individual and a recording of speech that is obtainable or could be formulated (Wright, 2017). Relying on the lawful system, conversations on tapped telephones or interviews of police that are recorded could be utilised as proof if the suspect is not cooperating or if additional evidences are required or more significant material of speech is required. Subsequent to analysis that is completed, the conclusion could be reached that is pertinent to the question if the speech samples that is likened originate from the similar or from diverse individuals (Baror et al., 2021). Relatively, voice comparisons could be requested as a process for police investigations or privately deprived of going to court, however, most commonly, voice comparisons fallouts in documented, scientifically and systematically motivated reports that are utilised as evidence in magistrates and that relying on the law system, local

regulations or country-wide regulations, have to be illustrated and defended in the magistrates verbally by the expert accountable for the report (Phakiti et al., 2018).

### ***Methods***

This section presents the methods and techniques used to perform this research. This section elaborates and lays out the steps involved in doing the research, as well as the methods employed. As a result, the chapter is crucial to the research paper's overall success. Moreover, it also covers the project's overall structure, which involves the creation of numerous data gathering methodologies as well as techniques for analysing the collected data. The research design specifies how data will be gathered and the data analysis procedure for the study. Quantitative, qualitative, or mixed methods research is possible (Camacho 2020). Quantitative research requires the gathering of numerical data, which may be done by turning participant replies into numbers or acquiring financial statistics from various organizations' financial statements. Because the quantitative method demands numerical data, it will not focus on the respondents' specific replies; consequently, questionnaires may be used to obtain such responses from respondents, and the responses can then be evaluated using statistical tools like SPSS. Qualitative research comprises collecting data on respondents' replies and examining how they perceived the questions and settings they were asked to reply to (Chu ). In their questionnaire, the qualitative study contains open-ended questions as well as face-to-face interviews to extract the required details from the replies. The mixed design, which mixes qualitative and quantitative research approaches, is another sort of research design. For this following paper a qualitative research design is chosen and data collected will be done from secondary sources.

### ***Content analysis***

From the analysis of literature review, the arguments and results of forensic linguistic researchers and experts reveals that audio analysis and forensic approach is the fine methods and that has increase in the last few years (Marfianto and Riadi, 2018). These techniques have been found to signal further developments in the

interpretation and analysis of evidential value of evidences provided forensics (Phakiti et al., 2018). More prominently, there are also strong signs of a mounting awareness among investigators and law experts working in the forensic audio analysis and speech field in order to view validation of the techniques utilised as integral part of the discipline. The forensic methods and techniques have been found to segment and enhance the investigation flow, collection of evidences and court procedures (Baranov, 2017). Authorship ascription is perhaps the oldest techniques in the implementation of forensic linguistics. However, the prejudiced power of the approaches utilised so far stays comparatively weak to other discussed methods and techniques, if it has not been revealed to be completely missing.

### ***Analysis***

The aim of this section in this report is to presents results of the secondary research methodology defined above. The analysis presents applications and benefits of forensic linguistic techniques followed by discussion of limitations in these techniques.

### ***Applications and Benefits FLT***

#### **Court of Law**

The evolution of FLT can be traced back to analysing evidence to assess their validity and for detective work. There are several cases that used forensic experts to provide testimony or expert opinion on evidence in court. For example in Uter, Abouakkour, Roberts, Wood and Williams [2014] an expert was called who analysed telephone call recordings of different suspects and the recordings of their interviews during investigation. Through comparative analyses the expert concluded that there was high similarity in samples. As a result of analysis the court convicted by defendants for kidnapping, witness intimidation, and torture. Similarly, in another case in the UK Julie Dawn Lunn [2012] an expert was called by the police department to compare blackmail letters received by victims with written specimens of the defendant and concluded that there was reasonable similarity for her conviction. In Rizwan

Ahmed and others [2012] two experts were consulted by the defendant in a murder case. They analysed slang text messages exchanged by co-accused. The court based on expert opinion and other evidence decided to drop charges against one defendant. There are also examples of cases who used experts either by defendant or to analyse linguistics such as internet chats and to solve or evaluate evidence for murder cases (Milom, 2019).

### **Confessions and denials by public figures**

There is also research indicating that forensic linguistic can be used to evaluate the opinions and speeches of public figures. Many scholars argued that the public confessions and denials are important linguistically because these provide significant evidence base for research. Furthermore, the author argued that public figures in response to their accusers whether in public or in private offer language that can be used to further investigate issues. Although, these language and speeches are diverse there is possibility to analyse the prominent figures such as public officials (Kupper and Meloy, 2021).

### **Plagiarism Detection**

In this subject the aim is to use scientific techniques to analyse written or spoken language or text in order to determine and measure the content and meaning, to identify the speaker, or to determine the author such as in case of plagiarism. Consider for example the case of plagiarism detection in which forensic stylistics was applied and focused on the accusation of plagiarism by Helen Keller in her story "The Frost King". She was a deaf-blind author and was accused of plagiarism in 1892. The investigation concluded that her story was plagiarized and was based on another story Frost Fairies by Margaret Canby (Rasheed, 2018). Keller had heard the latter story. The investigation concluded that Keller made merely minute changes to original work and replaced common words and phrases by less common ones and narrated the same story. Thus in other words Keller only represented the same story in different words. Based on the conclusion that Keller manipulated words by using less common lexis the forensic stylistic proved the case of plagiarism (Foltýnek, Meuschke, and Gipp,



2019). More specifically, the investigation applied Flesch and Flesch-Kincaid readability test and the results showed that there was more originality in Canby's text as compared to Keller's. Furthermore, the test also provided higher grade to Canby's text on a scale that measures reading-ease. Therefore, forensic stylistics were applied to prove plagiarism (Allan, Bradshaw, and Finch, 2017).

There are also other cases in which plagiarism was investigated using forensic stylistics such as the case of Richard Condon and Robert Graves; and case between Archibald Carey and Martin Luther King Jr. The evaluation of the text Condon using linguistic technique indicated that it was rich in clichés (Bygate, 2004). In contrast to Helen Keller who was known for using rare phrases instead of using common words, the investigation revealed that Condon was more inclined towards using and expanding commonly used words. Furthermore, the investigation also told that Condon had many themes common with the works of Graves. On the other hand, the investigation for plagiarism between Archibald Carey and King, it was reported that around 50% of the doctoral dissertation by King was copied from a peer theology student (Chapelle, 2013). The report showed that King used more assonance and alliteration and only changed some names while presenting the same thesis. Furthermore, the investigations also showed that the source text in these cases were considerably shorter, simpler, and pithier with respect to structure however, the accused works were categorised as "purple" devices which refer to extending current text and developing/reshaping language to produce different product (de Bot, 2000).

There is also a famous case of the Unabomber, in which the investigation found Theodore Kaczynski as the main convict. The Family members had recognized his style of writing in the notorious manifesto titled 'Industrial Society and Its Future' and then they reported to the authorities (Cotterill, 2010). The FBI agents who searched Kaczynski's hut collected several documents that were authored by Kaczynski but were not published. One of the FBI agent conducted an analyses and reported that

there were numerous common lexical items in the documents. The prosecution based her argument on the bases of commonalities in the texts (Ahmed, 2021).

### **Deception Detection**

The forensic linguistic techniques also led to the development of Voice Risk Analysis (VRA) technology which can be used for detection of deception and fraudulent activities is. This technology has been used frequently in anti-fraud investigations such as insurance industry has used these technologies to detect possible frauds. Similarly local authorities also use them for their investigations. In the UK, the corporate users of VRA technologies have reported that these can be used to reduce fraud however, there are certain limitations due to which there are challenges in accurate identification of fraudsters (Kaplan, Ling, and Cuellar, 2020). It is important to note that this technology is not based on phonetics as discussed in literature review section however; there is a wide range of experts who have shown doubts and concerns about the validity and reliability. There is also research presented by business experts who used simulation tools in lab settings and concluded that VRA can be used to detect proxies for deception (Garrett, Crozier, and Grady, 2020).

### ***Limitations***

#### **Regulation of Expertise**

Despite the fact that there are successful cases stories as presented in previous section yet there are certain limitations in forensic linguistic technologies. One of the main limitations is low-standard expertise may mislead justice system and establishing reliability and authority of expert is rather difficult (Foulkes, French, and Wilson, 2019). Although there has been no case of misuse or misfire by forensic linguistic technologies that affected the justice system in the UK (McCartney and Amoako, 2017), but there are cases where people were wrongly incarcerated because of erroneous forensic linguistic investigation. There is lack of statutory regulation to identify an accredited experts in the UK which can be considered as an obstacle to development and adoption of these technologies (Meluzzi, Cenceschi, and Trivilini, 2020).

### **Scientific Critiques**

There are also several scientific critiques by practitioners and academic who posit that it is a dangerous phenomenon, termed the CSI effect, due to which the expectations and trust on forensic technologies in the public and practitioners had falsely improved based on TV series and film (Singh and Sudhakar, 2017). There fundamental problems in this regard is that the jurors have expectations that certain procedures are possible but experts reject these tests. For example, tests for personality analysis, truth detection tests, and threat assessment tests for speech intonation have been rejected by experts to provide reliable results yet public perception is positive (McMenamin, 2020).

### **Expressing Conclusions**

The linguists and phoneticians have shown frustration due to the pressure upon them to present objective conclusions with high level of certainty which in reality are not possible. Experts say that it is unreasonable to ask for expressed and explicit conclusions based on evolving and subjective analytical techniques (McMenamin, 2020). On the other hand, it is not simple for juries and judges to fully understand technical terms to evaluate whether conclusions presented are reliable for making decisions or how they can inform decisions. Furthermore, legal experts have also raised concerns about the comprehension and applicability of probability statistics in making legal and judicial decisions (Meluzzi, Cenceschi, and Trivilini, 2020). The fundamental limitation is to evaluate the degree to which conclusions based on forensic sciences and technologies can be accurately applied to court decisions (Singh and Sudhakar, 2017).

### **Technical Limitations**

There is also significant number of experts and researchers who presented research and highlighted limitations in forensic linguistic technologies particularly those that have been used in criminal investigations cases (Foulkes, French, and Wilson, 2019). One of the common themes in this stream of literature as identified by Singh and Sudhakar, (2017) is that whilst investigations use sociolinguistic profiling,

it lacks the validity and reliability to be used as evidence in criminal convictions. The study argued that it is possible to effectively use voice parades by applying phoneticians and relevant techniques but these are rather expensive and time consuming (McMenamin, 2020).

There is on-going research about identification of ways that can streamline design and application of voice parades but there is no significant development. There is also considerable scepticism among psychologists regarding voice parades as they argue that the nature of voice and memory is complex and reliance is not valid (Meluzzi, Cenceschi, and Trivilini, 2020). It has been reported that there is lower reliability in voice recognition as compared to face recognition and even face recognition technologies do not provide results with high level of certainty (Foulkes, French, and Wilson, 2019).

### ***Conclusion***

The following paper aimed to evaluate the techniques and methods adopted in forensic linguistics. The paper is divided to 5 sections the first section discussed a brief background of the topic while the second chapter aimed to review literature from previous studies. The third section depicted the methods and techniques incorporated in carrying out this research. The fourth section analysed the data gathered for this paper while the last section concluded the paper. It is seen in this research that Courtroom speech, courtroom translation and interpretation, comprehensibility of texts and legal documents, including police cautions issued to offenders or suspects, and the use of linguistic evidence in court procedures are all part of forensic linguistics. As seen in the analysis section the arguments and outcomes of forensic linguistic researchers and professionals demonstrate that audio analysis and forensic approach are great approaches that have increased in the last few years. Moreover, the results reveal that in circumstances when the authorship of a series of papers or a single document is contested, a linguist can be called in to discuss the authorship of the disputed or conflicting material. Hence this typically indicates comparing the

contested sample to specimens that are known to have been documented and authored by a certain person.

Moreover, it is seen in this research that under the forensic linguistics field, sociolinguistic profiling provides for the provision of information about how language operates in the background through categorization and categorisation that best identify the variety of ways language is used in the legal context. This approach and technique, which falls within the subject of forensic linguistics, allows a forensic linguist to determine information about anonymous writers or the writings' secret origins and grounds. Moreover, the study concludes that this method also allows investigators to discover where the text was produced, the equipment used to write it, and any text-based evidence that has been tempered. This facilitates the legal hearings and evidence collecting processes. Therefore it can be concluded that forensic linguistics is an important part to court proceeding which if done right can reveal evidence that can aid in solving complex cases more efficiently.

## References

- Ahmed, H.R., 2021. The Role of Forensic Linguistics in Crime Investigation: Uses in Legal Proceedings. *ANGLISTICUM. Journal of the Association-Institute for English Language and American Studies*, 10(2), pp.23-31.
- Allan, K., Bradshaw, J. and Finch, G., 2017. *The English language and linguistics companion*. Red Globe Press.
- Bali, A.S., Edmond, G., Ballantyne, K.N., Kemp, R.I. and Martire, K.A., 2020. Communicating forensic science opinion: an examination of expert reporting practices. *Science & Justice*, 60(3), pp.216-224.
- Baranov, A.N., 2017. Linguistics in forensic linguistic expertise (method and truth). *Vestnik Volgogradskogo Gosudarstvennogo Universiteta. Seriiā 2, Īazykoznanie*, 16(2), p.18.
- Baror, S.O., Venter, H.S. and Adeyemi, R., 2021. A natural human language framework for digital forensic readiness in the public cloud. *Australian Journal of Forensic Sciences*, 53(5), pp.566-591.
- Bygate, M., 2004. Some current trends in applied linguistics: Towards a generic view. *AILA review*, 17(1), pp.6-22.
- Camacho, D., Panizo-LLedot, Ā., Bello-Orgaz, G., Gonzalez-Pardo, A. and Cambria, E., 2020. The four dimensions of social network analysis: An overview of research methods, applications, and software tools. *Information Fusion*, 63, pp.88-120.
- Chapelle, C., 2013. *The encyclopedia of applied linguistics*. John Wiley and Sons, Inc..
- Chu, H. and Ke, Q., 2017. Research methods: What's in the name?. *Library & Information Science Research*, 39(4), pp.284-294.
- Cotterill, J., 2010. How to use corpus linguistics in forensic linguistics. In *The Routledge handbook of corpus linguistics* (pp. 578-590). Routledge.

- de Bot, K., 2000. Psycholinguistics in applied linguistics: Trends and perspectives. *Annual Review of Applied Linguistics*, 20, pp.224-237.
- Filipovic, L. and Gascón, A.H., 2018. Interpreting meaning in police interviews: applied language typology in a forensic linguistics context. *Vigo International Journal of Applied Linguistics*, (15), pp.67-104.
- Foltýnek, T., Meuschke, N. and Gipp, B., 2019. Academic Plagiarism Detection: A Systematic Literature.
- Foulkes, P., French, P. and Wilson, K., 2019. LADO as forensic speaker profiling. In *Language analysis for the determination of origin* (pp. 91-116). Springer, Cham.
- Garrett, B.L., Crozier, W.E. and Grady, R., 2020. Error rates, likelihood ratios, and jury evaluation of forensic evidence. *Journal of forensic sciences*, 65(4), pp.1199-1209.
- Heydon, G., 2019. Researching Forensic Linguistics: Approaches and Applications. Routledge.
- Kaplan, J., Ling, S. and Cuellar, M., 2020. Public beliefs about the accuracy and importance of forensic evidence in the United States. *Science & Justice*, 60(3), pp.263-272.
- Kupper, J. and Meloy, J.R., 2021. TRAP-18 indicators validated through the forensic linguistic analysis of targeted violence manifestos. *Journal of threat assessment and management*.
- Marfianto, A. and Riadi, I., 2018. WhatsApp messenger forensic analysis based on Android using text mining method. *International Journal of Cyber Security and Digital Forensics (IJCSDF)*, 7(3), pp.319-327.
- McCartney, C. and Amoako, E., 2017. The UK forensic science regulator: a model for forensic science regulation. *Ga. St. UL Rev.*, 34, p.945.

- McMenamin, G.R., 2020. Forensic stylistics: The theory and practice of forensic stylistics. *The Routledge handbook of forensic linguistics*, pp.539-557.
- Meluzzi, C., Cenceschi, S. and Trivilini, A., 2020. Data in Forensic Phonetics: from theory to practice.
- Milom, V.L., 2019. *Native Language Analysis in Forensic Linguistic Contexts: A Pilot Study Using Four Languages*. Hofstra University.
- Phakiti, A., De Costa, P., Plonsky, L. and Starfield, S., 2018. Applied linguistics research: Current issues, methods, and trends. In *The Palgrave handbook of applied linguistics research methodology* (pp. 5-29). Palgrave Macmillan, London.
- Rasheed, N.J., 2018. FORENSIC PSYCHOLINGUISTICS: USING LANGUAGE ANALYSIS FOR IDENTIFYING AND ASSESSING OFFENDERS.
- Singh, A.K. and Sudhakar, A., 2017. Ethical Questions in NLP Research: The (Mis)-Use of Forensic Linguistics. *arXiv preprint arXiv:1712.07512*.
- Wright, D., 2017. Using word n-grams to identify authors and idiolects: A corpus approach to a forensic linguistic problem. *International Journal of Corpus Linguistics*, 22(2), pp.212-241.