



Forensic Linguistics function in Evidentiary and Investigative Contexts

Raed Toghuj

وظيفة اللسانيات الجنائية في السياقات الإستدلالية والإستقصائية

رعد تغوج

Forensic Linguistics function in Evidentiary and Investigative Contexts

Recently, forensic linguistics has been an arena of significance in many fields of study especially in judicial systems, legal and forensic matters, investigation, and open-source intelligence across the globe. The term typically refers to legal and professional analysis of recorded or written language by experts (forensic linguists) to provide expert and correct interpretation. It is particularly used in legal matters especially in the court and criminal justice systems. In the court system, forensic linguistics is heavily applied to examine language evidence – either recorded in voice or handwritten in civil matters or crimes. The analysis or examination is carried out for two major reasons. First, the analysis is utilized when relevant investigations are carried out with a focus to help in identifying witnesses or suspects in specific cases or scenes, or the determination of the significance of writing or utterance to a case. Secondly, forensic linguistics plays a pivotal role when written or spoken language samples are presented to a court as evidence. In such contexts, forensic linguists provide expert testimonies of correct interpretation of the samples. As such,

language analysis is significant in any judicial matters and systems provided the questionable language constitutes crimes. In most cases, crimes such as threats, hate speech, bribery, hate literature, coercion among others necessitate the use of a linguist expert for correct and most importantly professional interpretation. Evidently, the concept of forensic linguistics is ascribed to provide the truth from recorded speeches or voices and written languages in the face of a crime or relevant legal investigation matters. This paper will posit on the different ways and methods that forensic linguistics is applied to investigate and provide professional interpretation of recorded and written languages in evidentiary and investigative contexts.

Authorship analysis and attribution

Authorship analysis has over the years been increasingly used in forensic linguistic applications by linguists due to its ability to do language processing. The questions “why might have written this text” triggered computer scientists to invent software with the capacity to detect the possible author of a particular text from a group of suspected individuals. As one of the branches of forensic linguistics, authorship analysis involves a keen and detailed analysis of a particular quoted text, often questionable, with a sole focus to identify the most probable author (Krassa, 2020). In most cases, forensic linguists use authorship

analysis to identify, from two or more suspects, the author of a written text in the face of uncertainty and anonymity on a piece of writing's author. The technique is mostly used by relevant experts when solving cases related to disputed authorship in cases such as malicious texts, abusive and threatening cases among others. According to Ariani et al. (2014), forensic linguists often use authorship identification in legal proceedings to correctly try identifying whether a particular suspected individual wrote or said a given text through a thorough and expert analysis of significant language aspects such as idiolect and specific patterns associated with language use including grammar, spelling, collocations, vocabulary, and pronunciation.

Correa (2013) asserts that forensic linguists carry out text author identification through linguistic analyses which pay keen and close attention to the writing styles which are later related to suspect(s) to identify the most probable writer of a questionable text. According to this scholar, forensic linguists rely on this method, which has been empirically proven to bear positive results to a larger extent, based on the idea that there are individual variations on the way people use a particular language. That is to say, a certain language cannot be used in the same way, in this context, texts, and speeches, between two or more individuals. The scholar (Correa, 2013), further stipulates that forensic linguists

typically compare a presumed author(s)' known writings with the questionable text presented as evidence to ascertain the likelihood of the suspect being the author. Other scholars like R Gupta (2019) and Krassa (2020) agree with the above assertions that forensic linguists use features and linguistic styles of questioned pieces of writing or text and compare them with known texts or previous writings by suspected authors.

In efforts to boost the validity of authorship analysis, experts use computer software to analyze questioned texts especially in the event they are long. However, full reliance on the utilization of computers in authorship analysis is limited for forensic linguists based on the fact that computers provide a minimal explanation about their analysis outcomes. Most literature points out different methods used by forensic linguists in authorship attribution such as "part-of-speech n-grams and function words frequencies" as the most reliable in identifying the original author of questionable texts among suspected individual(s). The methods used are overly based on assumptions that pay attention to similarity detection of the text under review and previous texts by the suspects. Krassa (2021) points out that forensic experts in linguistics use different features to identify an author during authorship attribution such as content, structural, syntactic, and lexical features. Other linguistic forensic

experts use stylometric features for privacy and forensics assessments alongside other features such as relative frequency of words, vocabulary richness, *n*-grams parts in a text or speech among others.

Plagiarism Detection

The detection of plagiarism is one of the areas in computational linguistics that has attracted great interest in the industry and research. The area of plagiarism detection has over the recent past developed, expanding its applications to learning institutions where it is extensively used to detect the use of improper texts, in this case, someone else's work as assuming it is own. The field has evolved drastically, further expanding its applications in forensic linguistics through the use of sophisticated software. Forensic linguists often detect plagiarism in a questionable text either manually or through the use of artificial intelligence (AI) software embedded with forensic linguistics algorithms (Sousa-Silva, 2014). Most of the existing literature points out that plagiarism detection is one of the many variants used in authorship attribution where a particular piece of writing (questionable) is compared against other published writings, often large bodies, in efforts to find out those bearing similarities and using the outcome to predict the most probable writer of that particular text. Unlike other methods in forensic linguistics, Narayanana (2020) points out that

plagiarism detection does not make use of writing features such as cognitive but rather, follows keen arrangement and use of words as compared to other published texts.

According to Juodkaite-Granskiene et al. (2017), various approaches are applied in plagiarism detection in the contexts of forensic linguistics such as stylometric measurements involving the suspected individuals' style of writing and speaking, identifying idiolectal styles, stylistic methods, and vocabulary analytical techniques. These scholars further add that, in plagiarism detection, forensic linguists use essential elements to establish the probable author of a suspected text such as consideration of lexical density, lexical elements, lexical richness, hapax legomena, corpus reference, and hapax dis-legomena as a means of establishing the frequency (low/high) of words in a questionable text. Aspects and concepts of saliency are taken into account during the process. The use of plagiarism detection however is imperative in texts especially those that have been submitted through online platforms since computers easily detect associated similarities with regards to texts available online.

Speaker identification and voice comparison

According to research, in forensic linguistics, major tasks undertaken during speech and voice analysis in a particular questionable voice recording involve voice/speaker identification (Leemann et al., 2014; Morrison, 2018; Lutsenko & Nikulin, 2020). The process involves making voice/speaker comparisons or what scholars and practitioners in forensic linguistics term as speaker profiling (Morrison, 2018). Speaker profiling entails a thorough examination of key features of a questionable voice with a focus on containing essential elements related to the speaker's ethnic, social, and regional backgrounds. On the other side, voice comparison entails a comparative analysis of the voice in question against reference known samples of a particular speaker who is suspected to have spoken the voice in question (W, 2017). However, vocal features analysis, no matter how thorough it is, is limited since it does not provide the accurate identity of the speaker. Nevertheless, its application in forensic linguistics is imperative since it provides a rich and wealthy source of relevant information about the probable speaker thus giving the experts a high degree of confidence and precision in predicting who might have spoken the questionable voice among suspected individuals as highlighted by Sugundo et al. (2013). Several methods of speaker identification have been established and successfully contributed significantly to the identification of the speaker of a particular piece

of recording. The major methods as identified by Foulkes and French (2012) include “voice printing,” linguistic-acoustic analysis, and automatic systems analysis.

In forensic linguistics contexts, speaker identification according to Ahuja and Vyas (2016) is the analysis of a given voice to make the most relevant decision about who must have spoken. Linguists often use the known voice sample of the suspected individual with the presented sample of the voice in question. Comparisons bear aspects of similarity and dissimilarity, and once most of the features turn out to be similar, the expert decides whether the voice belongs to the suspect or not (Univaso, 2017). Thus, according to the above author, the process of speaker identification can either bear a positive/affirming result after the analysis or a negative outcome. Earlier studies on the field of forensics speaker identification relied on essential features of human voice or speech such as intonation, intensity, pitch, and amplitude (Nolan, 2012). However, over the recent past, advancements with regards to technology have been made and now speaker identification is most efficient based on technological capabilities involving making constructive comparisons of the human voice of different people. Forensic linguists have studied important aspects that help in pointing out to the specific speaker of a certain questionable voice such as

text-independent vs text-dependent, continuous speech vs single-word utterance, multi-lingual vs monolingual, spontaneous speech vs reading text, different-sex vs same-sex among others. With these significant developments in speaker identification, the concept has been heavily applied to identify speakers of questionable voices especially in court processes including cases of terrorist activities, threat calls, kidnapping, bribery among others. Speaker identification is highlighted by Sousa-Silva (2014) does not only help in identifying the main suspect in judicial contexts but also play a major role in proving the innocence of a suspected person or group of persons in certain cases. Experts sometimes acoustically analyze such questionable voices using spectrogram. Visual comparison involving graphic patterns of suspect sample and question sample plays a significant role in determining the most probable speaker of a particular voice record sample.

Language as evidence in civil cases

Trademark

One of the recent practical applications of linguistics to the decree is in brand name disputes, where the lawful matters present a high-quality fit for linguistic apparatus. The continuing hard work of The International brand name alliance demonstrates sure signs that legal practitioners are making healthy development

towards the regulation of trademark decree and guidelines all over the globe, making this a productive area for worldwide application of linguistics. This piece of writing investigates lawyers' opinions on trademarks and discusses diverse trademark classes such as essential, expressive, and indicative marks. (Olsson 2018). It examines the application of phonetics, sentence structure, semantics, or graphemes to determine trademark conflicts.

Brand name Law

Trademark law comprises the legal set of statutes that administer how firms might: recognize their goods or services in the market to avert purchaser confusion and defend the income they've selected to categorize their goods or services against use by their competitors.

The categorization of characters

The two brands implicated in the Community Trademark. DOUBLEMINT is morphologically are categorized as a complex name consisting of two words, "DOUBLE" and "MINT." DOUBLEMINT is a composite name, and descriptive name as the word "Mint" is a crop with aromatic vegetation, and one of the ingredients of DOUBLEMINT is the fragrant leave off "Mint."

Likelihood of Confusion

In the case of J. Thomas, McCarthy (2009; quoted in J. T. Berger & R. M. Halligan, 2012, p. 135) emphasizes uniqueness. It defines the term "unique" as familiar to many. Its legal nuance is dissimilar from its daily use. In trademark law, "individual" is a crucial expression. A unique description is either intrinsically or throughout the attainment of minor connotation, and then a term does not include the legal position of a "trademark" or "service mark." No uniqueness means no mark.

This section focuses on finding out the linguistic proof to decide if or not there is any probability of confusion linking the marks DOUBLEMINT and DOUBIEMINT. In stipulations of two linguistics classes believed to be appropriate from a permissible point of view, that is the class of sight (visual resemblance) and the class of connotation conceptual resemblance. The class of a picture includes all the features of the image appearance of a trademark. They have the spelling, semiotic aspects such as shade, print, blueprint, and linguistic doctrine of acknowledgment and remembrance. These features are well-thought-out for investigation and the probability of perplexity connecting DOUBLEMINT and DOUBIEMINT. Forensic Linguistics finally fulfills the qualitative analysis with the quantification of the graphemes between the higher-ranking and the low-ranking marks. Specialists establish that the brand

names DOUBLEMINT and DOUBIEMLNT demonstrate that the two graphemes associated in a similar order: D-O-U-B-L-E-M-I-N-T, D-O-U-B-I-E-M-L-N-T image of DOUBLEMINT and DOUBIEMLNT.

Defamation

In verbal communication of defamation suits, on one occasion, uses his technique of showing, through cases of his employment in real court suits, the way linguistics can add to the understanding of undecided words, in this instance, the verbal communication of defamation and smear. The focal point is slander battles and disputes in the media, medical field, and corporate. Legal suits comprise suits against the press and communication multimedia, civil disputes between competing healthcare practitioners, and claims in opposition to profit-making interests, job providers, and other stakeholders. The specialist finds that the blueprint has a similar brief outline of how the disagreement happened in approximately every instance. It contains an explanation of the verbal communication in issue, the linguistic investigation itself, and some closing comments now and then individual regarding the implications of the examination for the legal classification or the boundary of verbal

communication and the decree. (Ralarala 2019). A limited record of the heuristic lenses is occasionally assorted. The unofficial S looks at suspected offensive verbal communication comprised of linguistic and structural uncertainty, acknowledgment of intention, transmitted connotation, the supportive theory, dialogue arrangement, records of dialogue, orientation, conjecture, illocutionary, conversation acts, modulation, thematic and relevant investigation. The majority of the analysis uses different techniques, depending on the character of the interpretive assignment. A nearer glance at one of S's treatments will provide a taste of this method.

Copyright infringement

Professional forensic linguists can provide a knowledgeable ruling on the specificity, and protectability of singular terms, phrases, and trade names. Analysis of linguistic similarity to assess contravention claim shows the experts to what degree is the article by now component of the speech? If it is by now a linguistic element, linguistics indicates that if it's "indicative," as with brands similar to "Dawn" and "Joy." A lexical component is also classified if it is expressive. For example, the situation with "Simple Green" or "Ever Ready." A

forensic linguistics specialist can also confirm cases where one firm's brand is too related to another, confusing customers or consumers.

Speech act analysis

Specialists have done broad research of manifestations of harmful and hateful language activities in impulsive spoken or printed dialogs. It connects to the procedure of data cover-up and fabrication in law enforcement interviews and law court evidence. Forensic linguistics uses several investigative techniques to produce this document: a demonstration examination of technical writing, proportional lawful and rational investigation, extrapolation techniques, and substance investigation. (Coulthard, 2017) Despite much investigational research dedicated to audio-phonetic or psycho-linguistic aspects of deceit, their outcome is not adequately dependable for forensic use. The specialist account should not rely on a hypothesis. The writer negates such the proof acceptability of linguistics. The linguistics conclusion concerning utterances implies verbal communication indicators related to lies noted via psycho-linguistic assessment in spoken language acoustic or film copy of a law enforcement meeting or court evidence. Forensic linguistics applies psycholinguistic techniques and algorithms to sense language symbols of deceit. They reveal an enormous

assortment that contradicts the cardinal proof acceptability. The inadequate progress of the user in the expert method and the deficit of a united method for resolving specialist tasks on a firm scientific foundation creates a need for increasing inclusive techniques for investigating lies based on forensic speech science and mental conjecture.

Forensic linguistics perform tests to help recognize suspects or case testimony. The specialists determine the implication of written words to a particular lawsuit. Verbal or printed specimens specimen are brought as confirmation in a law court, together with the evidence of the linguistic specialist. (Olsson, 2019). Speech confirmation may weigh a great deal on the lawsuit case itself. In the case where the words in inquiry comprise a speech offense: coercion, intimidation, inducement, hatred dialogue, and hatred writing, or speech proof are more related to a lawsuit, and it might require a professional linguist to elucidate the connotation of the is spoken or printed words, the approach in it was presented, and the function of circumstance in the understanding of the memorandum.

References

- Ahuja, P., & Vyas, J. M. (2016). Forensic speaker profiling: the study of supra-segmental features of Gujarati dialects for text-independent speaker identification. *Australian Journal of Forensic Sciences*, 50(2), 152–165. <https://doi.org/10.1080/00450618.2016.1237547>
- Ariani, M. G., Sajedi, F., & Sajedi, M. (2014). Forensic Linguistics: A Brief Overview of the Key Elements. *Procedia - Social and Behavioral Sciences*, 158, 222–225. <https://doi.org/10.1016/j.sbspro.2014.12.078>
- Correa, M. (2013). Forensic Linguistics: An Overview of the Intersection and Interaction of Language and Law. *Studies about Languages*, 0(23). <https://doi.org/10.5755/j01.sal.0.23.5020>
- Coulthard M. (2017). *An Introduction to Forensic Linguistics: Language in Evidence*. Routledge
- Foulkes, P., & French, P. (2012). Forensic Speaker Comparison: A Linguistic–Acoustic Perspective. In *Oxford Handbooks Online*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199572120.013.0041>
- Juodkaitė - Granskienė, G., Herasimenkienė, G., Žalkauskienė, A., & Jurka, R. (2017). The problem of Definition of Plagiarism and Features of Usage

- of Other Texts Identified in Forensic Linguistics Examinations. *Pedagogika*, 126(2), 202–213. <https://doi.org/10.15823/p.2017.29>
- Krassa, S. I. (2020). First roundtable on practices and standards in forensic authorship analysis (overview 1). *Политическая лингвистика*, 6, 174–187. <https://doi.org/10.26170/pl20-06-19>
- Krassa, S. I. (2021). First roundtable on practices and standards in forensic authorship analysis (Overview 4). *Политическая лингвистика*, 3, 227–234. https://doi.org/10.26170/1999-2629_2021_03_24
- Leemann, A., Kolly, M.-J., & Dellwo, V. (2014). Speaker-individuality in suprasegmental temporal features: Implications for forensic voice comparison. *Forensic Science International*, 238, 59–67. <https://doi.org/10.1016/j.forsciint.2014.02.019>
- Lutsenko, K., & Nikulin, K. (2020). Voice speaker identification is one of the current biometric methods of identification of a person. *Theory and Practice of Forensic Science and Criminalistics*, 19(1), 239–255. <https://doi.org/10.32353/khrife.1.2019.18>
- Morrison, G. S. (2018). The impact in forensic voice comparison of lack of calibration and mismatched conditions between the known-speaker recording and the relevant-population sample recordings. *Forensic*

Science International, 283, e1–e7.

<https://doi.org/10.1016/j.forsciint.2017.12.024>

Narayanan, A. S. (2020). A Survey on Plagiarism Detection Techniques.

International Journal of Psychosocial Rehabilitation, 24(1), 1564–1572.

<https://doi.org/10.37200/ijpr/v24i1/pr200254>

Nolan, F. (2012). Intonation in speaker identification: an experiment on pitch alignment features. *Forensic Linguistics*, 9(1), 1–21.

<https://doi.org/10.1558/sll.2002.9.1.1>

Olsson J. (2018). *Forensic linguistics*. A&C Black. Language Arts & Discipline.

Olsson J. (2019). *SOLVING Crime Through Forensic Linguistics*. A&C Black

R Gupta, R. (2019). Forensic analysis of characteristic features of disguised writing in fixing authorship. *IP International Journal of Forensic Medicine and Toxicological Sciences*, 4(1), 22–27.

<https://doi.org/10.18231/j.ijfmts.2019.006>

Ralarala M. (2019). *New Frontiers in Forensic Linguistics: Themes and Perspectives in Language and Law in Africa and Beyond*. African Sun Media.

Segundo, E. S., Alves, H., & Trinidad, M. F. (2013). CIVIL Corpus: Voice Quality for Speaker Forensic Comparison. *Procedia - Social and Behavioral Sciences*, 95, 587–593.

<https://doi.org/10.1016/j.sbspro.2013.10.686>

Sousa-Silva, R. (2014). Investigating academic plagiarism: A forensic linguistics approach to plagiarism detection. *International Journal for Educational Integrity*, 10(1). <https://doi.org/10.21913/IJEI.v10i1.932>

Univaso, P. (2017). Forensic Speaker Identification: a tutorial. *IEEE Latin America Transactions*, 15(9), 1754–1770.

<https://doi.org/10.1109/tla.2017.8015083>

W, H. (2017). Quantitative Comparison Method of Forensic Voice Evidence. *International Journal of Forensic Sciences*, 2(2).

<https://doi.org/10.23880/ijfsc-16000127>