IJFL (International Journal of Forensic Linguistic)
Vol. 5, No. 2 October 2025, Page 36-42

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

Analysis of Guilty Plea and Lie Detection through Linguistic Markers: A Literature Review

Candiawan Telaumbanua

Bali Business School, Denpasar, Indonesia
canditel@gmail.com

How to cite (in APA style):

Telaumbanua, Candiawan. (2025). Analysis of Guilty Plea and Lie Detection through Linguistic Markers: A Literature Review. *IJFL (International Journal of Forensic Linguistic)*, 5(2), 36-42.

Abstract-This literature review explores the relationship between guilty pleas, false confessions, and lie detection through linguistic markers within modern judicial systems. Although guilty pleas function as an efficient mechanism to expedite legal proceedings and reduce the burden on courts, research shows that they do not always represent genuine admissions of guilt. False guilty pleas, often influenced by coercion, plea bargaining pressures, or fear of harsher sentences, continue to be a major source of wrongful convictions. Because language serves as the primary medium through which guilt is expressed and legally constructed, it plays a crucial role in assessing the authenticity and voluntariness of such pleas. The review synthesizes key findings from forensic linguistics, psychology, and computational linguistics on linguistic indicators of deception, including lexical, syntactic, discourse, and paralinguistic features. Common linguistic patterns associated with deception include reduced self-references, avoidance of agentive expressions, limited sensory details, and inconsistent narratives. Recent advances in Natural Language Processing (NLP) have further expanded the potential for automated lie detection, though challenges related to cultural variability, accuracy, and ethical concerns remain. In conclusion, integrating linguistic analysis into judicial practice can enhance the evaluation of guilty pleas, reduce the risk of wrongful convictions, and strengthen substantive justice. Future research should focus on cross-linguistic validation, the use of authentic courtroom data, and the development of ethical guidelines to ensure that linguistic evidence is applied fairly and responsibly.

Keywords: Guilty Plea, False Confession, Forensic Linguistics, Linguistic Markers, Lie Detection, Courtroom Discourse.

I. INTRODUCTION

In the modern criminal justice system, guilty plea or confession of guilt is one of the main instruments in case resolution. In various jurisdictions, especially legal systems that adhere to adversarial traditions such as the United States and the United Kingdom, more than 90% of criminal cases are resolved through guilty plea, rather than through a full court hearing (Bibas, 2003). This efficiency is what makes this mechanism popular, as it is able to reduce the burden of cases, speed up the judicial process, and reduce state costs. Behind these pragmatic benefits, however, lies a fundamental question: does each guilty plea really reflect the sincere confession of a guilty defendant, or is it a manifestation of pressure, manipulation, and systemic injustice?

The phenomenon of false guilty plea has become an important concern in the last two decades. A report from the Innocence Project shows that a large number of exoneration cases, including those supported by DNA evidence, turn out to come from individuals who previously pleaded guilty through a guilty plea (Cooper et al., 2019). Psychological pressure, disparity in sentencing between court and plea

negotiations, lack of legal understanding, and incorrect defense strategies often force defendants, even innocent ones, to choose to plead guilty in order to avoid the risk of heavier punishment. Thus, a guilty plea cannot always be understood as an authentic representation of substantive wrongdoing, but rather as a compromise within an institutional framework.

It is in this context that language plays a central role. Guilty plea is not only a legal procedure, but also a performative speech that linguistically shapes legal reality: a person who says "I am guilty" officially changes his status from defendant to convict. The language used in confession, both oral and written, is not just a medium, but also evidence that has a direct impact on a person's legal fate. Therefore, linguistic analysis of guilty pleas is significant to assess the authenticity, voluntariness, and potential pressure in the statement.

Research in forensic linguistics shows that human language holds psychological imprints that can be indicators of honesty or falsehood. Linguistic characteristics such as word choice, syntactic patterns, use of pronouns, narrative details, to intonation and pauses in speech have been identified as linguistic markers related to lying or distress (DePaulo et al., 2003; Vrij, 2014). For example, coerced confessions tend to use shorter sentences, full of repetition, avoid sensory details, or use passive forms to obscure the perpetrator's agent. In contrast, authentic confessions usually have a more consistent, detailed, and coherent narrative.

In addition, the development of Natural Language Processing (NLP) technology opens up new avenues in language-based lie detection. Text analysis models, such as LIWC (Linguistic Inquiry and Word Count) as well as machine learning algorithms, have been able to identify linguistic patterns that distinguish honest statements from deceptive ones with a high level of accuracy in experimental data (Kleinberg et al., 2018). However, the challenge that remains is how to apply the method in a real-world context, such as a court hearing or guilty plea transcript, which is much more complex than laboratory experiments.

Furthermore, cross-linguistic studies show that indicators of lies are not always universal. An effective marker in English is not necessarily valid in Indonesian or other languages that have a different grammatical structure (Velutharambath et al., 2025). This emphasizes the need for contextual research that considers aspects of culture, language structure, and legal practice in each country.

Thus, the background of this research departs from three main issues. First, guilty plea is an important mechanism but prone to abuse, so an independent evaluation tool is needed to test its authenticity. Second, the development of theories and methods in forensic linguistics opens up opportunities to detect lies through linguistic analysis, including guilty plea statements. Third, there is still a gap in research that directly links guilty pleas and linguistic markers of lies, both at the theoretical level and in practical application.

Based on this framework, this literature review aims to critically examine the relationship between guilty plea, false confession, and linguistic-based lie detection. Her main focus is to identify relevant linguistic markers, evaluate empirical findings, and discuss their implications for legal practice and wrongful convictions prevention efforts in the criminal justice system.

II. METHODS

This study uses a qualitative approach with a literature review design. The selection of this method is based on the purpose of the research that aims to comprehensively examine the development of theories, methods, and empirical findings regarding the relationship between guilty plea, false confession, and lie detection through linguistic markers. The literature review approach is considered appropriate because it is able to summarize, synthesize, and provide a critical evaluation of previous studies from the disciplines of linguistic forensics, legal psychology, and computational linguistics. Thus, this method allows researchers to see the big picture of the extent to which linguistic analysis has played a role in uncovering the dynamics of guilt and detecting lies in court.

The data sources in this study include scientific publications in the form of reputable international journal articles, empirical research reports, and academic books that discuss the topics of forensic linguistics, lie detection, and plea bargaining. In addition, this study also refers to secondary data such as reports from independent institutions, such as the Innocence Project, as well as court decisions or exoneration documentation that are relevant to the phenomenon of false guilty plea. The data is collected through systematic searches of various online databases, including JSTOR, Scopus, Taylor & Francis, SpringerLink, and open

repositories such as arXiv. The search process was carried out with keywords such as guilty plea, false confession, forensic linguistics, linguistic markers of deception, courtroom discourse, and lie detection NLP. The articles obtained were then selected based on inclusion criteria, namely relevance to the research theme, containing empirical data or theoretical studies with forensic nuances, and published in the last two decades.

Data analysis was carried out in three stages. First, a content analysis is carried out to identify the main focuses, methods, and findings of each publication. Second, a thematic synthesis was carried out to group the findings into major themes, such as the characteristics of guilty plea and false confession, linguistic markers of lies, computational approaches in lie detection, and practical implications for the legal system. Third, this study conducts a critical evaluation of the strengths, limitations, and research gaps found, to be then linked to practical needs in the criminal justice system.

The theoretical framework used in this study integrates three main perspectives. First, forensic linguistics is used to study language structure, lexical choices, rhetorical strategies, and speech function in guilty plea. Second, discourse analysis is used to look at how guilt narratives are constructed, negotiated, or even forced through language. Third, psychological and computational models of lies are used to understand the cognitive and linguistic underpinnings of deceptive behavior, including the potential for automating lie detection with Natural Language Processing (NLP) technology.

To maintain the validity of the research, theoretical triangulation is used by combining linguistic perspectives, legal psychology, and normative legal studies. Reliability is maintained through a strict selection of sources derived from peer-reviewed scientific publications and transparency in the preparation of thematic categories. With the design of this method, the research is expected to provide a picture that is not only descriptive, but also critical and reflective, thus producing theoretical and practical contributions to the development of forensic linguistics in analyzing guilty pleas and detecting language-based lies.

III. RESULT AND DISCUSSION

The manuscript of the chosen article is presented below, together with a detailed explanation of the findings;

Table 1. Overview of Research on Guilty Pleas, Wrongful Convictions, and Linguistic Markers of Deception

No.	Author & Year	Title / Source	Research Focus	Key Findings
1.	Bibas (2003)	Plea Bargaining Outside the Shadow of Trial, Harvard Law Review	Analysis of the role of guilty plea in the criminal justice system in the US	Showing that more than 90% of criminal cases are resolved by plea bargain; efficiency is often achieved at the expense of substantive justice.
2	Redlich et al. (2018)	The Influence of Confessions on Guilty Pleas and Plea Discounts, Psychology, Public Policy, and Law	The relationship between confession and acceptance of guilty plea	Defendants who confess (rightly or wrongly) are more likely to accept a plea bargain; Confession increases the chances of plea by twofold.
3	Cooper et al. (2019)	Innocents Who Plead Guilty, Federal Sentencing Reporter	DNA case exculpation report	About 11% of DNA exculpation cases in the U.S. involve defendants who previously pleaded guilty; a lot due to systemic pressure.
4	Young (2012)	Conconviting the Innocent: Where Criminal Prosecutions Go Wrong	Analysis of wrongful convictions in the US	It revealed that trial penalties encourage many innocent defendants to choose guilty pleas to

				avoid harsher punishments.
5	Villar et al. (2013)	Linguistic Indicators of a False Confession, Psychiatry, Psychology and Law	Linguistic analysis of false confessions	False confession is characterized by contradictory narratives, excessive repetition, and mitigation of responsibility.
6	DePaulo et al. (2003)	Cues to Deception, Psychological Bulletin	Meta-analysis of linguistic markers of lies	Deception is characterized by fewer *self-references*, more negations, tentative use of language, and shorter narratives.
7	Newman et al. (2003)	Lying Words: Predicting Deception from Linguistic Styles, PSPB	Lie analysis with LIWC	Liars use passive sentences more, avoid agents, and reduce narrative details.
8	Vrij (2014)	Detecting Lies and Deceit: Pitfalls and Opportunities	Theoretical study of lie detection	Lies are characterized by a lack of sensory detail, chronological inconsistencies, and the use of defensive rhetoric.
9	Loconte & Kleinberg (2025)	Automated verbal deception detection for embedded lies	NLP for lie detection	ML algorithms achieve >75% accuracy in detecting lies, but real-world validity is still low.
10	Velutharambath et al. (2025)	What if Deception Cannot be Detected? A Cross-Linguistic Study on the Limits of Deception Detection from Text	Cross-language studies	The accuracy of the model drops significantly when applied to different languages; Lie markers are not universal.
11	Hermawan et al. (21)	Counter-Claiming for a Crime Narrative, Indonesian Journal of Applied Linguistics	Analysis of pleas in Indonesian corruption cases	The defendant uses linguistic strategies in the form of mitigation, reframing, and agent avoidance to weaken the prosecutor's narrative.
12	Hancock et al. (2007)	On Lying and Being Lied To, Discourse Processes	Analysis of lies in online communication	Lying is characterized by more pauses, typos, and the use of *hedges* that show cognitive distress.

The results of this literature review show that the relationship between guilty plea, false confession, and linguistic-based lie detection can be mapped into several major themes: (1) the role of guilty plea in the modern legal system, (2) the phenomenon of false guilty plea and its driving factors, (3) linguistic markers as indicators of lies, (4) computational approaches in lie detection, (5) linguistic analysis of plea statements, and (6) practical implications for the judicial system.

Guilty Pleas in Legal Systems

Preliminary studies show that guilty plea has become the dominant mechanism in the settlement of criminal cases in various countries. In the United States, for example, more than 90% of criminal cases are resolved through plea bargains, not through a full trial (Bibas, 2004). It is this efficiency that makes it popular, but it is often achieved at the expense of transparency and substantive justice. Redlich, Yan, Norris, and Bushway (2018) in Psychology, Public Policy, and Law found that defendants who gave confessions, either true or false, were much

more likely to accept plea bargains than those who did not confess. This shows that confessions and plea decisions affect each other, and often the defendant feels that they have no choice but to plead guilty.

False Guilty Pleas and Wrongful Convictions

One of the most consistent findings in the literature is the existence of cases in which innocent individuals have instead chosen a guilty plea. The Innocence Project report (2019) noted that about 11% of DNA exculpation cases in the United States involve defendants who previously pleaded guilty. Garrett (2016) in his book Convicting the Innocent explains that this phenomenon arises because of the trial penalty, which is the risk of a much heavier sentence if the defendant chooses to be tried rather than accept a plea bargain. Villar, Arciuli, and Paterson (2012) found that false confessional statements are often characterized by contradictory narratives, excessive use of repetition, and emotional pressure in language styles. These findings strengthen the argument that linguistic analysis can be used to identify inauthentic confessions or guilty pleas.

Linguistic Markers of Deception

The literature examining lie detection through language found a number of linguistic markers that consistently appear in deceptive statements. DePaulo et al. (2003) in the Psychological Bulletin identified that people who lie tend to use fewer self-references, more negation words, and more tentative language such as "maybe" or "likely." Newman et al. (2003) support these findings by showing that liers often avoid using agentive language (e.g., avoiding the subject "I") and more often use passive sentences to reduce responsibility. Vrij (2008) adds that narrative lies usually have fewer sensory details and are less consistent in the sequence of events. Prosodic markers are also relevant: Hancock et al. (2007) observed that in online communication, fraudsters are more likely to show long pauses, typos, and the use of hedges. All of this is evidence that language holds cognitive cues about pressure and cheating intentions.

Computational Approaches in Lie Detection

The development of Natural Language Processing (NLP) technology has opened up new opportunities for automatic-based lie detection. Kleinberg et al. (2018) introduced The Lying Dutchman dataset and found that machine learning algorithms were able to identify false statements with an accuracy of more than 75%. Velutharambath, Klinger, and Sassenberg (2025) tested cross-linguistic validity and found that the model's accuracy decreased drastically when applied to languages other than the language of practice, confirming that lie markers are culture-specific. This is important in the context of guilty plea, because judicial practices are multilingual and multicultural.

Linguistic Analysis of Plea Statements

Although specific research on linguistic analysis of guilty pleas is limited, some early studies point in a promising direction. Hermawan, Rahyono, and Dallyono (2022) used Appraisal theory to analyze pleas in corruption cases in Indonesia. They found that defendants often used linguistic strategies to reframe the prosecutor's narrative, for example by emphasizing external factors or avoiding direct attribution to oneself. Villar et al. (2012) also emphasized that false guilty pleas have a more defensive language style, contain many mitigation strategies, and are less coherent than authentic confessions. This suggests that linguistic analysis can help identify guilty pleas that are insincere or that arise as a result of the pressure of the legal system.

Implications for Legal Practice

The results of this review have important implications for legal practice. First, linguistic analysis can assist judges and prosecutors in evaluating whether a guilty plea is made voluntarily or under duress. Second, linguistic expert witnesses can contribute to providing an objective assessment of the defendant's language style, especially in cases where there are false confession claims. Third, NLP technology can be used to analyze large amounts of plea transcripts, thus serving as a supporting tool to identify potential false guilty pleas. However, the limitations of accuracy and the risk of bias should still be considered, so the results of linguistic analysis should be used as a complement, not a replacement, to traditional legal judgments.

Research Gaps and Challenges

Although the research shows promising results, there are still a number of research gaps. First, research that directly analyzes guilty pleas in real context is still very limited, so an authentic data corpus from the court is needed. Second, most studies are focused on the English language, so cross-lingual and cross-cultural research is urgently needed. Third, the existence of contextual factors such as pressure from legal counsel, the psychological condition of the defendant, and legal negotiation strategies need to be taken into account in linguistic analysis. Fourth, the integration of computational technology raises ethical questions, including privacy, accountability, and potential misuse of the analysis results.

IV. CONCLUSION

This literature review confirms that the guilty plea is one of the most influential instruments in the modern criminal justice system, but its existence preserves a paradox between efficiency and substantive justice. On the one hand, this mechanism is able to speed up the settlement of cases, reduce the burden on the court, and reduce litigation costs. On the other hand, empirical findings show that guilty pleas do not always reflect an authentic confession of

guilt, but are often influenced by external factors such as trial penalties, psychological pressures, and manipulative legal strategies. The phenomenon of false guilty pleas revealed through the Innocence Project report and other research confirms that the risk of wrongful conviction remains high even if the defendant formally pleads guilty.

In this context, linguistic analysis offers an important contribution. Linguistic markers, whether in the form of lexical choices, syntactic patterns, narrative details, or discourse styles, have been shown to contain cognitive and psychological traces that can be used to detect lies. Research by DePaulo et al. (2003), Vrij (2008), Newman et al. (2003), as well as contemporary studies such as Kleinberg et al. (2018) show the consistency of the findings that lies have identifiable linguistic characteristics. The development of NLP technology is further expanding the scope of lie detection, although challenges related to cross-language accuracy and algorithmic bias still need to be addressed.

In addition, recent research in Indonesia (e.g. Hermawan et al., 2022) shows that guilty plea also has a local dimension that is important to be examined, because the defendant's linguistic strategy is influenced by certain cultures, legal norms, and social contexts. Therefore, cross-linguistic and cross-cultural analysis is crucial to avoid bias in the application of linguistic markers.

Thus, it can be concluded that an interdisciplinary approach that integrates forensic linguistics, legal psychology, and computational technology has great potential to improve fairness in the practice of guilty pleas. However, its implementation must be accompanied by a clear ethical framework, the involvement of linguistic expert witnesses, as well as further research based on authentic data from court practice. These efforts are expected not only to prevent wrongful convictions, but also to strengthen the principle of substantive justice in the criminal justice system in various jurisdictions.

V. REFERENCES

Bibas, S. (2003). Plea Bargaining Outside the Shadow of Trial. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.464880

Cooper, G. S., Meterko, V., & Gadtaula, P. (2019). Innocents Who Plead Guilty: An Analysis of Patterns in DNA Exoneration Cases. Federal Sentencing Reporter, 31(4–5), 234–238. https://doi.org/10.1525/fsr.2019.31.4-5.234

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. Psychological Bulletin, 129(1), 74–118. https://doi.org/10.1037/0033-2909.129.1.74

Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2007). On Lying and Being Lied To: A Linguistic Analysis of Deception in Computer-Mediated Communication. Discourse Processes, 45(1), 1–23. https://doi.org/10.1080/01638530701739181

Hermawan, R., Rahyono, F. X., & Dallyono, R. (2021). COUNTER-CLAIMING FOR A CRIME NARRATIVE: AN EVALUATION OF THE DEFENDANT'S PLEA AT THE CORRUPTION CRIMINAL COURT. Indonesian Journal of Applied Linguistics, 11(1). https://doi.org/10.17509/ijal.v11i1.34669

Loconte, R., & Kleinberg, B. (2025). When lies are mostly truthful: Automated verbal deception detection for embedded lies (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2501.07217

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying Words: Predicting Deception from Linguistic Styles. Personality and Social Psychology Bulletin, 29(5), 665–675. https://doi.org/10.1177/0146167203029005010

Redlich, A. D., Yan, S., Norris, R. J., & Bushway, S. D. (2018). The influence of confessions on guilty pleas and plea discounts. Psychology, Public Policy, and Law, 24(2), 147–157. https://doi.org/10.1037/law0000144

Velutharambath, A., Sassenberg, K., & Klinger, R. (2025). What if Deception Cannot be Detected? A Cross-Linguistic Study on the Limits of Deception Detection from Text (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2505.13147

Villar, G., Arciuli, J., & Paterson, H. (2013). Linguistic Indicators of a False Confession. Psychiatry, Psychology and Law, 20(4), 504–518. https://doi.org/10.1080/13218719.2012.712834

Vrij, A. (2014). 14. Detecting lies and deceit: Pitfalls and opportunities in nonverbal and verbal lie detection. In C. R. Berger (Ed.), Interpersonal Communication (pp. 321–346). DE GRUYTER. https://doi.org/10.1515/9783110276794.321

Young, A. (2012). Book Review: Convicting the Innocent: Where Criminal Prosecutions Go Wrong, by Brandon L. Garrett. Osgoode Hall Law Journal, 50(2), 491–498. https://doi.org/10.60082/2817-5069.1030