Philology Matters / ISSN: 1994 – 4233 / E-ISSN: 2181–1237
Uzbekistan State World Languages University
2026 Vol. 1 No. M 4
LINGUISTICS
DOI: 10.36078/987656202

Moxlaroyim Tursunova

Uzbekistan State World Languages University
Tashkent, Uzbekistan
E-mail: moxlaroyimzaynobiddinova83@gmail.com

Scientific Supervisor: Shahodat Usmonova – PhD

THREAT OR HYPERBOLE? LINGUISTIC INDICATORS FOR READING THREATENING LANGUAGE ON SOCIAL MEDIA

Annotation. Distinguishing threats from hyperbole in social media communication poses significant linguistic and legal challenges, because violent wording may express frustration, humour or artistic play as readily as genuine menace. This article examines how context, communicative purpose and speaker awareness shape the interpretation of potentially threatening online statements. Using a qualitative, document-based case-study design, it analyses three judgments: Chambers v Director of Public Prosecutions (2012), Elonis v. United States (2015) and Counterman v. Colorado (2023). The judgments are read closely and interpreted through speech act theory and pragmatics, supported by research in forensic linguistics. The analysis yields seven indicators: target and addressee, speech-act function, genre and communicative frame, hyperbole and specificity, audience and context, repetition, and recipient response. The results show that the courts did not treat any single word or grammatical form as decisive; they weighed circumstances, communication patterns and, increasingly, the speaker's mental state. The article concludes that linguistic evidence can clarify wording, pragmatic meaning and ambiguity but cannot establish private intention or replace legal assessment. The framework is an analytical synthesis, not a validated predictive model, and requires testing on larger datasets.

Keywords: forensic linguistics, social media, threats, hyperbole, speech acts, pragmatics, context

Moxlaroyim Tursunova

O‘zbekiston davlat jahon tillari universiteti
Toshkent, O‘zbekiston

Ilmiy rahbar: Shahodat Usmonova – PhD

TAHDIDMI YOKI MUBOLAG‘AMI? IJTIMOIY TARMOQLARDAGI TAHDIDLI NUTQNI TALQIN QILISH UCHUN LINGVISTIK INDIKATORLAR

Annotatsiya. Ijtimoiy tarmoqlardagi xabarlarda tahdid va mubolag‘ani farqlash lingvistik hamda huquqiy jihatdan murakkab masaladir, chunki zo‘ravonlikka oid ifodalar haqiqiy tahdidni ham, jahl, hazil yoki badiiy o‘yinni ham ifodalashi mumkin. Mazkur maqolada kontekst, kommunikativ maqsad va so‘zlovchining xabardorligi potensial tahdidli onlayn bayonotlar talqiniga qanday ta’sir ko‘rsatishi o‘rganiladi. Sifat tahliliga asoslangan hujjatlar tahlili (case study) dizayni asosida uchta sud qarori – Chambers v Director of Public Prosecutions (2012), Elonis v. United States (2015) va Counterman v. Colorado (2023) tahlil qilinadi. Qarorlar nutqiy aktlar nazariyasi va

pragmatika asosida, sud lingvistikasi tadqiqotlariga tayanib sinchiklab o‘qiladi. Tahlil natijasida yetti indikator ajratiladi: maqsad va adresat, nutqiy akt vazifasi, janr va kommunikativ freym, mubolag‘a va aniqlik, auditoriya va kontekst, takrorlanish hamda qabul qiluvchining munosabati. Natijalar sudlar hech bir alohida so‘z yoki grammatik shaklni hal qiluvchi deb hisoblamaganini, balki holatlarni, xabarlar izchilligini va tobora ko‘proq so‘zlovchining ruhiy holatini baholaganini ko‘rsatadi. Xulosaga ko‘ra, lingvistik dalillar so‘z qo‘llanishi, pragmatik ma’no va noaniqlikni yoritishi mumkin, biroq shaxsning ichki niyatini aniqlay olmaydi va huquqiy bahoni almashtirmaydi. Taklif etilgan tizim tasdiqlangan bashoratli model emas, tahliliy sintezdir va kattaroq ma’lumotlar to‘plamida sinovdan o‘tkazilishi lozim.

Kalit so‘zlar: sud lingvistikasi, ijtimoiy tarmoqlar, tahdid, mubolag‘a, nutqiy aktlar, pragmatika, kontekst

Мохларойим Турсунова

Узбекский государственный университет мировых языков
Ташкент, Узбекистан

Научный руководитель: Шаохдат Усмонова – PhD

УГРОЗА ИЛИ ГИПЕРБОЛА? ЛИНГВИСТИЧЕСКИЕ ИНДИКАТОРЫ ДЛЯ ИНТЕРПРЕТАЦИИ УГРОЖАЮЩЕЙ РЕЧИ В СОЦИАЛЬНЫХ СЕТЯХ

Аннотация. Разграничение угроз и гиперболы в сообщениях социальных сетей представляет собой сложную лингвистическую и правовую задачу, поскольку насильственная лексика может выражать как реальную угрозу, так и раздражение, юмор или художественную игру. В статье исследуется, как контекст, коммуникативная цель и осведомлённость говорящего влияют на интерпретацию потенциально угрожающих онлайн-высказываний. На основе качественного документального кейс-стади анализируются три судебных решения: Chambers v Director of Public Prosecutions (2012), Elonis v. United States (2015) и Counterman v. Colorado (2023). Решения подвергаются пристальному прочтению с опорой на теорию речевых актов и прагматику, а также на исследования в области судебной лингвистики. В результате анализа выделяются семь индикаторов: адресат и объект высказывания, функция речевого акта, жанр и коммуникативная рамка, гипербола и конкретность, аудитория и контекст, повторяемость и реакция получателя. Результаты показывают, что суды не считали решающим ни отдельное слово, ни грамматическую форму: оценивались обстоятельства, структура коммуникации и, всё чаще, психическое отношение говорящего. Делается вывод, что лингвистические данные способны прояснить формулировки, прагматическое значение и неоднозначность, но не могут установить скрытое намерение или заменить правовую оценку. Предложенная схема является аналитическим синтезом, а не валидированной прогностической моделью, и требует проверки на более крупных массивах данных.

Ключевые слова: судебная лингвистика, социальные сети, угроза, гипербола, речевые акты, прагматика, контекст

INTRODUCTION

Social media has transformed public communication by allowing individuals to express opinions, emotions and disagreement before large and partly unknown audiences [boyd, 2014; Herring, 2013]. Because posts are short, informal and detached from the setting in which they were written, readers often lack the cues available in face-to-face interaction, a problem described as context collapse [Marwick & boyd, 2011]. Statements that look menacing when read literally may therefore operate as jokes, exaggerations or emotional reactions. Exaggeration itself is a routine resource of everyday talk rather than an ex-

ceptional one [Claridge, 2011; McCarthy & Carter, 2004].

These ambiguities matter for forensic linguistics, which studies language in legal and investigative settings [Coulthard et al., 2017; Olsson & Luchjenbroers, 2014]. Early work showed how linguistic evidence can be used and misused in courtrooms [Shuy, 1993; Solan & Tiersma, 2005], and legal language follows conventions of its own for assigning meaning to utterances [Tiersma, 1999]. Research on threats has described their pragmatic character [Fraser, 1998; Storey, 1995], compared the stance of realised and non-realised threats in corpora [Gales, 2011; 2015], and examined threatening in English through mixed methods [Muschalik, 2018]. Studies of Twitter have linked abusive and threatening language to group identity and online conflict [Hardaker & McGlashan, 2016]. Behavioural threat assessment offers a complementary view, treating communication as one of several warning behaviours [Meloy & O'Toole, 2011; Meloy et al., 2012]. In computational work, systems have been built to detect hate speech and abusive language, yet they struggle to separate hateful from merely offensive language and can reproduce bias [Davidson et al., 2017; Schmidt & Wiegand, 2017; Sap et al., 2019].

The theoretical starting point is that a threat is not identified by grammatical form. A declarative sentence may threaten, while a sentence full of violent vocabulary may joke, warn or complain. Speech act theory explains that utterances perform actions [Austin, 1962; Searle, 1969], and Searle's [1976] classification of illocutionary acts, which distinguishes acts by their point, offers a principled basis for separating a threat from a warning, complaint or insult. Pragmatics adds that communicated meaning exceeds literal content [Grice, 1975] and depends on what hearers can infer from context [Sperber & Wilson, 1986]. Communicative frames such as humour, fiction or confrontation guide this inference [Goffman, 1974], and threats must also be separated from face-threatening acts and impoliteness, which cause offence without promising harm [Brown & Levinson, 1987; Culpeper, 2011]. Hyperbole is deliberate exaggeration for emphasis or emotional effect. Systematic procedures for identifying it in discourse have been proposed [Burgers et al., 2016], but deciding whether exaggeration amounts to a threat still depends on function, specificity, circumstances and the relationship between speaker and recipient.

The legal background has developed in parallel. In the United States, an early decision treated a conditional remark made at a political gathering as hyperbole rather than a true threat [Watts v. United States, 1969], and the Supreme Court later described true threats as statements by which the speaker means to communicate a serious intention to commit unlawful violence [Virginia v. Black, 2003]. Three recent judgments extend this line of thought to online communication. *Chambers v Director of Public Prosecutions* concerned a tweet about an airport [Chambers v Director of Public Prosecutions, 2012]; *Elonis v. United States* concerned Facebook posts written in the style of rap lyrics [Elonis v. United States, 2015]; and *Counterman v. Colorado* concerned hundreds of unwanted messages sent to one person [Counterman v. Colorado, 2023]. To the author's knowledge, few linguistic studies bring these decisions together to show how courts treat wording, context and speaker awareness at once. This is the research gap addressed here.

The aim of this article is to examine the three judgments from a forensic linguistic perspective and to derive indicators that support the analysis of potentially threatening language on social media. The objectives are: (1) to describe how each court treated wording, context and mental state; (2) to organise the observations into seven analytical indicators; and (3) to specify the limits of linguistic evidence. The central research question is: how can linguistic and contextual indicators help distinguish threats from hyperbole while recognising the limits of linguistic evidence in legal interpretation? Because the study is qualitative and interpretive, no hypothesis is tested.

METHODS

This study uses a qualitative, document-based legal case-study design. It involves no human participants and collects no new empirical data. Its object is the way courts interpret potentially threatening online communication, and its subject is the judicial reasoning on wording, context and mental state.

The primary materials are the published judgments in *Chambers v Director of Public Prosecutions* [2012] EWHC 2157 (Admin), *Elo-*

nis v. United States, 575 U.S. 723 (2015), and *Counterman v. Colorado*, 600 U.S. 66 (2023), together with the provisions applied in them: section 127 of the Communications Act 2003 [Communications Act, 2003] and 18 U.S.C. § 875(c) [18 U.S.C. § 875(c)]. Supporting authorities were consulted for legal background: *Watts*, *Virginia v. Black*, *Director of Public Prosecutions v Collins* [2006] UKHL 40 [Director of Public Prosecutions v Collins, 2006], and prosecution guidance on social media communications in England and Wales [Crown Prosecution Service, 2018].

The cases were selected purposively because they (a) concern statements published or sent through social media platforms; (b) are decisions of appellate or supreme courts with full published reasoning; (c) address different issues, namely contextual interpretation, a statutory mental state and a constitutional mental state; and (d) are available in full text. The selection makes no claim to be representative.

Each judgment was read closely. For every case, the disputed wording, the platform, the circumstances of publication, the competing interpretations, the legal question, and the court's reasoning and outcome were extracted. These data were coded against seven indicators drawn from speech act theory [Austin, 1962; Searle, 1969; 1976], pragmatics [Grice, 1975; Sperber & Wilson, 1986], frame analysis [Goffman, 1974], research on threats [Fraser, 1998; Gales, 2011] and work on online context [Marwick & boyd, 2011]. The indicators are: (1) target and addressee; (2) speech-act form and function; (3) genre and communicative frame; (4) hyperbole and specificity; (5) audience and context; (6) repetition and communication pattern; and (7) recipient response and uptake. Each is phrased as an analytical question, such as who is addressed or identifiable. The indicators organise linguistic observation and do not constitute an independent legal test of criminal liability.

Reliability and validity were addressed in three ways. The analysis relies only on public documents that readers can check. Observations about wording and context were kept separate from conclusions about intention, which only the courts can draw. Findings were compared with the secondary literature. Because a single analyst performed the coding, no inter-rater reliability figures are reported, and this is acknowledged as a limitation.

RESULTS

The results are presented case by case and then summarised by indicator in the table below.

Chambers v Director of Public Prosecutions (2012)

Paul Chambers posted a message on Twitter after adverse weather closed Robin Hood Airport in England and disrupted his travel plans. He wrote that, unless the situation changed, he would "blow the airport sky high." The message appeared on a public timeline and was later found by airport personnel, who reported it to the police. Chambers was convicted under section 127 of the Communications Act 2003 of sending a message of a menacing character, and the conviction was upheld at an earlier stage of the proceedings.

In July 2012 the High Court allowed his appeal. It held that, assessed objectively in context, the message was not menacing, because a message that creates no fear or apprehension in those who read it, or may reasonably be expected to read it, cannot be of a menacing character. The judgment considered the wording, the circumstances of publication, the nature of the platform and the absence of evidence that the message was initially treated as a credible threat [Chambers v Director of Public Prosecutions, 2012]. It also stated the mental element: the sender must have intended the message to be menacing or have recognised the risk that it might create fear or apprehension in a reasonable reader, following the earlier House of Lords decision in *Collins* [Director of Public Prosecutions v Collins, 2006].

Elonis v. United States (2015)

Anthony Elonis published a series of Facebook posts containing violent language after the breakdown of his marriage and the loss of his job. The posts referred to his former wife, colleagues, a kindergarten, law-enforcement officers and an FBI agent. Elonis argued that the posts were artistic expression influenced by rap lyrics and that he had not intended to threaten anyone. The prosecution argued that they breached the federal statute criminalising the transmission of threats to injure another person [18 U.S.C. § 875(c)].

The jury was instructed to ask whether a reasonable person would regard the communications as threatening. Elonis was convicted on

four counts, and the appellate court upheld the convictions. In June 2015 the Supreme Court reversed and remanded. It held that a negligence-based standard was insufficient under the statute and that the defendant's mental state had to be considered, but it did not decide which mental state suffices in every true-threat case, for example whether recklessness is enough [Elonis v. United States, 2015].

Counterman v. Colorado (2023)

Between 2014 and 2016, Billy Counterman sent hundreds of Facebook messages to a local singer and musician whom he had never met. She never replied and repeatedly blocked him, but he continued to contact her through newly created accounts. Some messages were overfamiliar, some suggested that he was monitoring her, and others expressed anger and referred to possible harm. The recipient became fearful and contacted the authorities. Counterman was prosecuted under Colorado's stalking statute. The trial court applied an objective standard, asking whether a reasonable person would regard the messages as threatening, and the Colorado Court of Appeals affirmed the conviction.

On 27 June 2023 the Supreme Court held that the First Amendment requires the prosecution to prove some subjective understanding of the statement's threatening nature, and that recklessness suffices: the defendant must have consciously disregarded a substantial risk that the communication would be viewed as threatening violence. The judgment was vacated and the case remanded [Counterman v. Colorado, 2023].

Summary by indicator

Table.

Application of the seven indicators to the three judgments

Indicator	Chambers v DPP (2012)	Elonis v. United States (2015)	Counterman v. Colorado (2023)
1. Target and addressee	An institution (the airport); no direct addressee	Named individuals and institutions: former wife, colleagues, a kindergarten, officers, an FBI agent	One identifiable individual, a local singer and musician
2. Speech-act form and function	Violent wording; reading contested between menace and frustration	Violent wording; reading contested between threat and artistic expression	Mixed messages (familiar, monitoring, angry, referring to harm); reading contested
3. Genre and communicative frame	Public tweet during travel disruption, defended as a joke	Facebook posts in a rap-lyric style, defended as artistic expression	Direct Facebook messages sent from new accounts after blocking
4. Hyperbole and specificity	Conditional, exaggerated statement without operational detail	Violent imagery; whether hyperbolic or serious was disputed	References to possible harm combined with cues of monitoring
5. Audience and context	Public timeline; found by airport personnel	Facebook posts; context of marital breakdown and job loss	Private messages to a recipient who never replied
6. Repetition and communication pattern	Single message	Series of posts	Hundreds of messages, 2014–2016
7. Recipient response and uptake	Reported to police; no evidence of initial treatment as credible threat	Concern reported to authorities; prosecution followed	Recipient fearful; blocked sender repeatedly; contacted authorities

Across the three cases, the courts' reasoning rested on a combination of features rather than on the presence of violent vocabulary. In *Chambers*, the key features were context and the absence of a credible-threat response. In *Elonis* and *Counterman*, the central question was the defendant's mental state, and *Counterman* also involved a pattern of repeated contact with a single identifiable recipient. The mental-state standards differ: intention or recognised risk under the English statute, a statutory mental state left incompletely specified in *Elonis*, and constitutional recklessness in *Counterman*. None of the judgments treated an individual word or grammatical structure as sufficient by itself.

DISCUSSION

The *Chambers* decision shows why a literal reading is unreliable. The tweet contained a violent proposition, yet the court looked to the platform, the ordinary frustration of the situation and the airport's own reaction before concluding that the message was not menacing. This supports the pragmatic claim that communicated meaning exceeds literal content [Grice, 1975] and that hearers rely on context and frame [Goffman, 1974; Sperber & Wilson, 1986]. It resembles the treatment of a conditional political remark as hyperbole in *Watts* [Watts v. United States, 1969] and corpus findings that threatening stance varies with circumstances [Gales, 2015]. The case does not, however, establish that humorous framing protects a statement from being read as a threat. Prosecution guidance likewise stresses that context must be assessed in every case [Crown Prosecution Service, 2018].

Elonis and *Counterman* shift attention from how a message is received to what its sender knew. Together with the mental element stated in *Chambers*, they show a convergence across two jurisdictions towards requiring some awareness of the risk of being understood as threatening, although the legal bases differ. For linguists this marks a boundary. Lexical choices, metaphors, repetition and references can inform interpretation, but they cannot demonstrate intention [Coulthard et al., 2017; Olsson & Luchjenbroers, 2014]. The rap-lyric defence in *Elonis* also shows that frames can be invoked strategically. An analyst should therefore describe the claimed frame and ask whether the message and its circumstances actually fit it.

Counterman shows that the unit of analysis may be a sequence rather than a sentence. Indicators 6 and 7 capture this: repeated unwanted contact, continuation after blocking and the recipient's fear all become part of the interpretive picture. Behavioural threat assessment similarly treats behaviours such as fixation as warning signs [Meloy et al., 2012], although such signs concern risk and are not the same as the communicative act of threatening [Meloy & O'Toole, 2011]. The indicators are also consistent with research showing that threats are realised through varied linguistic forms and depend on speaker, target and circumstances [Fraser, 1998; Storey, 1995; Muschalik, 2018], and with analyses of online abuse as socially situated action [Hardaker & McGlashan, 2016]. The framework adds a link between such descriptions and recent legal standards.

The findings have practical implications. An analyst should examine the complete available communication, identify the apparent target, consider the platform and communicative setting, and relate the disputed statement to surrounding messages. Automated systems that identify threats through keywords may overlook these distinctions. Violent expressions occur in news reporting, political criticism, fiction, humour and appeals for peace, whereas genuine threats may be indirect and contain no explicit threat verb. Research on hate-speech detection already reports confusion between offensive and hateful language [Davidson et al., 2017], dependence on surface lexical features [Schmidt & Wiegand, 2017] and bias against particular dialects [Sap et al., 2019]. Context-sensitive analysis is therefore needed to reduce both false positives and false negatives.

The study has limitations. It relies on three cases from two jurisdictions with different legal frameworks, so their holdings should not be treated as interchangeable rules. The coding was performed by one analyst, and the seven indicators have not been tested on a large dataset or for inter-rater reliability. The framework is therefore an analytical synthesis rather than a validated predictive model. All three cases involve English-language communication, and threats in

other languages may behave differently. Further research could apply the indicators to larger collections of publicly available online communication and test their reliability across analysts, platforms and languages, including Uzbek and Russian.

CONCLUSION

This article has examined the distinction between threatening language and hyperbole through three judgments: Chambers v Director of Public Prosecutions, Elonis v. United States and Counterman v. Colorado. Each contributes a different perspective on interpreting potentially threatening online communication. Chambers shows the importance of context in separating exaggerated expression from a menacing message. Elonis highlights the relevance of the speaker's mental state under federal threat legislation. Counterman establishes that the First Amendment requires subjective awareness at least at the level of recklessness.

The seven indicators developed here offer a structured way to examine potentially threatening messages. They direct attention to the target, speech-act function, communicative frame, specificity, audience, repetition and recipient response. No single indicator is sufficient in isolation, and the framework is no substitute for legal analysis. The scientific contribution of the article is a synthesis that links linguistic description to recent legal standards of mental state.

Ultimately, distinguishing threats from hyperbole requires careful interpretation of language within its communicative and legal context. Forensic linguistics can clarify how statements function and why alternative interpretations arise, while recognising that linguistic evidence cannot by itself establish intent or legal responsibility.

Statement on the use of artificial intelligence. Artificial intelligence tools were used in preparing this article only for language editing and stylistic improvement. The scientific content, analysis and conclusions belong to the author.

REFERENCES

1. 18 U.S.C. § 875(c). Interstate communications.
2. Austin, J. L. (1962). *How to do things with words*. Oxford: Oxford University Press.
3. boyd, d. (2014). *It's complicated: The social lives of networked teens*. New Haven, CT: Yale University Press.
4. Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
5. Burgers, C., Brugman, B. C., Renardel de Lavalette, K. Y., & Steen, G. J. (2016). HIP: A method for linguistic hyperbole identification in discourse. *Metaphor and Symbol, 31*(3), 163–178.
6. Chambers v Director of Public Prosecutions [2012] EWHC 2157 (Admin).
7. Claridge, C. (2011). *Hyperbole in English: A corpus-based study of exaggeration*. Cambridge: Cambridge University Press.
8. Communications Act 2003 (UK), s. 127.
9. Coulthard, M., Johnson, A., & Wright, D. (2017). *An introduction to forensic linguistics: Language in evidence* (2nd ed.). London: Routledge.
10. Counterman v. Colorado, 600 U.S. 66 (2023).
11. Crown Prosecution Service. (2018). *Social media: Guidelines on prosecuting cases involving communications sent via social media*. London: Crown Prosecution Service.
12. Culpeper, J. (2011). *Impoliteness: Using language to cause offence*. Cambridge: Cambridge University Press.
13. Davidson, T., Warmsley, D., Macy, M., & Weber, I. (2017). Automated hate speech detection and the problem of offensive language. *Proceedings of the International AAAI Conference on Web and Social Media, 11*(1), 512–515.
14. Director of Public Prosecutions v Collins [2006] UKHL 40.

15. Elonis v. United States, 575 U.S. 723 (2015).
16. Fraser, B. (1998). Threatening revisited. *Forensic Linguistics, 5*(2), 159–173.
17. Gales, T. (2011). Identifying interpersonal stance in threatening discourse: An appraisal analysis. *Discourse Studies, 13*(1), 27–46.
18. Gales, T. (2015). Threatening stances: A corpus analysis of realized vs. non-realized threats. *Language and Law / Linguagem e Direito, 2*(2), 1–25.
19. Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. New York, NY: Harper & Row.
20. Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics, Vol. 3: Speech acts* (pp. 41–58). New York, NY: Academic Press.
21. Hardaker, C., & McGlashan, M. (2016). "Real men don't hate women": Twitter rape threats and group identity. *Journal of Pragmatics, 91*, 80–93.
22. Herring, S. C. (2013). Discourse in Web 2.0: Familiar, reconfigured, and emergent. In D. Tannen & A. M. Trester (Eds.), *Discourse 2.0: Language and new media* (pp. 1–25). Washington, DC: Georgetown University Press.
23. Marwick, A. E., & boyd, d. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society, 13*(1), 114–133.
24. McCarthy, M., & Carter, R. (2004). "There's millions of them": Hyperbole in everyday conversation. *Journal of Pragmatics, 36*(2), 149–184.
25. Meloy, J. R., Hoffmann, J., Guldimann, A., & James, D. (2012). The role of warning behaviors in threat assessment: An exploration and suggested typology. *Behavioral Sciences & the Law, 30*(3), 256–279.
26. Meloy, J. R., & O'Toole, M. E. (2011). The concept of leakage in threat assessment. *Behavioral Sciences & the Law, 29*(4), 513–527.
27. Muschalik, J. (2018). *Threatening in English: A mixed method approach*. Amsterdam: John Benjamins.
28. Olsson, J., & Luchjenbroers, J. (2014). *Forensic linguistics* (2nd ed.). London: Bloomsbury.
29. Sap, M., Card, D., Gabriel, S., Choi, Y., & Smith, N. A. (2019). The risk of racial bias in hate speech detection. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 1668–1678). Florence: Association for Computational Linguistics.
30. Schmidt, A., & Wiegand, M. (2017). A survey on hate speech detection using natural language processing. In *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media* (pp. 1–10). Valencia: Association for Computational Linguistics.
31. Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
32. Searle, J. R. (1976). A classification of illocutionary acts. *Language in Society, 5*(1), 1–23.
33. Shuy, R. W. (1993). *Language crimes: The use and abuse of language evidence in the courtroom*. Oxford: Blackwell.
34. Solan, L. M., & Tiersma, P. M. (2005). *Speaking of crime: The language of criminal justice*. Chicago, IL: University of Chicago Press.
35. Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.
36. Storey, K. (1995). The language of threats. *Forensic Linguistics, 2*(1), 74–80.
37. Tiersma, P. M. (1999). *Legal language*. Chicago, IL: University of Chicago Press.
38. Virginia v. Black, 538 U.S. 343 (2003).
39. Watts v. United States, 394 U.S. 705 (1969).