

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE DIREITO

Jessicka Oliveira de Assis

**O papel da Linguística Forense na identificação da autoria em cibercrimes cometidos
única ou exclusivamente através da linguagem**

JUIZ DE FORA

2020

Jessicka Oliveira de Assis

**O papel da Linguística Forense na identificação da autoria em cibercrimes cometidos
única ou exclusivamente através da linguagem**

Projeto de Monografia apresentado ao curso de
Direito da Universidade Federal de Juiz de Fora
como requisito parcial para a obtenção de
Bacharela em Direito.

Orientador: Prof. Dr. Wagner Silveira Rezende

JUIZ DE FORA

2020

Jessicka Oliveira de Assis

**O papel da Linguística Forense na identificação da autoria em cibercrimes cometidos
única ou exclusivamente através da linguagem**

Projeto de Monografia apresentado ao curso de
Direito da Universidade Federal de Juiz de Fora
como requisito parcial para a obtenção de
Bacharela em Direito.

PARECER DA BANCA

APROVADO

REPROVADO

Juiz de Fora, 17 de março de 2021.

BANCA EXAMINADORA

Prof. Dr. Wagner Silveira Rezende – Orientador
Universidade Federal de Juiz de Fora

Prof.^a Dr.^a Ana Paula Grillo El-Jaick
Universidade Federal de Juiz de Fora

Prof.^a Dr.^a Clarissa Diniz Guedes
Universidade Federal de Juiz de Fora

Prof. Dr. Richard Malcolm Coulthard
Universidade Federal de Santa Catarina

AGRADECIMENTOS

Aos meus pais, Denilson Amaral Ferreira (*in memoriam*) e Claudia de Oliveira Moreira, e a todos os professores que me formaram, em especial meu orientador, Wagner Silveira Rezende, descomunal e infindável agradecimento conjunto já que, filha de professores, o amor florescido em mim é intrinsecamente docente.

À toda a minha família, em especial ao meu irmão, Thales Oliveira Ferreira, e aos meus amigos, família por mim escolhida. Ao meu namorado, Pablo Ramonyer Félix Rocha, e ao meu fiel escudeiro, Ronan Nascimento de Oliveira Santos. Vocês são minha fonte precípua de amor, força e suporte.

A todos que direta ou indiretamente contribuíram para esta pesquisa.

RESUMO

O presente estudo busca argumentar acerca da pertinência de uma ciência para a identificação da autoria em cibercrimes cometidos única ou parcialmente através da linguagem, a Linguística Forense, que será, a partir de agora, tratada também pela sigla LF. O objetivo principal desta pesquisa é o de analisar como a literatura especializada trata o tema da aplicação da Linguística Forense para identificar a autoria em cibercrimes cometidos única ou parcialmente através da linguagem. Ainda, o presente estudo se propõe a arguir a inevitabilidade da aplicação de um instituto linguístico no que tange aos crimes cometidos através da linguagem. Os objetivos específicos estão diretamente relacionados ao objeto delimitado como tema, sendo eles: (i) conceituar e analisar estatisticamente a ocorrência dos cibercrimes no Brasil, apresentando, ainda, os institutos legais e previsões normativas que contemplem o crime virtual; (ii) apresentar como se dá a verificação de materialidade e autoria nos crimes cometidos única ou parcialmente através da linguagem, especialmente no que se refere à dificuldade de identificação do segundo elemento; (iii) conceituar e analisar a Linguística Forense, apresentando casos históricos célebres; (iv) examinar a possibilidade de aplicação da Linguística Forense para a solução de crimes virtuais, sob à luz do chamado criminal profiling; (v) apresentar a incipiência jurídica tanto no que se refere à Linguística Forense quanto aos cibercrimes, apresentando, ainda, se foram encontrados dados que tratem da aplicação daquela como instrumento de identificação da autoria nestes; (vi) arguir, por fim, a necessidade de uma ciência que solucione os cibercrimes, apresentando, como possível solução para aqueles cometidos única ou parcialmente através da linguagem, a Linguística Forense. Utilizando-se das obras de Malcom Coulthard (2014/2015) como marco teórico, o presente estudo utiliza-se de metodologia qualitativa, valendo-se de pesquisas bibliográfica e documental, além da realização de entrevista semiestruturada. Os dados serão, portanto, colhidos por meio de levantamento documental em sítios da internet, previsões legais e normativas, como o Código Penal, o Marco Civil da Internet, a Lei Carolina Dieckmann, além de jurisprudências e súmulas, bem como uma entrevista e uma pesquisa bibliográfica em periódicos, bancos de teses e dissertações, revistas científicas e informativas. Partindo do princípio de que os meios habituais de solução de crimes não se mostram efetivos quando aplicados àqueles cometidos única ou parcialmente através da linguagem, na seara virtual, a Linguística Forense parece se apresentar como um instrumento particularmente importante para a investigação destes crimes. No entanto, partimos também da hipótese, a ser comprovada, da incipiência no cenário nacional, tanto no que se refere aos cibercrimes quanto à Linguística Forense, de uma literatura especializada que trate sobre o tema, a despeito de sua importância. Sendo assim, as perguntas que norteiam este estudo são: qual seria o meio ideal para a identificação da autoria nesses cibercrimes? Já existe algum elemento, na literatura especializada brasileira, que aponte para a utilização de alguma técnica específica? A pertinência do presente estudo, portanto, se justifica, para além da incipiência do ramo, pela aplicação de uma ciência específica para a identificação da autoria em crimes ainda mais específicos, a fim de contribuir para a melhoria da investigação de autoria no contexto cibercriminal.

Palavras-chave: Linguística Forense; Cibercrimes; *Criminal profiling*.

ABSTRACT

The present study seeks to argue about the relevance of a science for the identification of authorship in cybercrimes committed only or partially through language, the Forensic Linguistics, which will, from now on, also be treated by the acronym LF. The main objective of this research is to analyze how the specialized literature deals with the subject of the application of Forensic Linguistics to identify authorship in cybercrimes committed only or partially through language. Still, the present study proposes to argue the inevitability of the application of a linguistic institute with regard to crimes committed through language. The specific objectives are directly related to the object delimited as a theme, which are: (i) conceptualize and statistically analyze the occurrence of cybercrimes in Brazil, also presenting the legal institutes and normative predictions that contemplate cyber crime; (ii) present how materiality and authorship are verified in crimes committed solely or partially through language, especially with regard to the difficulty of identifying the second element; (iii) conceptualize and analyze Forensic Linguistics, presenting famous historical cases; (iv) examine the possibility of applying Forensic Linguistics to solve virtual crimes, in the light of the so-called criminal profiling; (v) present the legal incipience both with regard to Forensic Linguistics and cybercrime, also presenting whether data were found that deal with the application of that as an instrument of identification of authorship in them; (vi) to argue, finally, the need for a science that solves cybercrimes, presenting, as a possible solution for those committed only or partially through language, Forensic Linguistics. Using the works of Malcom Coulthard (2014/2015) as a theoretical framework, the present study uses a qualitative methodology, using bibliographical and documentary research, in addition to conducting a semi-structured interview. The data will therefore be collected through documentary surveys on websites, legal and normative provisions, such as the Penal Code, the Marco Civil da Internet, the Carolina Dieckmann Law, in addition to jurisprudence and overviews, as well as an interview and bibliographic research in periodicals, thesis and dissertation banks, scientific and informational magazines. Assuming that the usual means of solving crimes are not effective when applied to those committed only or partially through language, in the virtual field, Forensic Linguistics seems to present itself as a particularly important tool for the investigation of these crimes. However, we also start from the hypothesis, to be proven, of the incipience on the national scene, both with regard to cybercrime and forensic linguistics, of a specialized literature that deals with the subject, despite its importance. Therefore, the questions that guide this study are: what would be the ideal way to identify authorship in these cybercrimes? Is there already an element in the Brazilian specialized literature that points to the use of a specific technique? The relevance of this study, therefore, is justified, in addition to the incipience of the branch, by the application of a specific science for the identification of authorship in even more specific crimes, in order to contribute to the improvement of the investigation of authorship in the cybercriminal context.

Keywords: Forensic linguistics; Cybercrimes; Criminal profiling.

SUMÁRIO

1 INTRODUÇÃO	7
2 OS CIBERCRIMES	11
2.1 LEGISLAÇÃO APLICÁVEL AO CRIME EM AMBIENTE VIRTUAL NO BRASIL.....	12
2.2 A AUTORIA NOS CIBERCRIMES COMETIDOS ÚNICA OU PARCIALMENTE ATRAVÉS DA LINGUAGEM.....	13
3 A LINGUÍSTICA FORENSE	17
3.1 A LÍNGUA COMO EVIDÊNCIA	18
3.2 O CASO UNABOMBER	20
3.3 A TÉCNICA DO <i>CRIMINAL PROFILING</i>	24
3.3.1 Perfilamento linguístico	26
4 A LINGUÍSTICA FORENSE E OS CIBERCRIMES NO BRASIL	29
4.1 LANGUAGE AND LAW	29
4.2 REVISTA BRASILEIRA DE CIÊNCIAS PENAIS	31
4.3 CATÁLOGO DE TESES E DISSERTAÇÕES DA CAPES	32
CONSIDERAÇÕES FINAIS.....	34
REFERÊNCIAS	35
APÊNDICE A - ENTREVISTA COM JAMES FITZGERALD	39
APÊNDICE B – INTERVIEW WITH JAMES R. FITZGERALD.....	43

1 INTRODUÇÃO

O Brasil era, em 2017, apresentado como o quarto dos vinte países com maior número de pessoas conectadas na rede mundial de computadores¹. Com 120 milhões de usuários, o conjunto de modos de pensar, técnicas, práticas, atitudes e valores atrelado ao crescimento do ciberespaço e denominado por Pierre Lévy (1999) “cibercultura” torna-se realidade inegável no cenário nacional.

A despeito dos inúmeros benefícios advindos da multiplicidade de relações que se desenvolvem nessa nova seara comunicacional, a cibercultura trouxe consigo um novo gênero de criminalidade. Como o crime comum está inerentemente inserido na realidade não virtual, de forma equivalente, na realidade do ciberespaço insere-se o cibercrime (KUNRATH, 2017).

Indicadores da SaferNet apontam para o recebimento e processamento, em 14 anos, de 3.244.768 denúncias anônimas, envolvendo 598.489 páginas (URLs) distintas (das quais 196.500 foram removidas), escritas em 9 idiomas e hospedadas em 58.209 domínios diferentes, de 254 diferentes TLDs e conectados à Internet através de 37.179 números IPs distintos, atribuídos a 101 países em 5 continentes².

Os delitos elencados no sítio do canal de denúncias se referem a racismo, violência ou discriminação contra mulheres, homofobia, tráfico de pessoas, apologia e incitação a crimes contra a vida, pornografia infantil, neonazismo, xenofobia, intolerância religiosa, maus tratos contra animais. Destes, apenas o tráfico de pessoas e os maus tratos contra animais não podem ser cometidos única ou parcialmente através da linguagem³.

Destaca-se que a SaferNet encaminha às autoridades somente as denúncias de crimes cuja ação penal seja pública e incondicionada à representação, de forma que no sítio não podem ser encontrados dados estatísticos dos crimes contra a honra⁵, por exemplo.

¹ Disponível em <https://exame.com/tecnologia/brasil-e-o-4o-pais-em-numero-de-usuarios-de-internet/>. Acesso em 02 dez. de 2020

² Disponível em <https://indicadores.safernet.org.br/indicadores.html>. Acesso em 02 dez. de 2020

³ Considera-se, neste ponto, que a pornografia infantil é delito passível de ser cometido através da linguagem, tendo em vista que o núcleo do tipo exige tão somente a exposição sexual infantil e que a imagem é, igualmente, um tipo de linguagem.

⁴ Destaca-se, ainda, que, de acordo com dados de jurisprudência do Jusbrasil, somente os crimes de ameaça, por exemplo, perfazem um total de 34.000 (trinta e quatro mil) processos na seara penal brasileira, o que desponta a latência de se tratar a temática dos crimes cometidos única ou parcialmente através da linguagem. Disponível em <https://www.jusbrasil.com.br/topicos/26413060/ameaca>. Acesso em 27 jan. 2021.

⁵ Disponível em <https://new.safernet.org.br/content/cal%C3%BAnia-difama%C3%A7%C3%A3o>. Acesso em 02 dez. de 2020

No entanto, como tais dados não se apresentam como condição para o presente estudo, esta pesquisa não se exime de analisar estes e outros crimes que, tendo em vista dados da jurisprudência e doutrina pátrias, são igualmente necessários para a análise que aqui se propõe.

Impulsionados pela sensação de anonimato e liberdade, os cibercrimes apresentam características peculiares quando comparados aos crimes não virtuais, como a transnacionalidade, a deslocalização, a atemporalidade, a permanência, o automatismo, a repetição, a alta tecnologia, a disseminação e a potenciação de danos (DIAS, 2010). Não à toa, a nocividade de tais ilícitos aliada à impunidade advinda do anonimato e da consequente dificuldade de identificação do autor do crime têm gerado diversos estudos dedicados ao perfilamento do criminoso eletrônico.

Nesse sentido, o presente estudo busca argumentar acerca da pertinência de uma ciência para a identificação da autoria em cibercrimes cometidos única ou parcialmente através da linguagem, a Linguística Forense, que será, a partir de agora, tratada também pela sigla LF.

O objetivo principal desta pesquisa é o de analisar como a literatura especializada trata o tema da aplicação da Linguística Forense para identificar a autoria em cibercrimes cometidos única ou parcialmente através da linguagem. Ainda, o presente estudo se propõe a arguir a inevitabilidade da aplicação de um instituto linguístico no que tange aos crimes cometidos através da linguagem.

Isso porque, apesar de comungarmos com a visão de que é “[...] impossível dizer que uma análise Linguística Forense da linguagem falada ou escrita em um crime seria a "melhor", ou mesmo "inevitável" para sua solução”, tendo em conta a dificuldade de se “contestar as descobertas quantitativas relacionadas a impressões digitais e DNA”, acredita-se, aqui, que, “na ausência desse tipo de evidência direta, a análise da FL pode provar ser o método mais confiável de determinar a autoria” (FITZGERALD, James. Entrevista I. [jan. 2021]. Entrevistadora: Jessicka Oliveira de Assis. Juiz de Fora, 2021. 1 arquivo documento Word (21 KB). A entrevista na íntegra encontra-se transcrita no Apêndice B desta monografia).

Os objetivos específicos estão diretamente relacionados ao objeto delimitado como tema, sendo eles: (i) conceituar e analisar estatisticamente a ocorrência dos cibercrimes no Brasil, apresentando, ainda, os institutos legais e previsões normativas que contemplem o crime virtual; (ii) apresentar como se dá a verificação de materialidade e autoria nos crimes cometidos única ou parcialmente através da linguagem, especialmente no que se refere à dificuldade de identificação do segundo elemento; (iii) conceituar e analisar a Linguística Forense, apresentando casos históricos célebres; (iv) examinar a possibilidade de aplicação da Linguística Forense para a solução de crimes virtuais, sob à luz do chamado *criminal profiling*;

(v) apresentar a incipiência jurídica tanto no que se refere à Linguística Forense quanto aos cibercrimes, apresentando, ainda, se foram encontrados dados que tratem da aplicação daquela como instrumento de identificação da autoria nestes; (vi) arguir, por fim, a necessidade de uma ciência que solucione os cibercrimes, apresentando, como possível solução para aqueles cometidos única ou parcialmente através da linguagem, a Linguística Forense.

Utilizando-se das obras de Malcom Coulthard (2014/2015) como marco teórico, o presente estudo utiliza-se de metodologia qualitativa, valendo-se de pesquisas bibliográfica e documental, além da realização de entrevista semiestruturada. Os dados serão, portanto, colhidos por meio de levantamento documental em sítios da internet, previsões legais e normativas, como o Código Penal, o Marco Civil da Internet, a Lei Carolina Dieckmann, além de jurisprudências e súmulas, bem como uma entrevista e uma pesquisa bibliográfica em periódicos, bancos de teses e dissertações, revistas científicas e informativas.

No tópico 2, seguinte a esta introdução, o estudo irá analisar os cibercrimes, seu conceito, a previsão normativa vigente para solução de tais delitos e trazer dados estatísticos acerca de seu tratamento no direito penal brasileiro. O terceiro tópico discorre especificamente sobre a Linguística Forense, conceitos a ela atrelados e o caso mais célebre em que foi utilizada. A seguir, no tópico 4, destrincham-se as técnicas da Linguística Forense à luz do *criminal profiling*, como uma ciência que se proponha a encontrar caminhos para o enfrentamento dos cibercrimes quando a língua é a única fonte de evidência. Por fim, apresenta-se uma reflexão acerca da incipiência jurídica no que tange não apenas aos cibercrimes, mas também à Linguística Forense no Brasil, analisando-se os dados encontrados acerca da aplicação desta enquanto ferramenta de atribuição de autoria nos cibercrimes cometidos única ou parcialmente através da linguagem, discutindo-se, ainda, a pertinência de se eleger a Linguística Forense para este fim.

Partindo do princípio de que os meios habituais de solução de crimes não se mostram efetivos quando aplicados àqueles cometidos única ou parcialmente através da linguagem, na seara virtual, a Linguística Forense parece se apresentar como um instrumento particularmente importante para a investigação destes crimes. No entanto, partimos também da hipótese, a ser comprovada, da incipiência no cenário nacional, tanto no que se refere aos cibercrimes quanto à Linguística Forense, de uma literatura especializada que trate sobre o tema, a despeito de sua importância. Sendo assim, as perguntas que norteiam este estudo são: qual seria o meio ideal para a identificação da autoria nesses cibercrimes? Já existe algum elemento, na literatura especializada brasileira, que aponte para a utilização de alguma técnica específica?

A pertinência do presente estudo, portanto, se justifica, para além da incipiência do ramo, pela aplicação de uma ciência específica para a identificação da autoria em crimes ainda mais específicos, a fim de contribuir para a melhoria da investigação de autoria no contexto cibercriminal.

2 OS CIBERCRIMES

Segundo a empresa de segurança na internet, Symantec, a cada minuto, 54 pessoas são vítimas de crimes cibernéticos no Brasil. Nas duas delegacias especializadas de Belo Horizonte, por exemplo, mil procedimentos de casos registrados em Minas Gerais estão em andamento⁶.

De acordo com um relatório da Norton Cyber Security, em 2017, o Brasil passou a ser o segundo país com maior número de casos de crimes cibernéticos, afetando cerca de 62 milhões de pessoas e causando um prejuízo de US\$ 22 bilhões⁷.

Na mesma toada, Kunrath afirma que:

A escalada do cibercrime é um fenômeno mundial. O Brasil, conforme notoriamente noticiado pela imprensa nacional e estrangeira, é um paraíso da pirataria virtual. E, desde a criação da internet, no final da década de 60, o ilícito cibernético vem crescendo exponencialmente, na mesma proporção da ampliação da web pelo globo terrestre (KUNRATH, 2017, p. 27).

Diante desse cenário, tornou-se essencial entender não apenas o fenômeno dos crimes informáticos, cibernéticos ou virtuais, mas também as razões pelas quais a criminalidade cibernética vem aumentando exponencialmente, seja no âmbito internacional ou nacional⁸.

Ainda segundo Kunrath (2017, p. 27), o cibercrime “[...] constitui a exteriorização de condutas ilícitas dos usuários das tecnologias da informação e internautas, cada vez mais recorrentes no ciberespaço”. Além de diversas denominações para a mesma ideia, o fenômeno possui também diversas classificações.

No presente estudo, utiliza-se a classificação de Vianna (2003), que se refere a cibercrimes próprios ou impróprios. Os primeiros seriam figuras típicas que somente podem ser praticadas através da tecnologia ou de um sistema computadorizado; os segundos se referem a crimes comuns nos quais o computador é mero instrumento à execução do delito.

Pretende-se analisar mais atentamente os denominados cibercrimes impróprios, já que, para o cometimento dos delitos que se concretizam única ou exclusivamente através da linguagem, não é indispensável a via tecnológica.

⁶ Disponível em <https://amp-mg.jusbrasil.com.br/noticias/3125198/brasil-registra-54-crimes-virtuais-por-minuto>. Acesso em 02 dez. de 2020

⁷ Disponível em <https://br.norton.com/norton-cybersecurity-insights-report-brazil>. Acesso em 02 dez. de 2020

⁸ Destaca-se que tal temática exige fôlego próprio de pesquisa futura, motivo pelo qual servirá, neste momento, de pano de fundo e não de escopo do presente estudo.

2.1 LEGISLAÇÃO APLICÁVEL AO CRIME EM AMBIENTE VIRTUAL NO BRASIL

No cenário nacional, existem, para além das normas aplicáveis na seara não virtual, algumas aplicáveis, seja direta ou indiretamente, ao âmbito da cibercultura. Dentre elas, citam-se:

- ✓ A Constituição Federal de 1988, que, em seu art. 5º, inciso XXXIII, trata do direito à informação enquanto direito fundamental.
- ✓ As leis n.º 12.735 e 12.737, de 30 de novembro de 2012, que introduzem, no ordenamento penal brasileiro, os delitos informáticos. A primeira, popularmente conhecida como “Lei Azeredo”, propunha a criminalização de ataques de hackers e crackers⁹, tendo sido quase que integralmente vetada; a segunda, conhecida como “Lei Carolina Dieckmann”, inseriu os arts. 154-A e B, 266, 298 no Código Penal brasileiro.
- ✓ A Cartilha de Segurança para a Internet, elaborada em 2012 pelo Comitê Gestor da Internet no Brasil (CGI.br), através do Centro de Estudos, Resposta e Tratamento de Incidentes de Segurança no Brasil, que prevê recomendações e esclarecimentos sobre a segurança dos internautas.
- ✓ A lei n.º 12.965, de 23 abril de 2014, mais conhecida como Marco Civil da Internet, que prevê direitos, deveres e responsabilidades atinentes à cibercultura. O marco regulatório estabelece, em contraposição à legislação até então existente, pautada em restrições, condenações e proibições relativas ao uso da internet, garantias e direitos relativos ao acesso e uso livre da web. É a partir dessa lei que o acesso à internet é reconhecido como direito essencial ao exercício da cidadania, delineando-se como direito universal e fundamental.

Existem, ainda, outros dispositivos¹⁰ do próprio Código Penal que, introduzidos gradativamente por leis esparsas, incriminam algumas práticas ilícitas cometidas através da via tecnológica, como os arts. 153, §1o-A, 313-A, 313-B, 325, I (CRESPO, 2011).

⁹ Segundo o verbete da Wikipedia, hacker é “um indivíduo que se dedica, com intensidade incomum, a conhecer e modificar os aspectos mais internos de dispositivos, programas e redes de computadores”; já o termo cracker foi criado em 1985 a fim de retirar a carga pejorativo do termo “hacker”. Infere-se, portanto, que o cracker utiliza seus conhecimentos técnicos para forjar, por exemplo, licenças de programas e softwares pagos, disponibilizando-os gratuitamente para a população; o hacker, ao contrário, prejudica a sociedade, camuflando vírus em sites de alto acesso, por exemplo.

¹⁰ Uma das legislações acerca da temática é a Convenção sobre o Cibercrime, que ocorreu em Budapeste, em 23 de novembro de 2001, e prevê, para os Estados signatários, uma política criminal comum, com o objetivo de

Além dos mencionados, existem paradigmas normativos que poderiam servir à aplicação da legislação vigente no ciberespaço, como relatórios da Organização das Nações Unidas (ONU), que, em consonância com o artigo XIX da Declaração Universal dos Direitos Humanos, preveem a promoção e proteção do direito à liberdade de opinião e expressão, enfatizando o acesso à internet como direito universal, e entendimentos do Supremo Tribunal Federal, como aquele contido no Informativo 453/2006¹¹, que negou a pretensão de se preencher, por analogia, uma lacuna da lei penal.

Diante das já mencionadas características particulares do cibercrime, a questão que aqui se coloca é quanto à insuficiência ou mesmo inaplicabilidade das normas existentes para a seara não virtual àquela virtual.

A regra da territorialidade, definida pelo art. 70 do Código de Processo Penal, que define a competência territorial pelo lugar da consumação do delito, no caso de cibercrimes, é obstaculizada pela transnacionalidade e deslocalização dos delitos virtuais; as regras de fixação quanto à temporalidade, como o próprio princípio da legalidade ou da anterioridade penal, são obstadas pela atemporalidade, permanência e repetição características dos cibercrimes; a autoria, por fim, pode ser ocultada das maneiras mais diversificadas pela alta tecnologia e automatismo na disseminação dos crimes cibernéticos.

No presente trabalho, tratar-se-á especificamente da insuficiência ou ausência de normas que solucionem o problema da autoria, o que acredita-se gerar uma limitação à função punitiva estatal e, conseqüentemente, uma sensação de insegurança e impunidade para a sociedade brasileira.

2.2 A AUTORIA NOS CIBERCRIMES COMETIDOS ÚNICA OU PARCIALMENTE ATRAVÉS DA LINGUAGEM

O entendimento firmado pela Quinta Turma Superior do Supremo Tribunal de Justiça (STJ) no HC 433299/TO 2018/0008602-5 foi o de que:

combater a criminalidade no ciberespaço através da adoção de uma legislação adequada e da melhoria da cooperação internacional. O Brasil, no entanto, não é signatário do tratado.

¹¹ “Vencidos os Ministros Carlos Britto, Ricardo Lewandowski, Joaquim Barbosa e Marco Aurélio, que recebiam a denúncia. Inq 1145/PB, rel. orig. Min. Maurício Corrêa, rel. p/ o acórdão Min. Gilmar Mendes, 19.12.2006. (Inq-1145)”. Disponível em <http://www.stf.jus.br/arquivo/informativo/documento/informativo453.htm>.

HABEAS CORPUS Nº 433.299 - TO (2018/0008602-5) RELATOR: MINISTRO FELIX FISCHER IMPETRANTE: JANDER ARAUJO RODRIGUES ADVOGADO: JANDER ARAÚJO RODRIGUES - TO005574 IMPETRADO : TRIBUNAL DE JUSTIÇA DO ESTADO DO TOCANTINS PACIENTE : DIVA RIBEIRO DE MELO PACIENTE : EURIPEDES LOURENCO DE MELO DECISÃO

EMENTA: OFERECIMENTO DE DENÚNCIA PELO MINISTÉRIO PÚBLICO - ACUSADO [...] **DENÚNCIA QUE PREENCHE OS REQUISITOS DO ART. 41, DO CPP - INDÍCIOS SUFICIENTES DE MATERIALIDADE E AUTORIA** - SUPORTE PROBATÓRIO SATISFATÓRIO PARA A INSTAURAÇÃO E PROSSEGUIMENTO DA AÇÃO PENAL - PRINCÍPIO DO "IN DÚBIO PRO SOCIETATE"- TESES DEFENSIVAS NÃO DEMONSTRADAS DE FORMA CABAL E INEQUÍVOCA - NECESSIDADE DA INSTRUÇÃO PROCESSUAL PARA MELHOR APURAÇÃO DOS FATOS - RECEBIMENTO DA DENÚNCIA. Embora os denunciados defendam a falta de justa causa para o prosseguimento da ação penal, até o momento as provas presentes nos autos são suficientes para, nesta fase processual, demonstrar indícios de autoria e a materialidade delitiva. Ressalta-se que a jurisprudência desta Corte Superior é no sentido de que a propositura da ação penal exige tão somente a presença de indícios mínimos de autoria. A certeza será comprovada ou afastada durante a instrução probatória, prevalecendo, na fase de oferecimento da denúncia o princípio do in dubio pro societate (RHC 83.296/CE, Rel. Ministro FELIX FISCHER, QUINTA TURMA, julgado em 28/11/2017, DJe 6/12/2017.) (STJ - HC: 433299 TO 2018/0008602-5, Relator: Ministro HUMBERTO MARTINS, Data de Publicação: DJ 05/02/2018, *grifo meu*).

Dessa forma, nos termos do art. 41 do Código de Processo Penal e do entendimento disposto no excerto acima, a denúncia e, conseqüentemente, a instrução processual precisam demonstrar a autoria e materialidade do crime para que haja a descrição adequada do fato criminoso, apta a gerar sentença de condenação.

No que se refere à materialidade, ou seja, ao elemento comprobatório da ocorrência do delito, analisa-se, a seguir, como esta se dá no caso dos cibercrimes cometidos única ou parcialmente através da linguagem.

A despeito de, entre os crimes transeuntes, ou seja, aqueles que não deixam vestígios¹², normalmente serem mencionados os delitos cometidos pela via verbal¹³, no caso dos cibercrimes, onde a via tecnológica é condição para a execução do crime, a linguagem será necessariamente captada por tal via e, conseqüentemente, estará retida e memorizada em alguma base de dados. Nesse caso, a dificuldade do(a) ofendido(a) poderá ser a de conseguir meios judiciais para retirar o elemento probatório de sua base de armazenamento, nos casos em que ele(a) mesmo(a) não conseguiu capturá-lo.

¹² Disponível em JESUS, Damásio E. de. **Direito Penal**. 1o volume. Parte Geral. 34a ed. São Paulo: Saraiva, 2013, p. 260.

¹³ Disponível em <https://juspol.com.br/tabela-de-classificacao-de-crimes/>. Acesso em 03 dez. 2020

Como se depreende dos dados da SaferNet, Facebook, Twitter e YouTube aparecem sempre entre os cinco domínios com mais páginas denunciadas em todos os crimes virtuais cometidos pela via da linguagem, de forma que se infere que a materialidade possa ser, nesses casos, mais facilmente comprovada, por exemplo, por *print screens*, ou seja, por capturas de tela que, via de regra, devem ser registradas em cartório¹⁴.

Tendo em vista a já referida desatualização da legislação penal no que se refere à identificação da autoria em delitos cibernéticos, o presente estudo se dedica, após breve explanação acerca da menor complexidade da comprovação da materialidade em crimes dessa natureza, à análise do primeiro elemento. Acredita-se, aqui, que a principal característica do cibercrime que contribui para a dificuldade na identificação da autoria é o anonimato.

Monteiro Neto (2008) elenca como fatores do aumento da criminalidade cibernética a falsa ideia de que o mundo da tecnologia da informação é um ambiente sem regras, onde tudo é revestido do véu do anonimato do monitor, e a mudança de perfil do criminoso de tecnologia.

Acerca do anonimato, Vera Elisa Marques Dias afirma que:

Inegavelmente, o anonimato, a camuflagem ou o uso de identidade falsa é a característica mais aliciadora, tentadora e propulsora para a iniciação da prática criminosa na internet. É, também, a característica mais assegurada, recorrendo os infratores mais especializados ou as organizações através deles a técnicas que lhes permitam ocultar ou dissimular a sua identidade e as suas condutas, como a técnica de spoofing, programas de anonimização e codificação, que diariamente são aperfeiçoados e transformados. Para além de se assegurar o anonimato do autor também se pode ocultar a própria informação através de mecanismos de cifra forte ou de encriptação, como a estenografia, e outros disponíveis gratuitamente na rede. Podem, assim, os cibercriminosos diminuir ou eliminar o risco de ser descoberto ou condenado, apagando todas as provas do ciberrastro (DIAS, 2010, p. 16).

Uma das soluções para o anonimato é, na seara da cibercultura, traçar um perfilamento do criminoso tecnológico. Segundo Monteiro Neto (2008), é possível identificar e classificar as práticas delitivas que ocorrem pela via tecnológica através da semelhança entre as condutas.

Atualmente, ainda segundo o autor:

O perfil criado e divulgado pela mídia tem o criminoso eletrônico como sendo, em regra, indivíduo do sexo masculino, que trabalha de alguma forma com a utilização de computadores e sistemas eletrônicos, com idade entre 16 e 33 anos de idade, avessos a violência e que possuem inteligência acima da média. São extremamente audaciosos e aventureiros, movidos acima de tudo pelo desejo de conhecimento e de superação à máquina. Hoje tais delinquentes são, em geral, pessoas que trabalham no ramo ligado a utilização de sistemas eletrônicos, não tão jovens nem inteligentes; são insiders, vinculados a empresas (em regra); sua característica central consiste na pouca

¹⁴ Disponível em <https://brunomore.jusbrasil.com.br/artigos/608243055/o-crime-de-ameaca-na-internet>. Acesso em 03 dez. 2020

motivabilidade em relação à norma (raramente se sensibilizam com a punição penal); motivos para delinquir: ânimo de lucro, perspectiva de promoção, vingança, apenas para chamar a atenção etc (MONTEIRO NETO, 2018, p. 112).

Lucena (2012) afirma que, em relação ao perfil do criminoso dessa modalidade, a maioria é jovem e surge de camadas médias e altas da sociedade, possuindo habilidades e conhecimentos neste meio.

Dessa forma, os paradigmas teóricos da criminologia moderna permitem identificar, através do chamado *criminal profiling*, um standard de ciberinfrator, traçando seu grupo social, sua média de idade, seus interesses e até mesmo as motivações de seu comportamento transgressor, “como demonstração de poder, obtenção de informações, aplicação de golpes financeiros ou, simplesmente, para exploração curiosa de vulnerabilidades e diversão” (KUNRATH, 2017, p. 37).

No entanto, o perfilamento criminal de um grupo social, como um todo, não é suficiente para a identificação da autoria de um único agente na prática de determinado delito. Por isso, propõe-se, nos próximos tópicos, a aplicação expandida do *criminal profiling* aos cibercrimes cometidos pela via da linguagem, através da Linguística Forense.

3 A LINGUÍSTICA FORENSE

A Linguística Forense, para Coulthard (2014), é a inter-relação entre Linguagem e Direito, que consiste na “análise da linguagem em todos os tipos de interação no contexto jurídico” (CALDAS-COULTHARD, 2014: 2) e que pode ser dividida em Linguagem Escrita do Direito, Interação em Contextos Legais e Linguagem como Evidência. Para o autor, o ramo é parte integrante da seara da Análise Crítica do Discurso, de forma que o linguista forense não apenas descreve, mas também tenta mudar e melhorar o mundo¹⁵.

Dentre as subáreas da Linguística Forense, a Linguagem Escrita do Direito se refere ao movimento de aproximação do chamado *juridiquês*, ou seja, da linguagem jurídica, ao português acessível e amplamente utilizado. A Interação em Contextos Legais se refere aos direitos relacionados à tradução e interpretação de testemunhas vulneráveis, como crianças, vítimas de estupro, intelectualmente desafiados e idosos.

Destaca-se que a Análise Crítica do Discurso é o ramo da Linguística que enxerga o discurso como uma prática social e não apenas como uma representação do mundo (FAIRCLOUGH, 2008, p. 9). Os adeptos à Teoria dos Sistemas, como em Niklas Luhmann, são unânimes ao afirmar que todo sistema precisa do meio para se alimentar, por isso “O Direito, como qualquer outro sistema, depende da linguagem para exercer suas funções sociais” (REZENDE, W.S., 2014, p. 295). Em outra linha de interpretação, Bourdieu, quando propõe o conceito de campo, não comunga com o fetichismo hermético atribuído ao Direito, indo de encontro à concepção de que “para conhecer o Direito basta olhar para o Direito”; ao contrário, o filósofo e sociólogo entende a ciência jurídica enquanto campo social, de forma que o capital jurídico é, essencialmente, linguístico (REZENDE, W.S., 2014) e, enquanto ciência da palavra (FERREIRA; BARROS, 2015), é permeado pela visão da língua enquanto evidência, motivo pelo qual o presente estudo debruça-se especificamente sobre a terceira subárea da Linguística Forense.

A evidência, no Direito, é a prova material do delito, ou seja, a comprovação de vinculação direta de um vestígio a determinada prática ilícita; tal vinculação pode ser comprovada, por exemplo, a partir de um exame de corpo de delito, previsto no art. 158 do

¹⁵ Neste ponto, entende-se que a interpretação do excerto de que o linguista forense não apenas descreve a realidade deve ser feita poética e não literalmente. Isso porque, apesar de o linguista forense descrever a linguagem e, com isso, poder aprimorar a realidade fática, não cabe à Linguística Forense ações que ultrapassem a seara descritiva. A LF não poderia realizar a tarefa deontológica e normativa, por exemplo, de pretender alterar a forma como as pessoas falam e se comunicam; no entanto, a partir do caminho analítico e descritivo, esta pode contribuir com a identificação de crimes cometidos por meio da linguagem, como aqui se defende, e, assim, melhorar o mundo.

Código de Processo Penal¹⁶. Na seara dos cibercrimes cometidos pela via linguagem, este trabalho se propõe a contribuir para a identificação de evidências de autoria.

Tendo por base os números apresentados acerca dos crimes cometidos única ou parcialmente através da linguagem, como a calúnia, a injúria, a difamação, a ameaça, o estelionato e a extorsão¹⁷, além dos demais crimes citados anteriormente em dados da Safernet, busca-se verificar, nas próximas seções e a partir do histórico de utilização da Linguística Forense no mundo, se o fenômeno linguístico é encarado e utilizado como evidência no cenário nacional.

3.1 A LÍNGUA COMO EVIDÊNCIA

No Brasil, associações como a Associação Internacional de Linguistas Forenses (IAFL) e a Associação de Linguagem e Direito (ALIDI) têm trabalhado para desenvolver a Linguística Forense, arguindo que, a partir da singularidade linguística do indivíduo, é possível encontrar sequências compartilhadas de palavras minimamente idênticas, já que:

Ao produzir um texto, seja oral ou escrito, o sujeito lança mão de um vasto repertório lexical e regras de ordenação sintática pertencentes à gramática de seu idioma. Entretanto, esse arranjo não é feito da mesma forma por diferentes pessoas. Ao falarmos ou ao escrevermos, organizamos o material linguístico que está disponível em nosso acervo mental de uma forma única, afinal, cada indivíduo constituiu seu léxico a partir de experiências também únicas. Isso significa que imprimimos nosso estilo em nossos textos, deixando nele nossa “assinatura”. Esse uso individual do idioma é chamado de idioleto, ou seja, é como se fosse um dialeto pessoal, uma marca identitária daquele indivíduo. A noção de idioleto, entretanto, ainda não é consensual entre os linguistas (PEREIRA E SILVA, Welton, recurso online).

É, portanto, a partir de padrões específicos de usos de elementos da linguagem (vocabulário, pronúncia, ortografia, repertório lexical, regras de ordenação sintática, organização do material linguístico) que a Linguística Forense torna possível a criação de marcadores de estilo de um idioleto¹⁸. Isso porque, como afirma o gerativismo linguístico chomskiano, há a possibilidade de se gerar infinitas expressões linguísticas a partir de elementos constitutivos finitos (GLENDA, 2008, *apud* CHOMSKY, 1965). Portanto, é possível inferir

¹⁶ Disponível em https://www.gta.ufrj.br/grad/13_1/forense/terminologia.html#evidencia. Acesso em 04 dez. 2020.

¹⁷ Disponível em <http://www.roseta.org.br/pt/2019/09/16/linguistica-forense-como-o-linguista-pode-contribuir-em-uma-demanda-judicial/>. Acesso em 04 dez. 2020.

¹⁸ Considera-se, para os fins desta pesquisa, idioleto como o dialeto pessoal ou a variação intrafalante de determinado indivíduo, ou seja, como uma marca linguística identitária que representa, como será argumentado neste estudo, o estilo individual da utilização da linguagem feita por de determinado sujeito.

que, diante de uma mesma construção linguística, que poderia ter sido construída não apenas de diversas, mas de infinitas formas, o mais provável é que estejamos diante, também, de marcadores que representam o mesmo acervo mental, ou seja, o *input* de determinado indivíduo que, por ser tão específico quanto sua própria digital, pode nos fornecer um perfil. Além do mais, tendo em vista que a escrita pode ser vista a olho nu, diferentemente da impressão digital (SANTOS BONFIM, 2020), o perfilamento linguístico pode se apresentar como uma evidência mais aferível no caso concreto.

A partir deste perfilamento, é possível inferir dados como a faixa etária, o gênero, a orientação sexual, o nível de instrução, a profissão, a origem geográfica, as preferências e hobbies, a opção religiosa e a classe social do agente do crime. São esses dados que transformam a linguagem em evidência, tornando-a um elemento passível de atribuição de autoria.

A Fonética Forense, ramo da Linguística Forense, é um dos exemplos da utilização da Linguística Forense para atribuição de autoria, seja em chamadas de emergência e pedidos de resgate, seja em comunicações de ameaça etc. Outros exemplos são análises de cartas anônimas, testamentos, motivações de sentenças, marcas registradas e verificações de plágio, área extensamente estudada pelo professor Rui Sousa-Silva, além das situações em que é necessário:

[...] ratificar a autenticidade de bilhetes de suicídio, identificar o responsável por trás de cartas com pedido de resgate, comprovar ou descartar a autoria de confissões (como visto na equivocada condenação de Timothy Evans), avaliar gravações de entrevistas policiais supostamente opressoras; examinar e extrair informações sobre o autor nos mais diversos tipos de textos que possam interessar em uma investigação criminal — podendo, inclusive, as análises e conclusões do especialista serem usadas como prova pericial (SANTOS BOMFIM, Amanda, 2020, p. 14).

Um dos casos célebres da utilização da área na investigação policial é o Caso Timothy Evans, ocorrido na década de 50 nos Estados Unidos, no qual um pai de família foi acusado de assassinar sua esposa e a filha. Todas as evidências, no caso, eram baseadas em depoimentos de confissão. Timothy foi condenado e morto pelo caso, mas, um ano mais tarde, os corpos foram encontrados no apartamento que pertencia a John Reginald Halliday Christie, vizinho de Evans, à época dos assassinatos, que confessou o crime. Examinando as confissões atribuídas a Evans, Svartvik (1968) constatou vários estilos de linguagem, inclusive com características próprias de registros policiais, tendo identificado, ainda, características incomuns ao vocabulário de uma pessoa iletrada, como era Timothy.

Outro grande exemplo é o do Caso Prinzivalli, ocorrido em meados da década de 80, também nos Estados Unidos, em Los Angeles, no qual o linguista W. Labov foi enfático ao afirmar que Prinzivalli não poderia ser o responsável pelo telefonema de ameaça de bomba

investigado no caso, porque possuía um sotaque nova iorquino, enquanto a pessoa que fez a ameaça de bomba tinha um sotaque inconfundível da Nova Inglaterra (região de Boston)¹⁹.

O caso mais célebre da aplicação da Linguística Forense como elemento passível de atribuição de autoria de um delito, no entanto, é o Caso Unabomber, que será apresentado a seguir.

3.2 O CASO UNABOMBER

Incontornável para a análise da Linguística Forense, o caso Unabomber se popularizou por sua expansão midiática. Tratado em diversos programas de TV, como em um episódio da série *Manhunt* e pela série *Unabomber: In His Own Words*, o caso diz respeito ao terrorista doméstico homônimo. Theodore Kaczynski é um matemático anarquista que abandonou sua carreira acadêmica para viver um estilo de vida primitivo, residindo numa cabana remota sem eletricidade ou água corrente, próximo ao Estado de Montana, nos Estados Unidos.

Com um QI pontuado em 167, Theodore repudiava as tecnologias contemporâneas e, numa tentativa de iniciar uma revolução contra a industrialização, se tornou perito na construção de explosivos e bombardeou diversos espaços nacionais que acreditava simbolizarem a tecnologia moderna, como universidades e companhias aéreas, o que lhe rendeu a alcunha Unabomber ou Unabomb (acrônimo das palavras *University and Airline Bomber*).

De acordo com Santos Bonfim (2020) *apud* Turchie e Noel (2018), em maio de 1978 foi encontrado um embrulho no estacionamento da Universidade de Illinois, endereçado ao professor de tecnologia E. J. Smith. Era o primeiro de dezesseis bombardeios que seriam conduzidos por Kaczynski e que se estenderiam até 1995, matando três e ferindo vinte e três pessoas até o momento de sua captura.

A investigação do caso Unabomber foi uma das mais dispendiosas na história do FBI, com uma força-tarefa composta por centenas de agentes e com um custo estimado de 50 milhões de dólares (SANTOS BONFIM, 2020 *apud* PULITZER; WIEHL, 2020). Isso porque capturar o serial bomber se tornou uma prioridade para os estadunidenses diante da ameaça de uma explosão iminente direcionada, pelos correios, a qualquer cidadão americano, já que:

[...] suas bombas eram endereçadas especialmente para professores universitários e companhias aéreas, embora muitos cidadãos “comuns” também tenham sido vítimas, por acidente ou não. Diante de tamanha ameaça,

¹⁹ Disponível em <http://www.roseta.org.br/pt/2019/09/16/linguistica-forense-como-o-linguista-pode-contribuir-em-uma-demanda-judicial/>. Acesso em 04 dez. 2020.

os estadunidenses ficaram em polvorosa e muito assustados — afinal, as bombas poderiam ser endereçadas a qualquer um. Além disso, é importante ter em mente que, no final do século XX, os correios desempenhavam um papel muito mais ativo na vida dos norte-americanos. Seria o equivalente a, hoje, correr o risco de ser vítima de uma explosão ao escutar um áudio de um aplicativo de mensagens instantâneas [...] Um exemplo do terror causado por Theodore Kaczynski foi a terceira bomba, enviada em novembro de 1979. O pacote estava em uma aeronave da United Airlines, durante um voo que levava 80 pessoas a bordo. O dispositivo fora programado para explodir quando o avião alcançasse certa altitude, e assim o fez — a bomba explodiu em pleno ar. Mesmo com a explosão, o avião milagrosamente conseguiu pousar sem muitos danos e zero feridos (SANTOS BONFIM, 2020, p. 15).

Em junho de 1993, o misterioso Unabomber envia uma carta anônima ao New York Times, na tentativa de convencer a imprensa de que por trás dos ataques estava um grupo terrorista. Era o primeiro passo que conduziria à captura do assassino.

Em abril de 1995, nova carta, que chega a 1.700 palavras, explica a motivação para os bombardeios e para a escolha das vítimas, além de trazer consigo uma barganha. O pedido de Kaczynski era de que o jornal publicasse seu manifesto “A sociedade industrial e seu futuro” (1995) e, em troca, ele cessaria os assassinatos. Da recusa, o matemático daria início à construção de uma próxima bomba.

Em contato com o FBI, o jornal foi orientado a publicar o manifesto, na esperança de que algum amigo ou familiar pudesse reconhecer o trabalho, a filosofia de vida ou as próprias palavras de Theodore. E foi o que aconteceu. O irmão de Kaczynski entrou em contato com o FBI, apresentando um ensaio escrito em 1971 que apresentava similaridades com o documento publicado.

O agente responsável pela análise comparativa dos documentos foi James Fitzgerald (2017), que relatou “termos, sentenças e frases verbais e nominais quase idênticos em ambos os documentos”. Ambos os documentos deveriam ter sido escritos por um sujeito que possuísse “excelente domínio da língua inglesa; ele certamente tinha um vocabulário extenso e habilidades estilísticas para alterar levemente seu léxico e a estrutura de suas sentenças, embora o tempo inteiro transmitisse conceitos muito, muito similares”.

A partir daí, Fitzgerald foi o responsável por liderar e organizar uma equipe que realizaria uma análise linguística comparativa de todos os documentos entregues ao FBI pelo irmão do assassino. Era essa equipe que encontraria, após extenso e minucioso estudo de mais de 50 páginas com cerca de 600 exemplos categorizados em “similares”, “muito similares” e “virtualmente idênticos”, tanto no manifesto publicado quanto em uma carta escrita por

Theodore nos anos 70, a seguinte frase: “We can’t eat our cake and have it too” (“Nós não podemos comer nosso bolo e tê-lo também”, em tradução livre).

A questão principal para a solução do caso Unabomber residiu na escrita equivocada de um provérbio célebre em inglês, que diz “You can't have your cake and eat it (too)” (Você não pode comer seu bolo e tê-lo também”, em tradução livre). A escrita peculiar da sentença no plural, com as mesmas palavras e na mesma ordem invertida, foi a evidência linguística que estabeleceu uma causa provável forte o suficiente para um mandado de busca para revistar a cabana onde morava o Unabomber e que resultou em sua prisão.

Muitos consideram que o caso Unabomber foi o início da Linguística Forense e o ponto de partida para diversos estudos que desembocariam na chamada “Atribuição de Autoria”:

[...] que Juola (2008) define como uma ciência voltada para a análise de características do autor a partir do estudo de documentos escritos por aquele autor. Pode-se entender melhor esse processo a partir da seguinte explicação de McMEnamin (2002): “No estudo da variação linguística, há um método para organizar observações que leva à identificação ou separação de marcadores de estilo e posteriormente a uma descrição e qualificação sistemáticas (SANTOS BONFIM, 2020, p.19 *apud* MCMENAMIN, 2002, p. 119).

Além da resolução do caso Unabomber, uma das contribuições de Fitzgerald tanto para a Linguística Forense, em sua dimensão da atribuição de autoria, quanto para a solução de crimes do FBI, foi a implementação do CTAD (*Communicated Threat Assessment Database*), uma ferramenta que

[...] consiste num banco de dados lexical pesquisável e comparável a partir de parâmetros comportamentais e linguísticos, com todas as comunicações ameaçadoras e criminais e tem se apresentado um eficiente recurso linguístico para solução de crimes que dependem da atribuição de autoria, tendo sido o responsável pela análise que resultou na renúncia de dois assistentes de escritório de uma promotoria de Louisiana, em 2012²⁰.

No site do Serviço de Referência de Justiça Criminal (*National Criminal Justice Reference Service*, em tradução livre), lê-se que:

O CTAD foi desenvolvido como o repositório principal para todas as ameaças comunicadas e outras comunicações orientadas para o crime (COCs) recebidas pelo FBI. O CTAD foi implementado pela Unidade de Análise Comportamental-1 do FBI (BAU-1) do Grupo de Resposta a Incidentes Críticos, que é um componente do Centro Nacional para a Análise de Crimes Violentos da Academia do FBI. O BAU-1 concentra-se em contraterrorismo, avaliação de ameaças e outros serviços linguísticos forenses. Todas as ameaças comunicadas e COCs recebidos no FBI são inseridos no CTAD e analisados por seu potencial de ameaça e autoria. O CTAD foi desenvolvido

²⁰ Disponível em <http://www.justificando.com/2015/07/30/o-papel-da-linguistica-forense-em-uma-investigacao/>. Acesso em 22 dez. 2020.

no início dos anos 2000 como uma forma de lidar efetivamente com o grande aumento nas comunicações ameaçadoras e nos COCs recebidos pela unidade BAU-1 desde meados dos anos 1990. O BAU-1 desenvolveu o CTAD para servir como um mecanismo de busca abrangente capaz de categorizar e classificar comunicações, bem como para fornecer recursos de redação de relatórios. A base de dados é linguística e comportamentalmente orientada para permitir a verificação cruzada de marcadores comportamentais dentro do texto e para permitir a avaliação da ameaça potencial. O CTAD funciona dividindo as comunicações, que variam de várias frases a 20 ou mais páginas cada, em 24 categorias, com o tema geral da comunicação denotando as categorias primária e secundária selecionadas. O CTAD tem quatro seções diferentes - Administrativo, Fatos do caso, Perfil linguístico e Fatos do caso confirmados - e um importante componente de pesquisa. Os planos futuros para o CTAD incluem atualizações contínuas para melhorar sua eficácia. Exemplos de casos específicos auxiliados pelo CTAD são oferecidos junto com um glossário de termos usados no artigo (tradução livre)²¹.

Em entrevista ao Centro de Linguística Forense, em julho de 2009, Fitzgerald descreve que a importância de bancos de dados como o CTAD para o FBI reside na criação de um *corpus* linguístico e comportamental que conecta milhares de comunicações criminosas, recebidas todos os anos. Diferentemente de outros bancos de dados do laboratório da instituição, o CTAD se concentra em palavras funcionais, em questões de formatação e estilística que permitem vincular as comunicações entre si. Segundo Fitzgerald, o CTAD é eficaz e nada serviu tão bem ao FBI, sob sua tutela, quanto a ferramenta.

Idealizador, ainda, do CTARC (*Communicative Threat Assessment Resource Corpus* ou Corpus de recursos de avaliação de ameaças comunicativas, tradução livre), sistema similar ao CTAD, cuja diferença reside em sua aplicabilidade ao âmbito privado, o agente se apresenta como um grande entusiasta da Linguística Forense, acreditando que esta tem um grande futuro em termos de aplicabilidade para resolver casos, podendo solucionar crimes, inclusive anônimos, relacionados ao idioma. Segundo Fitzgerald, “Onde, antes da Linguística Forense, havia apenas palpites sortudos, agora a ciência confirma, o que o torna muito mais preciso e muito mais eficiente”²².

Por ser uma respeitável figura de autoridade no âmbito da Linguística Forense e um vanguardista na criação de um sistema de *corpus* baseado em estilística que tem se mostrado extremamente eficaz para atribuição de autoria e consequente solução de crimes no FBI, o presente estudo traz, em seu Apêndice A, uma entrevista com Fitzgerald.

A entrevista foi realizada por e-mail, enviada no dia 28 de janeiro de 2021 e contendo 5 perguntas que visam contribuir para a discussão, temática do trabalho, acerca da eficácia da

²¹ Disponível em <https://www.ncjrs.gov/App/publications/Abstract.aspx?id=239042>. Acesso em 22 dez. 2020.

²² Disponível em <https://www.forensiclinguistics.net/interviews-fitzgerald.html>. Acesso em 22 dez. 2020.

utilização da Linguística Forense como evidência para atribuição de autoria e solução de crimes, no geral, e de cibercrimes, em particular, trazendo à tona a proposição de uma pretensa inevitabilidade da aplicação de um instituto linguístico no que tange aos crimes cometidos através da linguagem.

Considerando todo o exposto e tendo em vista que Fitzgerald, em resposta a uma das perguntas da referida entrevista, afirmou que o caso Unabomber, se tivesse ocorrido como um cibercrime, poderia ter sido solucionado através da Linguística Forense e do perfilamento criminal, especialmente porque “[...] se seu manifesto estivesse disponível hoje na internet, a busca pelo axioma / provérbio de verbo invertido, "você não pode comer seu bolo e ter também", teria sido muito mais fácil de encontrar e vincular diretamente a ele” (FITZGERALD, James. Entrevista I. [jan. 2021]. Entrevistadora: Jessicka Oliveira de Assis. Juiz de Fora, 2021. 1 arquivo documento Word (21 KB). A entrevista na íntegra encontra-se transcrita no Apêndice B desta monografia), este estudo se debruça, a seguir, sobre a aplicação de técnicas da Linguística Forense para a solução de crimes virtuais sob a luz dos mecanismos de atribuição de autoria do *criminal profiling*.

3.3 A TÉCNICA DO *CRIMINAL PROFILING*

A técnica do *criminal profiling* ou perfilamento criminal pode ser definida como ferramenta da Criminologia, já que é essa ciência que estuda o fenômeno da criminalidade, especialmente no que diz respeito às causas e fatores da formação de um criminoso, e cujo objetivo genérico é o de contribuir não apenas para o entendimento da mentalidade criminosa, mas também do crime em si (ANDREA; FIORENTINO; CAMARGO, 2020).

O perfilamento criminal, mais especificamente, colabora para esta perspectiva por se apresentar como um método investigativo multidisciplinar de evidências físicas e psíquicas que, a partir de estudos teóricos e empíricos, se propõe a indicar prováveis comportamentos de um criminoso (ANDREA; FIORENTINO; CAMARGO, 2020). Assim, o *criminal profiling* cria o perfil psicológico do agente do delito levando em conta fatores biológicos e mesológicos, como seus antecedentes, seu *modus operandi* e seu comportamento após o crime (HAMADA; DO AMARAL, 2008).

Cabe destacar que, no que se refere à utilização de fatores biológicos, o perfilamento genético para a identificação da autoria criminal já é tutelada pela legislação brasileira, como se percebe da leitura da Lei nº 12.654 de 2012; já os fatores mesológicos, ou seja, aqueles que

se referem ao meio no qual viveu e/ou vive o criminoso, não possuem disposição normativa específica, sendo tratados, pela normativa nacional, tímida e esparsamente. É o caso, por exemplo, do art. 12 da Lei nº 13.709, de 14 de agosto de 2018, denominada Lei Geral de Proteção de Dados (LGPD), que dispõe:

Art. 12. Os dados anonimizados não serão considerados dados pessoais para os fins desta Lei, salvo quando o processo de anonimização ao qual foram submetidos for revertido, utilizando exclusivamente meios próprios, ou quando, com esforços razoáveis, puder ser revertido.

§ 2º Poderão ser igualmente considerados como dados pessoais, para os fins desta Lei, aqueles utilizados para **formação do perfil comportamental**²³ de determinada pessoa natural, se identificada (BRASIL, 2018, recurso online, grifo meu).

Dessa forma, percebe-se, hodiernamente, o tratamento residual e subsidiário do perfilamento através da evidência psíquica e mesológica, sendo esta utilizada quando "não se encontram vestígios de sangue e impressões digitais" (BERTOLDI et al., 2014). A esse respeito, o presente estudo pretende se afastar de uma perspectiva precipuamente genética e lombrosiana, considerando que:

A personalidade do criminoso não está apenas no conjunto de traços biológicos e físicos, também se encontra na situação delituosa que é resultante dos contextos e meios sociais, juntamente com os processos cognitivos pessoais, afetivos que levou aquela pessoa interpretar determinada situação de uma forma particular que para ele é propícia para demonstrar seu sentimento de forma criminosa (BERTOLDI, et al., 2014, p. 2)

Realizada tal observação, este estudo pretende analisar mecanismos de atribuição de autoria do *criminal profiling*, na seara do perfilamento mesológico, que possam ser aplicados pela Linguística Forense para solucionar os crimes cibernéticos cometidos única ou parcialmente através da linguagem.

Fitzgerald (2021), sobre o assunto, afirma que, no perfilamento realizado pelas ferramentas CTARC e CTAD,

[...] havia também um componente em cada um que permitia pesquisar um perfil linguístico de características demográficas específicas. Isso seria realizado na tentativa de determinar quais comunicações ainda anônimas podem ter sido autoras por sexo, idade, raça / etnia, ocupação, idioma nativo, etc. Estas não seriam comprobatórias nem depoimentos nos tribunais, mas sim usadas durante os estágios iniciais de uma investigação para determinar um perfil básico de um autor anônimo em uma tentativa de ajudar os

²³ A LGPD, neste ponto, trata da utilização de dados pessoais estratégicos para a objetificação de usuários de meios digitais: a partir do delineamento do perfil comportamental, especialmente consumerista, do usuário, através do que se convencionou chamar de *Big Data* - grandes centros de dados pessoais coletados por algoritmos de fontes diversas nos meios tecnológicos, exibem-se propagandas e anúncios específicos para aquele sujeito. O assunto não é parte do escopo desta pesquisa, possuindo caráter meramente ilustrativo para a temática do perfilamento psicológico.

investigadores a restringir seu campo suspeito do UNSUB (assunto desconhecido). Havia também uma categoria em cada corpus de “Antecedentes do Autor confirmados”, na qual, uma vez identificado e admitido como autor um autor anônimo, esses achados também eram incluídos nos respectivos corpora. Em minha própria experiência, raramente os perfis mudam, ou se sim, não dramaticamente (FITZGERALD, James. Entrevista I. [jan. 2021]. Entrevistadora: Jessicka Oliveira de Assis. Juiz de Fora, 2021. 1 arquivo documento Word (21 KB). A entrevista na íntegra encontra-se transcrita no Apêndice B desta monografia).

Dessa forma, tendo em conta a possibilidade de uma aplicação do *criminal profiling* pela Linguística Forense, a seguir, trata-se da temática, denominada, nesta pesquisa, pela nomenclatura “perfilamento linguístico”.

3.3.1 Perfilamento linguístico

O perfilamento criminal tem por base, portanto, para além da análise biológica do material genético, a análise de elementos impressos em características subjetivas do delito, como o nexos causal e o aspecto instalado na cena do crime consumado. Não à toa, o perfilamento mesológico deve se pautar nas seguintes perspectivas: a) personalidade do autor; b) significado do ato; e, c) o contexto do ato (ANDRÉA et al., 2020). Diante de cibercrimes cometidos única ou parcialmente através da linguagem, no entanto, especialmente quando a língua é a única fonte de evidência, faz-se necessário analisar todos esses aspectos sob uma perspectiva linguística.

No que se refere ao nexos causal, à cena do crime consumado e ao contexto do ato, é essencial que, num ambiente virtual, sejam analisados elementos objetivos do delito que podem apontar para características pessoais (idade, *hobbies*, gostos pessoais etc) do criminoso, conforme foi discutido na seção 2.2, como a rede social na qual foi encontrada a evidência do crime e o nome do usuário, ainda que falso, relacionado à postagem. Nessa perspectiva, a Forense Computacional, cujo objetivo é se utilizar de ferramentas computacionais para a recuperação de rastros digitais, conta com ferramentas automáticas - como o *Machine Learning*, que treina algoritmos sistematicamente a fim de analisar dados de reconhecimento - para a detecção dos cibercrimes (MALLMANN; DOS SANTOS XAVIER; SANTIN, 2018).

No que se refere à verificação da personalidade do autor e do significado do ato, uma das ferramentas utilizadas pelo *criminal profiling* são os testes psicológicos, técnicas de função privativa dos psicólogos, nos termos da Lei nº 4.119 de 1962. Os testes psicológicos ou testes

de personalidade utilizados pelo Instituto Nacional de Criminalística (INC) da Polícia Federal brasileira se enquadram em três categorias: (i) os projetivos, que medem a personalidade através de figuras e jogos, como é o caso do Teste de Rorschach; (ii) os prospectivos, que se propõem a conhecer profundamente o indivíduo, traçando com precisão sua personalidade; (iii) e os testes de inteligência, que medem o denominado quociente de inteligência (Q.I.) do criminoso, a fim de determinar seu estado mental. Todos esses tipos de testes periciais buscam detalhar a personalidade do criminoso, seu grau de imputabilidade, periculosidade, sensibilidade à pena e probabilidade de correção (HAMADA; DO AMARAL, 2008).

Além disso, os depoimentos das vítimas são essenciais à investigação, por "oferecerem uma visão de como é o comportamento do autor [...], bem como pistas de como este atua" (HAMADA; DO AMARAL, 2008, apud WINERMAN, 2004, p. 66). No entanto, como já anteriormente exposto, não escapa à apreciação da Linguística Forense o fato de que o depoimento do próprio autor seja a melhor pista de sua atuação e personalidade.

Nesse ínterim, tem sido largamente demonstrado pela Sociolinguística que "um mesmo indivíduo utiliza a língua diferentemente, dependendo de vários fatores, como tema, situação, destinatário, registro, gênero textual etc", de forma que textos pertencentes a um mesmo autor, especialmente quando diante da congruência de alguns desses fatores, apresentam a mesma variação intrafalante (DE ALMEIDA, 2017, p. 2).

A variação intrafalante ou idioleto representa o estilo individual de utilização linguística de determinado sujeito e tende a ser menor, na medida em que o conteúdo se torna mais profundo; *a contrario sensu*, quanto mais simples e abstrato o conteúdo, menos escolhas existem para o usuário da língua, o que "pode resultar em maior chance de que se façam sempre as mesmas escolhas, isto é, de organizar os textos segundo as mesmas categorias" e "Se, com essa grande chance de realizar sempre as mesmas escolhas, dois indivíduos distintos escolhem recorrentemente as mesmas opções, mas diferentemente um do outro, isso significa que a opção de cada um tem grande poder discriminatório" (DE ALMEIDA, 2017, p. 2).

Isso ocorre porque a noção de estilo está no cerne das investigações acerca de autoria, já que essa definição comporta um conjunto global de traços recorrentes do plano do conteúdo (formas discursivas) e do plano da expressão (formas textuais) que, ao fim, produzem um efeito de sentido de identidade, configurando um *ethos* discursivo, ou seja, uma imagem do enunciador. Portanto, o estilo é a "recorrência e distintividade (diferenciação), tanto no plano da expressão, quanto do conteúdo" (DE ALMEIDA, 2017, p. 3).

Diante do exposto, a principal técnica que se propõe a ser utilizada pelo perfilamento linguístico para atribuição de autoria, tendo em conta o manejo do *criminal profiling* diante da

análise subjetiva pela Linguística Forense da personalidade do criminoso e do significado de seu ato, é a análise de *corpus* de expressão e conteúdo, seja esta oral ou escrita.

Fitzgerald (2021), sobre a técnica, afirma que:

Como os bancos de dados locais, nacionais e internacionais existentes de impressões digitais e amostras de DNA, cada atribuição a uma pessoa específica, é apenas uma questão de tempo até que as corpora de linguagens conhecidas não apenas existam (como já existem), mas também sejam mais e mais utilizados por agências governamentais oficiais na tentativa de identificar os autores de comunicações questionadas. Em vez de investigadores e / ou linguistas forenses pesquisarem corpora separados para características de escrita ou padrões de fala específicos e demonstráveis para fins de comparação, um dia pode haver um corpus totalmente pesquisável disponível para aqueles interessados em coletar dados através da linguagem. Será um corpus combinado de todos, ou da maioria dos corpora existentes. Pode ser ou não um idioma de origem pública, mas também e-mails, textos e mensagens de mídia social particulares (FITZGERALD, James. Entrevista I. [jan. 2021]. Entrevistadora: Jessicka Oliveira de Assis. Juiz de Fora, 2021. 1 arquivo documento Word (21 KB). A entrevista na íntegra encontra-se transcrita no Apêndice B desta monografia).

Não à toa, o autor afirma, ainda, que o que, para os fins dessa pesquisa, convencionou-se chamar “perfilamento linguístico” já vem sendo utilizado nos Estados Unidos em cibercrimes cometidos única ou parcialmente através da Linguagem. Fitzgerald, no entanto, chama atenção para o fato de que:

[...] costumava ensinar que os envelopes de transmissão tradicionais em si mesmos continham sua própria forma de evidência linguística, certamente de uma perspectiva estilística, incluindo formatação, espaçamento, etc. Suponho que em crimes cibernéticos como os listados acima, um obstáculo envolveria a função de correção automática disponível em programas de processamento de texto. Obviamente, se uma pessoa tinha o hábito de escrever certas palavras incorretamente, elas seriam aparentemente corrigidas antes que a comunicação fosse enviada. Mas, como qualquer linguista forense sabe, existem muitos meios alternativos de determinar a autoria comum potencial, além de ortograficamente, ou mesmo gramaticalmente, em que a função de autocorreção normalmente executaria suas tarefas atribuídas (FITZGERALD, James. Entrevista I. [jan. 2021]. Entrevistadora: Jessicka Oliveira de Assis. Juiz de Fora, 2021. 1 arquivo documento Word (21 KB). A entrevista na íntegra encontra-se transcrita no Apêndice B desta monografia).

Considerando a hipótese delineada para esta pesquisa e apresentada a importância e pertinência da utilização da Linguística Forense, em especial da análise de *corpus*, para a solução de cibercrimes, tendo em vista, ainda, suas limitações, busca-se debater, a seguir, a forma pela qual o fenômeno tem sido encarado no contexto nacional, especialmente pela literatura especializada.

4 A LINGUÍSTICA FORENSE E OS CIBERCRIMES NO BRASIL

Diante do conjunto de possibilidades de técnicas e mecanismos de atribuição de autoria manejados pela Linguística Forense articulada com o *criminal profiling* e considerando o cenário nacional de entraves à identificação de cibercriminosos, este tópico se dedica a investigar como a literatura especializada brasileira trata o tema da Linguística Forense aplicada a cibercrimes cometidos única ou parcialmente através da linguagem.

Para tanto, analisam-se, no recorte de 2014 a 2020, três bancos de dados sobre a temática: a revista *Language and Law*, especializada na interface homônima (*Linguagem e Direito*); a *Revista Brasileira de Ciências Penais*, especializada em ciências penais e áreas afins; e o banco de teses e dissertações dos periódicos da CAPES, um dos mais importantes veículos de divulgação de trabalhos acadêmicos no Brasil.

O recorte temporal foi escolhido com base no início da vigência do Marco Civil da Internet, por ser esta a normativa legal regulatória de ciberconteúdos no Brasil; os bancos para coleta de dados foram eleitos por serem grandes referências em suas respectivas áreas, que devem comportar, juntas, uma análise da leitura que a literatura brasileira especializada faz sobre a temática; o levantamento foi feito com base na busca de três palavras-chave no título dos artigos, quais sejam “cibercrimes”, “Linguística Forense” e “*criminal profiling*”, por serem estes considerados os termos mais intimamente relacionados ao escopo deste estudo. O recurso metodológico para a análise dos dados encontrados, por fim, será aquele utilizado por Rezende (2020), em “Avaliação em larga escala da alfabetização: os casos de Paebe Alfa, Proalfa e Spaece Alfa”, consistindo na apresentação, em tópicos distintos para cada banco de dados, dos números de produções encontradas sobre o tema, em forma de tabela, e, quando encontrados, da explanação descritiva do estudo.

4.1 LANGUAGE AND LAW

Language and Law é uma revista internacional bilíngue (Português e Inglês) de publicação semestral, cujos trabalhos publicados, normalmente, dizem respeito à interface Direito e Linguagem.

A tabela a seguir apresenta os resultados da análise deste banco de dados, a partir da qual pode-se inferir que, durante o recorte em estudo, foram publicados 97 estudos, organizados em 13 volumes, nenhum sobre as temáticas cibercrimes e *criminal profiling* e apenas 3 (o que

corresponde a 3,09% do total) sobre Linguística Forense. Destaca-se que a pesquisa automática somente é permitida no sítio da revista a partir do ano de 2016, motivo pelo qual a investigação dos dois anos anteriores foi feita manualmente.

Tabela 1 - Mapeamento da revista Language and Law

Número	Vol.	Ano	Nº de Artigos da revista	Nº de Artigos sobre Cibercrimes	Nº de Artigos sobre <i>criminal profiling</i>	Nº de Artigos sobre Linguística Forense
1	1	2014	10	0	0	0
2	1	2014	6	0	0	0
1	2	2015	6	0	0	0
2	2	2015	5	0	0	0
1	3	2016	8	0	0	0
2	3	2016	10	0	0	0
1	4	2017	7	0	0	1
2	4	2017	9	0	0	0
1	5	2018	6	0	0	0
2	5	2018	6	0	0	0
1	6	2019	7	0	0	0
2	6	2019	6	0	0	2
1-2	7	2020	11	0	0	0
TOTAL			97	0	0	3

Fonte: Language and Law. Tabulação Própria.

A seguir, são analisados os conteúdos referentes aos 3 artigos encontrados sobre a temática da Linguística Forense.

“Introducción a la Lingüística Forense – un libro de curso”, de John Gibbons (2017), trata das evidências de comunicação no sistema jurídico dos Estados Unidos, especialmente no que diz respeito ao grande número de migrantes hispânicos na Califórnia moderna e a consequente perseguição xenofóbica contínua e difamação por eles sofridas. O estudo se debruça sobre o conceito de idioleto e, apesar de não mencionar os cibercrimes ou o *criminal*

profiling, aborda brevemente a possibilidade da detecção da autoria em crimes a partir da utilização de uma análise de estilo informática, rechaçando-a.

“Da “intimidação intencional” aos “ciúmes excessivos”: uma análise linguística forense das Fichas de Avaliação de Risco”, de Ana Sofia Ferreira e Rui Sousa-Silva (2019), destina-se à análise de problemas de ordem sintática, semântica, lexical/terminológica e pragmático-discursiva de Fichas de Avaliação de Risco (utilizadas, em Portugal, como instrumento de denúncia de atos de violência doméstica) e que podem comprometer a eficácia do instrumento, principalmente no que se refere à análise inadequada de risco e, conseqüentemente, de risco para a vítima, comprometendo sua segurança.

“Linguagem E Direito: Caminhos para Linguística Forense”, de Débora de Carvalho Figueiredo (2019), reúne uma coleção de capítulos produzidos por linguistas e juristas interessados na interface entre linguagem e direito, abordando temáticas relacionadas, em grande parte, à análise do discurso e tratando temas como a investigação de práticas linguísticas e discursivas dos gêneros Judiciário, plágio e publicação em mídia.

Dessa forma, os poucos trabalhos que mobilizam (em maior ou menor grau) o tema da Linguística Forense não fazem referência aos cibercrimes ou ao *criminal profiling*.

4.2 REVISTA BRASILEIRA DE CIÊNCIAS PENAIS

A Revista Brasileira de Ciências Penais é divulgada semestralmente, por via impressa e eletrônica, desde 2006, abordando estudos sobre a temática penal e áreas afins. Por ter passado, no entanto, por importantes modificações quanto a critérios de qualificação, no ano de 2019, a revista conta com seis volumes: 2006, 2007, 2010²⁴, 2019 e dois volumes em 2020.

A tabela a seguir apresenta os resultados da análise dos últimos três volumes deste banco de dados, a partir da qual pode-se inferir que, durante o recorte em estudo, foram publicados 35 estudos, organizados em 3 volumes. Destaca-se que a pesquisa automática somente não é permitida no sítio da revista, motivo pelo qual a investigação foi feita manualmente.

²⁴ Os anos de 2006, 2007 e 2010 não foram analisados por não fazerem parte do recorte temporal eleito neste estudo.

Tabela 2 - Mapeamento da Revista Brasileira de Ciências Penais

Vol.	Ano	Nº de Artigos da revista	Nº de Artigos sobre Cibercrimes	Nº de Artigos sobre <i>criminal profiling</i>	Nº de Artigos sobre Linguística Forense
4	2019	14	0	0	0
5	2020	12	0	0	0
6	2020	9	0	0	0
TOTAL		35	0	0	0

Fonte: Revista Brasileira de Ciências Penais. Tabulação Própria.

Dessa forma, verifica-se que nenhum estudo publicado neste banco de dados diz respeito às temáticas aqui pesquisadas, quais sejam cibercrimes, *criminal profiling* e Linguística Forense.

4.3 CATÁLOGO DE TESES E DISSERTAÇÕES DA CAPES

A pesquisa automática na plataforma CAPES foi realizada por assunto, na busca avançada, a partir de artigos com data de publicação nos últimos 7 (sete) anos, a contar de 2021, em qualquer idioma, de 01/01/2014 a 31/12/2020, sem seleção de base específica e contendo, no título, separada e conjuntamente, as palavras-chave selecionadas.

Durante o recorte em estudo, a busca realizada separadamente por palavra-chave na base de dados retornou nula quanto aos cibercrimes; 25 (vinte e cinco) resultados para *criminal profiling* e 3 (três) resultados para Linguística Forense. Cabe destacar, no entanto, que a busca das palavras-chave, conjuntamente, no título, retornou nula nas três hipóteses: Linguística Forense e *criminal profiling*; Linguística Forense e cibercrimes; cibercrimes e *criminal profiling*.

Diante dos resultados apresentados, é possível inferir que não existem estudos publicados sobre a temática dos cibercrimes, no período em estudo; que existem esparsos

estudos publicados acerca do *criminal profiling* e ainda mais esparsos sobre a Linguística Forense. Nenhum deles, no entanto, trata da aplicação da Linguística Forense a cibercrimes, da utilização das técnicas do *criminal profiling* para a solução de cibercrimes ou mesmo da articulação destas aos mecanismos de atribuição de autoria da Linguística Forense.

Dessa forma, diante dos dados apresentados, verifica-se que o cenário jurídico brasileiro é incipiente no que se refere à Linguística Forense, aos cibercrimes e ao *criminal profiling*, em diferentes graus, o que faz, conseqüentemente, com que seja ainda menos explorada a interface entre os assuntos. Tendo em conta os dados coletados, conclui-se, ainda, que não foram encontrados dados de aplicação da Linguística Forense a cibercrimes, nem mesmo sob a alcunha de outro instituto, não havendo, inclusive, elemento que aponte para a utilização, no Brasil, de técnicas específicas para a identificação de autoria em tais delitos.

CONSIDERAÇÕES FINAIS

A partir da observação dos números de cibercrimes cometidos única ou parcialmente através da linguagem nos últimos anos, é possível inferir que o ordenamento jurídico, diante de sua insuficiência e das características dessa nova modalidade criminosa, tem encontrado dificuldades para refrear esse tipo de conduta, especialmente no que se refere à responsabilização do criminoso, já que, na seara da identificação da autoria, o anonimato oferecido pela cultura cibernética se torna um entrave para a solução de tais delitos.

Nessa perspectiva, ante a manifesta necessidade de uma ciência específica para os crimes de tal natureza, a possibilidade (e necessidade) de se utilizar a língua como evidência, apresenta a Linguística Forense como um desenlace ao atual cenário ciberdelitivo nacional. Considerando a ausência de demais evidências e as técnicas mesológicas já utilizadas pelo *criminal profiling* para se traçar a personalidade de um criminoso e o significado de seu ato delitivo, é possível conceber uma modalidade de perfilamento que se baseie no fenômeno linguístico forense: o perfilamento linguístico, pautado numa análise de *corpus* que vise identificar o estilo linguístico do sujeito, identificando-o.

A despeito das limitações da técnica, a possibilidade oferecida pela Linguística Forense, a partir do manejo de técnicas do *criminal profiling*, de se atribuir autoria ao cibercriminoso pode se apresentar como a melhor e mais viável opção para o cenário brasileiro. No entanto, a pesquisa realizada no âmbito deste trabalho comprovou, à guisa de considerações parciais, tendo em conta a limitação da própria natureza deste estudo, a hipótese de incipiência jurídica no que se refere às três temáticas, demonstrando que, a despeito da viabilidade da aplicação do instituto, esta não ocorre, nem mesmo sob a alcunha de outro método.

REFERÊNCIAS

Análise Forense Digital: Conceitos e modelos. Disponível em https://www.gta.ufrj.br/grad/13_1/forense/terminologia.html#evidencia. Acesso em 04 dez. 2020.

ANDRÉA, Gianfranco Faggini Mastro; FIORENTINO, Henrique Alencar; CAMARGO, Juliana Moreira. **Criminal profiling: análise teórica e prática de um novo instrumento de segurança jurídica para o Brasil.** CRIMES E SOCIEDADE EM DEBATE, p. 66.

AZZARITI, Mônica. **O papel da Linguística Forense em uma investigação.** Justificando, 2015. Disponível em <http://www.justificando.com/2015/07/30/o-papel-da-linguistica-forense-em-uma-investigacao/>. Acesso em 22 dez. 2020.

BERTOLDI, Maria Eugênia et al. **Perfilamento e crime: uma abordagem psicocriminológica.** JICEX, v. 4, n. 4, 2014.

BRASIL. Constituição da República Federativa do Brasil de 1988. Disponível em http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em 03 dez. 2020.

BRASIL. DECRETO-LEI No 2.848, DE 7 DE DEZEMBRO DE 1940. Código Penal. Disponível em http://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm. Acesso em 03 dez. 2020.

BRASIL. DECRETO-LEI Nº 3.689, DE 3 DE OUTUBRO DE 1941. Código de Processo Penal. Disponível em http://www.planalto.gov.br/ccivil_03/decreto-lei/del3689compilado.htm. Acesso em 03 dez. 2020.

BRASIL. LEI Nº 4.119, DE 27 DE AGOSTO DE 1962. Dispõe sobre os cursos de formação em psicologia e regulamenta a profissão de psicólogo. Disponível em http://www.planalto.gov.br/ccivil_03/leis/1950-1969/14119.htm#:~:text=Disp%C3%B5e%20s%C3%B4bre%20os%20cursos%20de,regulame%20a%20profiss%C3%A3o%20de%20psic%C3%B3logo.&text=Art.,de%20bacharelado%20C%20licenciado%20e%20Psic%C3%B3logo. Acesso em 14 de fev. de 2021

BRASIL. LEI Nº 12.654, DE 28 DE MAIO DE 2012. Altera as Leis nºs 12.037, de 1º de outubro de 2009, e 7.210, de 11 de julho de 1984 - Lei de Execução Penal, para prever a coleta de perfil genético como forma de identificação criminal, e dá outras providências. Disponível em http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/112654.htm. Acesso em 14 de fev. de 2021

BRASIL. LEI Nº 12.735, DE 30 DE NOVEMBRO DE 2012. Disponível em http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2012/Lei/L12735.htm. Acesso em 03 dez. 2020.

BRASIL. LEI Nº 12.737, DE 30 DE NOVEMBRO DE 2012. Disponível em http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/112737.htm. Acesso em 03 dez. 2020.

BRASIL. LEI Nº 13.709, DE 14 DE AGOSTO DE 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). (Redação dada pela Lei nº 13.853, de 2019). Disponível em http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Acesso em 14 de fev. de 2021

BRASIL. Ministério da Justiça. Marco civil da Internet: seus direitos e deveres em discussão. Disponível em: <<http://culturadigital.br/marcocivil/>>. Acesso em: 23 abr. 2014.

CALDAS-COULTHARD, C. R. **ReVEL na Escola: o que é a Linguística Forense?** ReVEL, 2014(12) 23.

Centre for Forensic Linguistics. CFL Interviews. **James R. Fitzgerald**, 2009. Disponível em <https://www.forensiclinguistics.net/interviews-fitzgerald.html>. Acesso em 22 dez. 2020.

COMITÊ GESTOR DA INTERNET NO BRASIL. Centro de Estudos, Resposta e Tratamento de Incidentes de Segurança no Brasil. **Cartilha de segurança para internet**. 2. ed. São Paulo: Comitê Gestor da Internet no Brasil, 2012.

CONVENÇÃO DE BUDAPESTE SOBRE O CIBERCRIME DE NOVEMBRO DE 2001, em vigor em 01 de julho de 2004. Disponível em <http://conventions.coe.int>. Status as of: 03/12/2020.

COULTHARD, Malcolm. **Linguística Forense: uma entrevista com Malcolm Coulthard**. ReVEL, vol. 12, n. 23, 2014. Tradução João Gabriel Rodrigues Marques Padilha. [www.revel.inf.br].

CRESPO, Marcelo Xavier de Freitas. **O cibercrime**. São Paulo: Saraiva, 2011.

DE ALMEIDA, Dayane Celestino. Análise forense de autoria textual: estilos sociais e individuais. **Language and Law/Linguagem e Direito**, v. 2, n. 2, 2017.

DE CARVALHO FIGUEIREDO, Débora. **Linguagem & Direito: Caminhos para Linguística Forense**. **Language and Law/Linguagem e Direito**, v. 6, n. 2, p. 109-112, 2020.

DIAS, Vera Elisa Marques. **A problemática da investigação do cibercrime**. Universidade de Lisboa, Lisboa, 2010. Disponível em http://www.verbojuridico.com/doutrina/2011/veradias_investigacaocibercrime.pdf. Acesso em 02 dez. 2020.

EXAME. **Brasil é o 4º país em número de usuários de internet, 2017**. Agência Brasil. Disponível em <https://exame.com/tecnologia/brasil-e-o-4o-pais-em-numero-de-usuarios-de-internet/>. Acesso em 02 dez. de 2020

FAIRCLOUGH. **Discurso e mudança social**. Brasília: Editora Universidade de Brasília, 2001, 2008 (reimpressão).

FERREIRA, Ana Sofia; SOUSA-SILVA, Rui. **Da “intimidação intencional” aos “ciúmes excessivos”:** uma análise linguística forense das Fichas de Avaliação de Risco. **Language and Law/Linguagem e Direito**, v. 6, n. 2, p. 65-89, 2020.

FERREIRA, Deyvid Braga; BARROS, Lívyra Ramos Sales Mendes de. **Ideologia e formações ideológicas de dominação e subserviência: um estudo da sumula vinculante nº 11 do STF**. Faculdade Raimundo Marinho. In: COULTHARD, Malcolm; COLARES, Virgínia; SOUSA-SILVA, Rui. **Linguagem e Direito: os eixos temáticos**. Recife: ALIDI, 2015.

GIBBONS, John. **Introducción a la Lingüística Forense—un libro de curso**. **Language and Law/Linguagem e Direito**, v. 4, n. 1, p. 165-174, 2017.

GLENDAY, CANDICE HELEN. **Noam Chomsky: Lingüística e Filosofia**. 2008. Tese de Doutorado. Dissertação de Mestrado. UEFN-Universidade Estadual do Norte Fluminense.

HAMADA, Fernando Massami; DO AMARAL, José Hamilton. **Perfilamento criminal como ferramenta criminológica**. ETIC-ENCONTRO DE INICIAÇÃO CIENTÍFICA-ISSN 21-76-8498, v. 4, n. 4, 2008.

JUSBRASIL. **Ameaça**. Disponível em <https://www.jusbrasil.com.br/topicos/26413060/ameaca>. Acesso em 27 jan. 2021.

JUSBRASIL. **Brasil registra 54 crimes virtuais por minuto.** Associação do Ministério Público de Minas Gerais. Disponível em <https://amp-mg.jusbrasil.com.br/noticias/3125198/brasil-registra-54-crimes-virtuais-por-minuto>. Acesso em 02 dez. de 2020

KUNRATH, Josefa Cristina Tomaz Martins. **A expansão da criminalidade no ciberespaço, 2017.** Feira de Santana: Universidade Estadual de Feira de Santana.

LÉVY, Pierre. **Cibercultura.** Tradução Carlos Irineu da Costa. São Paulo: Editora 34, 1999. Disponível em http://www.moodle.ufba.br/file.php/8897/levy_cibercultura.pdf. Acesso em: 2 dez. 2020.

LUCENA, Mariana Barrêto Nóbrega de. **O desvio social na rede mundial de computadores: aspectos sociológicos e psicológicos dos indivíduos pertencentes às subculturas criminais da internet.** Jus Navigandi, Teresina, ano 17, n. 3128, 24 jan. 2012. Disponível em <http://jus.com.br/revista/texto/20921>. Acesso em 03 dez. 2020.

MALLMANN, Jackson; DOS SANTOS XAVIER, Alex; SANTIN, Altair Olivo. **Deteção de Cibercrime em Redes Sociais: Machine Learning.** 2018.

MONTEIRO NETO, João Araújo. **Aspectos constitucionais e legais do crime eletrônico.** 2008. M775a. 191 f. Disponível em www.dominiopublico.gov.br. Acesso em: 03 dez. 2020.

MORÉ, Bruno. **O crime de ameaça na internet! Intimidar alguém impondo-lhe temor de sofrer mal injusto e grave é crime de ameaça.** Ex: jurar alguém de morte. JUSBRASIL. Disponível em <https://brunomore.jusbrasil.com.br/artigos/608243055/o-crime-de-ameaca-na-internet>. Acesso em 03 dez. 2020

National Criminal Justice Reference Service. **Virtual Library Search.** U.S Department of Justice. Office of Justice Programs. Disponível em <https://www.ncjrs.gov/App/publications/Abstract.aspx?id=239042>. Acesso em 22 dez. 2020.

NORTON. Disponível em <https://br.norton.com/norton-cybersecurity-insights-report-brazil>. Acesso em 02 dez. de 2020

PEREIRA E SILVA, Welton. **Linguística Forense: como o linguista pode contribuir em uma demanda judicial?** Roseta. Disponível em <http://www.roseta.org.br/pt/2019/09/16/linguistica-forense-como-o-linguista-pode-contribuir-em-uma-demanda-judicial/>. Acesso em 04 dez. 2020

REZENDE, W. S. **As relações entre direito, política e sociedade: retórica e teoria da ação na análise da argumentação em casos difíceis no Supremo Tribunal Federal brasileiro.** 2014. 523 f. 2014. Tese de Doutorado. Programa de Pós-Graduação em Ciências Sociais da Universidade Federal de Juiz de Fora.

REZENDE, W.S. **AValiação em Larga Escala da Alfabetização: os casos de Paebes Alfa, Proalfa e Spaece Alfa,** 2020. Tese de Doutorado. Programa de Pós-Graduação em Educação da Universidade Federal de Juiz de Fora.

SAFERNET. **32.579 ATENDIMENTOS E 4.291.500 DENÚNCIAS.** Datasafet. Disponível em <https://indicadores.safernet.org.br/indicadores.html>. Acesso em 02 dez. de 2020

SAFERNET. **Calúnia / difamação. CRIMES NA WEB.** Disponível em <https://new.safernet.org.br/content/cal%C3%BAnia-difama%C3%A7%C3%A3o>.

SANTOS BOMFIM, Amanda. **Linguística Forense: origem, aplicabilidade e seu papel na captura do assassino em série Unabomber.** Monografia (curso de Graduação em Direito da

Pontifícia Universidade Católica de Minas Gerais). Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte (2020).

TABELA DE CLASSIFICAÇÃO DE CRIMES. Justiça & Polícia. Disponível em <https://juspol.com.br/tabela-de-classificacao-de-crimes/>. Acesso em 03 dez. 2020

VIANNA, Túlio Lima. **Fundamentos de direito penal informático.** Rio de Janeiro: Forense, 2003

WIKIPEDIA. “**Hacker**”. Disponível em <https://pt.wikipedia.org/wiki/Hacker>. Acesso em 01 de mar de 2021.

APÊNDICE A - ENTREVISTA COM JAMES FITZGERALD²⁵

1. Depois de tantos anos trabalhando com a polícia e com o FBI, você diria que enxerga a Linguística Forense como a melhor ou mesmo como uma inevitável fonte de evidência para a solução de crimes que são cometidos através da linguagem, ainda que existam outras provas?

Cada investigação é diferente, por isso é impossível dizer que uma análise linguística forense (FL) da linguagem falada ou escrita em um crime seria a "melhor", ou mesmo "inevitável" para sua solução. Dependeria dos fatos do caso e das provas que estariam à disposição dos investigadores e, eventualmente, dos tribunais.

Como agente aposentado do FBI e investigador policial antes disso, agora quase sempre diria que evidências diretas, como impressões digitais ou DNA, são as melhores evidências para se ter em um caso em termos de vinculação de alguém a uma fonte de escrita questionável. É difícil contestar as descobertas quantitativas relacionadas a impressões digitais e DNA. Mas, na ausência desse tipo de evidência direta, a análise da FL pode provar ser o método mais confiável de determinar a autoria e, posteriormente, apresentá-la como testemunho aos julgadores dos fatos em uma questão criminal ou civil - se assim for avaliada, é claro, por um linguista forense qualificado e experiente.

2. Em sua entrevista para a Newsy, disponível no YouTube²⁶, você disse considerar a análise de *corpus* o futuro da Linguística Forense. Considerando ser uma ciência em crescimento, num século informatizado e tecnológico, no qual cada vez mais crimes são cometidos pela via virtual, como você acredita que a tecnologia pode contribuir para o desenvolvimento da Linguística Forense?

Com os bancos de dados locais, nacionais e internacionais existentes de impressões digitais e amostras de DNA, cada atribuição a uma pessoa específica, é apenas uma questão de tempo até que as corpora de linguagens conhecidas não apenas existam (como já existem), mas também sejam mais e mais utilizados por agências governamentais oficiais na tentativa de identificar os autores de comunicações questionadas.

Em vez de investigadores e/ou linguistas forenses pesquisarem corpora separados para características de escrita ou padrões de fala específicos e demonstráveis para fins de

²⁵ Esse apêndice contém a tradução livre das respostas de Fitzgerald, contidas, na íntegra, no apêndice seguinte, no qual encontram-se, ainda, as versões das perguntas aqui formuladas.

²⁶ Disponível em <https://www.youtube.com/watch?v=iOPb-rXhacs>. Acesso em 30 dez. 2020.

comparação, um dia pode haver um corpus totalmente pesquisável disponível para aqueles interessados em coletar dados através da linguagem. Será um corpus combinado de todos, ou da maioria dos corpora existentes. Pode ser ou não um idioma de origem pública, mas também e-mails, textos e mensagens de mídia social particulares. Eu teria uma preocupação menor se essas buscas de corpus fossem realizadas por meio de intimações governamentais ou mandados de busca, mas não gostaria que esse sistema fosse acessado (abusado?) sem a autoridade judicial adequada e vários controles e balanços. Caso contrário, seria aquela proverbial ladeira escorregadia de cidadãos perdendo privacidade para um Grande Governo exagerado. Isso pode ser problemático para uma sociedade livre a curto e longo prazo.

(Só para ficar claro quanto ao acima ... um corpus de comunicações ameaçadoras, criminosas e / ou de assédio, como as encontradas no CTAD do FBI, não constituiria uma violação dos direitos dos escritores / palestrantes cujas comunicações são localizadas nele. As comunicações do CTAD são em sua maioria anônimas e enviadas de uma parte para outra com fins criminosos. Na verdade, foram tornadas públicas pelo remetente na prática de um crime, portanto, seus direitos de privacidade não são um fator aqui).

3. A eficácia da utilização da Linguística Forense para a atribuição de autoria em crimes cometidos não virtualmente tem se mostrado cada vez mais flagrante. Na sua opinião e tendo em vista sua vasta experiência na área policial e na seara linguística, a Linguística Forense poderia se mostrar útil à atribuição de autoria e consequente solução de cibercrimes? Nos EUA, é isso que já acontece? Considerando os entraves que você mesmo encontrou para a aplicação da Linguística Forense, você acredita que, nessa seara, a ciência encontraria mais obstáculos à sua aplicação do que na seara não virtual?

Se por “cibercrimes” você quer dizer linguagem veiculada pela internet e / ou celular, sim, análises desse tipo de comunicação já estão ocorrendo nos Estados Unidos. Na verdade, é muito raro hoje em dia para mim, em minha experiência como linguista forense, ter cartas escritas à mão, ou mesmo cartas digitadas, para avaliar. Quase sempre são e-mails, entradas de blog, mensagens de mídia social ou semelhantes.

Eu costumava ensinar que os envelopes de transmissão tradicionais em si mesmos continham sua própria forma de evidência linguística, certamente de uma perspectiva estilística, incluindo formatação, espaçamento, etc. Mas, como ameaças ou assédio ou comunicações criminais são tão raramente enviadas pelo correio hoje, os envelopes de transmissão como prova são quase uma questão discutível.

Suponho que em crimes cibernéticos como os listados acima, um obstáculo envolveria a função de correção automática disponível em programas de processamento de texto. Obviamente, se uma pessoa tinha o hábito de escrever certas palavras incorretamente, elas seriam aparentemente corrigidas antes que a comunicação fosse enviada. Mas, como qualquer linguista forense sabe, existem muitos meios alternativos de determinar a autoria comum potencial, além de ortograficamente, ou mesmo gramaticalmente, em que a função de autocorreção normalmente executaria suas tarefas atribuídas.

4. Enquanto idealizador da CTAD e da CTARC e perfilador do FBI, você acredita que tais ferramentas se embasam e se utilizam dos mecanismos e técnicas do perfilamento criminal? Como essas ferramentas vêm sendo utilizadas? Essas ferramentas (CTAD e CTARC) já foram utilizadas ou há previsão de que sejam utilizadas na seara dos cibercrimes?

Embora o CTAD do FBI e o Academy Group, Inc's CTARC (agora extinto) tivessem categorias de "perfil" totalmente pesquisáveis incorporadas a eles, eles eram separados e distintos de qualquer pesquisa tradicional de palavra, pontuação ou mesmo classes gramaticais que seria realizado. Essas pesquisas tradicionais sempre foram a função primária dos dois corpora diferentes que desenvolvi com uma década de diferença.

No entanto, havia também um componente em cada um que permitia pesquisar um perfil linguístico de características demográficas específicas. Isso seria realizado na tentativa de determinar quais comunicações ainda anônimas podem ter sido autoras por sexo, idade, raça / etnia, ocupação, idioma nativo, etc. Estas não seriam comprobatórias nem depoimentos nos tribunais, mas sim usadas durante os estágios iniciais de uma investigação para determinar um perfil básico de um autor anônimo em uma tentativa de ajudar os investigadores a restringir seu campo suspeito do UNSUB (assunto desconhecido).

Havia também uma categoria em cada corpus de "Antecedentes do Autor confirmados", na qual, uma vez identificado e admitido como autor um autor anônimo, esses achados também eram incluídos nos respectivos corpora. Em minha própria experiência, raramente os perfis mudam, ou se sim, não dramaticamente.

5. Se Ted Kaczynski, o Unabomber, fosse um cibercriminoso (considerando que toda a sua ideologia não se baseasse numa pauta anti-tecnológica) e se o manifesto fosse publicado em uma rede social, na DeepWeb ou em qualquer meio tecnológico, você acredita que a Linguística Forense e o perfilamento criminal teriam sido suficientes para a solução do caso? Na sua opinião, o que poderia ter acontecido de forma diferente?

Sim, acredito que o caso poderia e ainda teria sido resolvido, mesmo que Kaczynski escolhesse a internet para enviar suas comunicações à mídia e às suas vítimas. Seu estilo de escrita e seus temas e tópicos eram idiossincráticos o suficiente para que, independentemente do fórum / gênero que ele tivesse escolhido para distribuir sua missiva, eu acredito que ele seria eventualmente identificado.

Claro, ele usou o sistema postal dos Estados Unidos e cartas padrão datilografadas para "enganar" separadamente duas de suas vítimas para que abrissem as bombas enviadas para elas (Wood em 1980, McConnell em 1985), então ele teria que mudar seu modus operandi nesses dois casos, mas ele teria sido inteligente o suficiente para fazê-lo com sucesso.

Além disso, se seu manifesto estivesse disponível hoje na internet, a busca pelo axioma / provérbio de verbo invertido, "você não pode comer seu bolo e ter também", teria sido muito mais fácil de encontrar e vincular diretamente a ele. Como foi em 1996, só porque sua mãe guardou uma cópia de uma carta ao editor do início dos anos 1970 escrita por um muito mais jovem Ted Kaczynski, eu a encontrei entre os outros 178 longos documentos escritos de seu filho, fornecidos por a família.

APÊNDICE B – INTERVIEW WITH JAMES R. FITZGERALD

1. After so many years working with the police and the FBI, would you say that you see Forensic Linguistics as the best or even as an inevitable source of evidence for solving crimes that are committed through language, even if there were other evidences?

Every investigation is different, so it's impossible to say that a forensic linguistic (FL) analysis of spoken or written language in a crime would be the "best," or even "inevitable" to its solution. It would depend on the facts of the case and the evidence which would be available to the investigators, and eventually the courts.

As a retired FBI agent, and police investigator before that, I would now almost always proffer that direct evidence such as fingerprints or DNA is the best evidence to have in a case in terms of linking someone to a questioned writing source. It's hard to dispute the quantitative findings as related to fingerprints and DNA. But absent that sort of direct evidence, FL analysis may prove to be the most reliable method of determining authorship and to subsequently present it as testimony to the triers of the facts in a criminal or civil matter – if so assessed, of course, by a qualified and experienced forensic linguist.

2. In your interview for a Newsy, available on YouTube, you said that you consider corpus analysis the future of Forensic Linguistics. Knowing that it is a growing science, in a computerized and technological century, in which more and more crimes are committed through virtual means, how do you believe that technology can contribute to the development of Forensic Linguistics?

Like existing local, national, and international databases of fingerprints and DNA samples, each attributtal to a specific person, it's only a matter of time before corpora of known language will not only exist (as they already do) but will also be utilized more and more by official government agencies in attempts to identify the authors of questioned communications. Instead of investigators and/or forensic linguists searching separate corpora for specific and demonstrable writing features or speech patterning for purposes of comparison, there may someday be one fully searchable corpus avialable to those so interested in collecting data through language. It will be a combined corpus of all, or most, existing corpora. This may or

may not be publicly sourced language, but also private emails, texts, and social media messaging. I would have a lesser concern if these type corpus searches were undertaken via government subpoenas or search warrants, but I would not want to see this system accessed (abused?) without the proper judicial authority and multiple checks and balances. Otherwise, it would be that proverbial slippery slope of citizens losing privacy to an overreaching Big Government. That could prove problematic to a free society in the short run and in the long run.

(Just to be clear re the above...a corpus of threatening, criminally oriented, and/or harassing communications such as those found in the FBI's CTAD, would not constitute a violation of the rights of the writers/speakers whose communications are located therein. The CTAD communications are mostly anonymous and sent from one party to another for purposes of criminal intent. They were, in effect, made public by the sender in the commission of a crime, so his/her privacy rights are not a factor here.)

3. The effectiveness of using Forensic Linguistics to attribute authorship in crimes committed non-virtually has shown itself to be increasingly blatant. In your opinion and in view of your vast experience in the police field and in the linguistic field, could Forensic Linguistics prove itself to be useful for the attribution of authorship and consequent solution of cybercrimes? In the USA, is this already happening? Considering the obstacles that you found yourself in the application of Forensic Linguistics, do you believe that, in this virtual field (of cybercrimes), the science would find more obstacles to its application than in the non-virtual field?

If by "cybercrimes" you mean language transmitted via the internet and/or cell phones, yes, analyses of these type of communications are already occurring in the US. In fact, it is very rare nowadays for me in my experience as a forensic linguist to have actual handwritten letters, or even typed letters, to assess. It's almost always emails, blog entries, social media messaging, or the like.

I used to teach that traditional transmittal envelopes in and of themselves contained their own form of linguistic evidence, certainly from a stylistics perspective, to include formatting, spacing, etc. But as threatening or harassing or criminally oriented communications are so rarely sent through the mail/post today, transmittal envelopes as evidence are almost a moot issue.

I suppose in cybercrimes as the sort listed above, one obstacle would involve the Auto-Correct function available in word processing programs. Obviously, if a person had a habit of

misspelling certain words, they would ostensibly be corrected before the communication was ever sent. But, as any forensic linguist knows, there are many alternate means of determining potential common authorship other than orthographically, or even grammatically, in which the Auto-Correct function would normally perform their assigned tasks.

4. As the creator of CTAD and CTARC and as a FBI profiler, do you believe that such tools are based on the mechanisms and techniques of *criminal profiling*? How have these tools been used? Have these tools (CTAD and CTARC) been used or are they expected to be used in the area of cybercrime?

While both the FBI's CTAD and the Academy Group, Inc's CTARC (now defunct) had fully searchable "profiling" categories built into them, they were separate and distinct to any traditional word, punctuation, or even parts-of-speech (POS) searches which would be undertaken. These traditional searches were always the primary function of the two different corpora which I developed a decade apart.

However, there was also a component in each which allowed for a linguistic profile of specific demographic features to be searched. This would be undertaken in an attempt to determine which still-anonymous communications may have been authored by sex, age, race/ethnicity, occupation, language nativeness, etc. These would not be evidentiary nor testimonial in the courts, but used instead during the early stages of an investigation to determine a basic profile of an anonymous author in an attempt to assist investigators in narrowing their suspect field of the UNSUB (unknown subject).

There was also a category in each corpus of "Author Background Confirmed" in which once an anonymous author was identified and admitted to being the author, those findings were also entered into the respective corpora. In my own experience, rarely would the profiles change, or if so, not dramatically.

5. If Ted Kaczynski, the Unabomber, were a cybercriminal (considering that his entire ideology was not based on an anti-technological agenda) and if the manifesto was published on a social network, on DeepWeb or in any technological medium, do you believe that Forensic Linguistics and *criminal profiling* would have been sufficient to resolve the case? In your opinion, what could have happened differently?

Yes, I believe the case could and would still have been solved, even if Kaczynski chose the internet to send his communications to the media and his victims. His writing style and his

themes and topics were idiosyncratic enough that regardless of the forum/genre he chose to distribute his missive, I believe he would have been eventually identified.

Of course, he used the US postal system and standard typewritten ruse letters to separately “trick” two of his victims into opening his mailed bombs to them (Wood in 1980, McConnell in 1985), so he would have had to change his modus operandi in those two instances, but he would have been clever enough to have successfully done so.

Also, if his manifesto was available today via the internet, the search for the verb-inverted axiom/proverb, “you can’t eat your cake and have it too,” would have been much easier to find and to link directly to him. As it was in 1996, only because his mother happened to keep a copy of an early 1970s letter-to-the editor written by a much-younger Ted Kaczynski, did I find it amidst the other 178 lengthy written documents of her son provided by the family.