

COMO A LINGUÍSTICA FORENSE E A LINGUÍSTICA DE CORPUS PODEM AUXILIAR EM UMA INVESTIGAÇÃO CRIMINAL: UMA ANÁLISE DO MANIFESTO DE UNABOMBER

Janaína da Silva Gomes¹

Jane Marian²

RESUMO

O objetivo da pesquisa é analisar e explicar o caso Unabomber e como os instrumentos da linguística forense e da linguística de corpus podem ser elementos complementares em uma investigação criminal. Para isso, foram selecionados, compilados e processados pelo *software Sketch Engine* os textos escritos por Ted Kaczynski: o manifesto (34.136 palavras), textos escritos pelo autor encontrados na internet (75.490 palavras) e cartas digitalizadas que foram escritas por ele antes de sua prisão (7.605 palavras). Como resultado, percebeu-se que a linguística forense teve inquestionável participação para prender um dos maiores terroristas dos Estados Unidos. No entanto, apesar da grande participação da linguística no desfecho do caso, ocorreram outros eventos sucessivos para que Ted Kaczynski fosse preso. Conclui-se que a análise comparativa do discurso dentro da área forense não se resume apenas aos termos frequentes, mas também, na forma que escritor expressa suas ideias e a fluidez do texto. Entende-se, portanto, que esse estudo da linguagem é essencial e serve como fator eliminatório para a solução de casos.

Palavras-chave: Linguística Forense. Linguística de Corpus. Análise do Discurso. Ted Kaczynski. Sketch Engine.

¹ Aluna do 6º período do curso de Letras Português/Inglês da FAE Centro Universitário. Bolsista do Programa de Apoio à iniciação Científica (PAIC 2019-2020). *E-mail*: j.sgomes@hotmail.com

² Orientadora da Pesquisa. Doutora em Linguística. Professora da FAE Centro Universitário. *E-mail*: jane.marian@fae.edu

INTRODUÇÃO

A linguística forense é o estudo de evidências de natureza textual escrita e falada. Essa ciência visa compreender a mensagem que está sendo passada pelo autor investigado para poder selecionar possíveis suspeitos e solucionar casos criminais. Com o uso de tecnologias e evidências empíricas (linguística de corpus) pode-se buscar dentro dos textos inúmeros rastros linguísticos.

Ao final de 1970, um terrorista começou a enviar cartas-bombas para algumas vítimas específicas, o criminoso recebeu o apelido de Unabomber. Suas atividades criminosas duraram cerca de 17 anos e o FBI estava com dificuldades em traçar um perfil, centenas de homens e mulheres eram suspeitos. Contudo, o terrorista escreveu um texto, o manifesto: *Industrial Society and Its Future*, cujas pistas linguísticas facilitaram encontrar o suspeito mais notório da história criminal, Ted Kaczynski. Ao revistarem a residência do suspeito, encontraram mais evidências incriminatórias, sendo assim, conseguiram comprovar que Kaczynski era o Unabomber. Ele foi condenado a quatro prisões perpétuas e esse foi o caso mais caro do FBI, mais de 50 milhões de dólares foram gastos na época com as investigações.

Sabe-se que a linguística auxiliou na prisão de Ted Kaczynski. Mas quais pontos dentro do caso Unabomber foram cruciais para prendê-lo? O estudo da linguística sempre foi de grande importância para o desenvolvimento do ser humano e suas comunicações. No entanto, foi com avanços nas tecnologias, com o surgimento dos computadores e da internet que houve maior demanda para estudos nesta área. Nota-se que esses eventos produziram um aumento significativo de informações linguísticas por meio de diversas fontes, como por exemplo: as ligações telefônicas, e-mails, *WhatsApp*, mensagens escritas nas redes sociais, etc. Mas, nesse sentido, as investigações criminais ainda precisam avançar bastante, e a tecnologia pode ser utilizada em prol disso.

A comunicação se faz presente em qualquer situação da vida, desde os momentos mais cotidianos e informais até os eventos mais ilustres e formais. Portanto, verifica-se que a solução dos problemas criminais pode estar conectada diretamente as análises linguísticas. O objetivo desta pesquisa é analisar o manifesto *Industrial Society and Its Future*, escrito por Ted Kaczynsky (o Unabomber), utilizando o programa computacional de análises linguísticas, *Sketch Engine*, com o intuito de verificar quais as marcas do discurso no texto. O uso da Linguística de Corpus se faz necessária para que os termos possam ser extraídos e analisados em seus contextos de forma comparativa.

1 REVISÃO BIBLIOGRÁFICA

1.1 LINGUÍSTICA FORENSE

O linguista forense é um especialista na área de Letras/Linguística. O seu trabalho como agente da polícia consiste em unir evidências de um crime, que estão em formato de áudio ou escrita. Esse profissional, também chamado de perito, irá analisar o conteúdo, ou seja, os rastros linguísticos, em busca de pistas que levem a investigação para o rumo de sua resolução. Como disse Silva (2017, p. 11): “A linguística é a ciência que investiga os fenômenos relacionados à linguagem e que busca determinar os princípios e as características que regulam as estruturas da língua”.

Os indivíduos dentro de uma mesma comunidade podem se diferenciar uns dos outros devido ao seu vocabulário particular, também conhecido como rastro ou marcas do linguístico, identidade linguística ou idioleto. Os elementos que podem auxiliar a formação de um vocabulário particular são: conhecimento de mundo, relações sociais, experiências de vida, cultura, locais de moradia, religião, entre outros. Toda essa bagagem é armazenada na mente de um indivíduo e forma o vocabulário particular. Sua forma de discursar sobre os elementos da vida também é um elemento que se expõe de forma automática. A fala é comparada com impressões digitais, nenhum ser humano fala igual ao outro, pois a variedade fonológica ocorre de forma diferente de uma pessoa para outra (TIMBANA, 2016, p. 38).

Em um ambiente onde está um cadáver, cujas evidências visuais indicam um suposto suicídio, é encontrada uma carta que afirma tal acontecimento (suicídio, como carta de despedida). Por se tratar de uma evidência escrita, a carta será analisada por um linguista forense que irá tentar descobrir se o autor daquela carta foi realmente o falecido, caso a conclusão seja que a carta foi forjada por uma outra pessoa, a investigação passará a buscar outros suspeitos do crime, pois, sendo assim, trata-se de um possível homicídio e não suicídio. Contudo, a Linguística Forense sempre conta com outras áreas das ciências forenses: medicina, na investigação do cadáver; *digital forensic*, buscando evidências em aparelhos eletrônicos, entre outras.

Com o passar do tempo, a tecnologia auxiliou essa área também. Atualmente existem programas de computadores capazes de fazer análise linguística em questão de segundos e encontrar elementos nos textos que “o olho humano” levaria muito mais tempo para encontrar. *Sketch Engine* é um *software* que auxilia o pesquisador para a análise e estudo da linguagem. O programa pode ser um facilitador para que o investigador possa analisar bilhões de palavras e identificar instantaneamente o que é

típico no texto ou raro. Esse programa é bastante utilizado por linguistas, lexicógrafos, tradutores, alunos e professores. É, normalmente, a primeira escolha para soluções de linguagens no mundo inteiro.

De acordo com Champod e Meuwly (1999, p. 2), a coleta de evidências serve para eliminar suspeitos até chegar ao culpado, para o autor, a possibilidade de encontrar pessoas iguais são nulas. O perito, ao coletar e analisar evidências, não irá apontar um suspeito de forma objetiva, mas apresentar dados de possibilidades, elementos que irão classificar a qualidade da evidência.

O manifesto de Unabomber foi uma grande, senão a maior, evidência do uso da linguística para o período. Contudo, apesar das grandes semelhanças do manifesto com as cartas de Kaczynski, ainda assim, muitos outros indivíduos não estariam isentos de ter uma escrita semelhante. Contudo, encontrar na cabine de Kaczynski mais evidências incriminatórias foi crucial para levar o caso adiante.

1.2 LINGUÍSTICA DE CORPUS

A Linguística de Corpus é uma área da linguística aplicada que se utiliza de textos e recursos computacionais para análises. Os textos são chamados de corpus ou corpora, o corpus é um texto único (singular) sobre um assunto específico, por exemplo, o manifesto de Unabomber pode ser considerado um corpus. Os corpora são compilados de textos (plural), ou seja, mais de um texto. Segundo Sardinha (2004, p. 3) a linguística de corpus:

[...] ocupa-se da coleta e da exploração de corpora, ou conjuntos de dados linguísticos textuais coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua ou variedade linguística. Como tal, dedica-se à exploração da linguagem por meio de evidências empíricas, extraídas pelo computador.

De acordo com o autor a linguística de corpus é uma área da linguística que faz coleta de dados da linguagem natural (falada por humanos e não por máquinas) para que, futuramente, seja utilizada para diversos fins, como: análise comparativa, criação de dicionários, tradução, ensino de língua estrangeira ou nativa, entre outros.

A Linguística de Corpus sempre foi presente entre os estudos da linguagem, porém com o avanço da tecnologia é que obteve progressos significativos. O primeiro corpus criado com ajuda de um computador foi o corpus Brown em 1964, desenvolvido pela *Brown University* que contava com um milhão de palavras. Apesar das dificuldades, os textos foram transferidos para o computador com o uso de cartões perfurados um por um (SARDINHA, 2004, p. 1). Para o autor esta área é considerada uma metodologia, pois

pode relacionar-se com várias áreas de conhecimento. Na presente pesquisa unimos o conhecimento da Linguística de Corpus e o da Linguística Forense como uma nova perspectiva para contribuir com o contexto jurídico/criminal.

No caso Unabomber, o perito criminal, tinha o texto do manifesto *Industrial Society and Its Future* e cartas que foram escritas por Theodore John Kaczynski, principal suspeito. O linguista fez a comparação analisando todos os textos em busca de semelhanças entre si, seja por uso frequente de palavras, ideias, estruturas frasais ou qualquer rastro linguístico que possa identificar o culpado.

1.3 ANÁLISE DO DISCURSO

Conhecemos o termo “discurso” como uma fala realizada por alguém sobre algum assunto específico, por exemplo, candidatos políticos, líderes religiosos, vencedores de um prêmio, etc. Porém, discurso é, na verdade, uma comunicação que envolve o falante, o ouvinte, a mensagem e um contexto. É um conjunto de ideias e opiniões expostas por meio da produção oral ou escrita.

O discurso produz sentido nos ouvintes e, estes, são analisados e interpretados. A análise do discurso envolve a interpretação de elementos envolvidos nesse processo. Essa é uma ciência da linguagem responsável em encontrar o sentido do contexto e compreender diferentes ideologias, sentidos e o que a mensagem quer passar, quesitos que sempre se encontram no discurso de forma intrínseca. Uma forma que existe de análise do discurso é captar algumas palavras-chaves presentes no texto que passam informações além do que o falante está expondo de fato.

Caregnato e Mutti (2006, p. 2) alegam que não existe apenas uma forma de análise, mas por volta de 57 maneiras diferentes. Análise do discurso busca o sentido e foca na interpretação, a qual pode ter suas variações em um contexto. Um exemplo dado pelas autoras seria a frase “é dando que se recebe”, que pode ter sentidos diferentes a depender de seu falante, seja um religioso, político ou uma prostituta.

Um dos fundadores dos estudos sobre o discurso foi Michel Pêcheux, estabelecendo a relação existente no discurso entre língua/sujeito/história ou língua/ideologia; portanto, quem segue este princípio pode afirmar uma filiação com a AD da linha francesa (CAREGNATO; MUTTI, 2006, p. 2).

Paulon, Nascimento e Laruccia (2014, p. 3) apresentam a análise do discurso como um estudo do texto que segue além dos materiais linguísticos, considerando que está relacionada com o contexto histórico-social do enunciador e que se trata de uma ciência que questiona a interpretação. Pêcheux (1999 *apud* PAULON; NASCIMENTO; LAUCCIA,

2014, p. 4) apresentou essa ciência sendo formada por três elementos: o histórico, relacionado com questões sociais; a linguística, que explica a estrutura do texto e; a ideologia que explica a subjetividade e como o falante se relaciona com o simbolismo. Outro fator citado foi a seguinte frase: “A Análise do Discurso visa a compreensão de como um objeto simbólico produz sentidos, como ele está investido de significância para e por sujeitos” (ORLANDI, 2007, *apud* PAULON; NASCIMENTO; LAUCCIA, 2014, p. 6).

Souza (2014, p. 11) reitera a ideia dos outros autores citados que a Análise do Discurso aborda o sentido na linguagem e apresenta a imagem de Pêcheux em seu trabalho, que a ideia francesa da disciplina seja visar o sentido na linguagem. Contextos onde pessoas se comunicam, independentemente do modo, linguagem verbal (oral/escrito) ou não-verbal existe um discurso que possui sentido entre os indivíduos.

Em outras palavras, podemos alegar que a análise do discurso é uma ciência que não estuda somente o que está dito diretamente no texto, mas sim, o que seu falante/escritor quer abordar com aquelas palavras. É um estudo voltado para o significado, o sentido e a interpretação do que está intrínseco no discurso e quais os valores e ideais são apresentados pelo orador, por isso, que essa análise consiste em outros elementos como o falante, o contexto e o alvo do discurso. Quem está falando, o que está falando, para quem está falando e o que quer passar com seu discurso.

1.4 O CASO UNABOMBER

Theodore John Kaczynski nasceu em Chicago no ano de 1942 e, quando já aluno, destacava-se nos estudos, logo conseguiu pular um ano na escola após ser constatado que seu QI era 167, ou seja, alto o bastante para ser classificado como um gênio. Entretanto, por ser tímido e destacar-se pela inteligência, acabou ficando deslocado entre os colegas mais velhos, sendo alvo de provocações. Mesmo com sua personalidade retraída, ele era mais aberto quando falava sobre matemática e outros assuntos de seu interesse, era considerado simpático, apesar de tudo.

Logo após concluir o ensino médio, aos 16 anos, ingressou em Harvard, no curso de Matemática. Neste período, Henry Murray³, o diretor da área de Psicologia de Harvard, iniciou um processo seletivo onde homens e mulheres participaram de forma voluntária. Foram 22 estudantes que passaram no teste de seleção, Kaczynski era um deles. Estes alunos passaram por uma experiência que consistia em exposições e agressões psicológicas, cuja finalidade era abalá-los emocionalmente. O experimento durou cerca de três anos e Kaczynski participou de 200 sessões.

³ Disponível em: <<https://psychology.fas.harvard.edu/people/henry-murray>> Acesso em: 25 jun. 2020.

Ted entrou no doutorado, na área de Matemática, e começou a lecionar em uma universidade. Largou o trabalho como professor universitário e foi viver em uma pequena cabana no meio de uma floresta em Montana, praticando jardinagem para cultivar o próprio alimento e recebia uma pequena ajuda financeira da família. Um empreendimento imobiliário estava desmatando a floresta onde residia e, assim, decidiu como forma de protesto, enviar as cartas-bombas.

Kaczynski publicou um manifesto chamado *Industrial Society and Its Future*, o qual abordava como tema principal a tecnologia como prejudicial ao ecossistema. Ele queria atingir o maior número de pessoas com seu discurso, então propôs ao FBI que pararia com o terrorismo em troca da publicação de seu texto. O mesmo foi publicado no *Washington Post*⁴ e no *New York Times*⁵.

O nome assinado no manifesto não foi seu nome verdadeiro e nem Unabomber, mas a sigla FC que significa *Freedom Club*, que se trata de um grupo anarquista contra cientistas e a tecnologia. O FBI esperava encontrar o criminoso nas bancas de jornais, mas isso não ocorreu. No entanto, o manifesto foi lido por muitas pessoas e, entre elas, o irmão e a cunhada do criminoso que entrou em contato com o FBI, alegando que elementos no manifesto eram parecidas com a forma de escrita de seu irmão e entregou um compilado de cartas escrito ele.

Durante a análise foi constatado que a forma de escrita de Unabomber era muito sofisticada, em controvérsia do ponto do perfil criminal que apontava que o criminoso tivesse baixa escolaridade. A coleta de dados do corpus e dos corpora, constatou-se que Unabomber usava termos como “*surrogate*” e “*tautology*”. Incluindo que o falante se assemelhava aos residentes da Carolina do Norte e que seria religioso, pois utilizou termos como desejos de Deus e pecados (CHRISTENSEN, FORD; LEONARD, 2017, p. 885). Também foi constatado que houve uma alteração ortográfica, entretanto, a mudança não alterava a pronúncia dos termos: ele trocou os termos *analyse*, *licence*, *willfully* e *installment* para *analyze*, *licence*, *wilfully* e *instalment* (TURCHIE, 2003). Outro detalhe foi a inversão de palavras da frase “*you can’t have your cake and eat it too!*” que tanto Kaczynski quanto Unabomber inverteram para “*you can’t eat your cake and have it too.*”

Após análise de dados linguísticos, o FBI decidiu ir até a residência de Kaczynski e prendê-lo. Em sua cabine encontraram evidências incriminatórias como a máquina que ele utilizou para escrever o manifesto e os componentes que usava para montar as

⁴ Disponível em: <https://www.washingtonpost.com/wp-srv/national/longterm/unabomber/manifesto.text.htm> Acesso em: 25 jun. 2020.

⁵ Disponível em: <https://archive.nytimes.com/www.nytimes.com/library/national/unabom-manifesto-1.html?mcubz=3>. Acesso em: 25 jun. 2020.

bombas. Ted Kazynski foi condenado a oito prisões perpétuas e foi um dos casos mais caros do FBI, que custou mais que 50 milhões de dólares (BHATIA, 2017, p. 4).

Unabomber era apto em disfarçar sua identidade e manipular a forma que era visto pelo FBI. Foram sequências de eventos que o fizeram ser capturado. A primeira foi a publicação do manifesto que possibilitou a desconfiança de sua identidade pela maneira de escrever e das ideias explícitas; depois, a família que decidiu denunciar o homem, ao invés de protegê-lo; após as evidências linguísticas, foi permitida a realização de uma busca na residência de Ted Kazynski, onde encontraram a máquina de escrever e os componentes das bombas. Caso ele não tivesse escrito o manifesto, a família decidisse não o denunciar ou não encontrassem mais evidências em sua casa, provavelmente, Unabomber ainda estaria solto.

2 METODOLOGIA

A metodologia da pesquisa apresenta uma abordagem empírica de cunho tanto qualitativo quanto quantitativo. A fonte para a coleta de dados consiste no manifesto escrito por Theodore Kazynski *Industrial Society and Its Future* (34.136 palavras); textos disponíveis *online* escritos pelo mesmo autor (75.490 palavras) e cartas que compõem um compilado intitulado *The Communiques of Freedom Club* (7.605 palavras). O programa computacional utilizado para as análises linguísticas foi o *Sketch Engine*.

FIGURA 1 – Criação do Corpus

1. CREATE CORPUS > 2. ADD TEXTS > 3. COMPILE

Build your own private corpus from texts on the web or from your own documents.

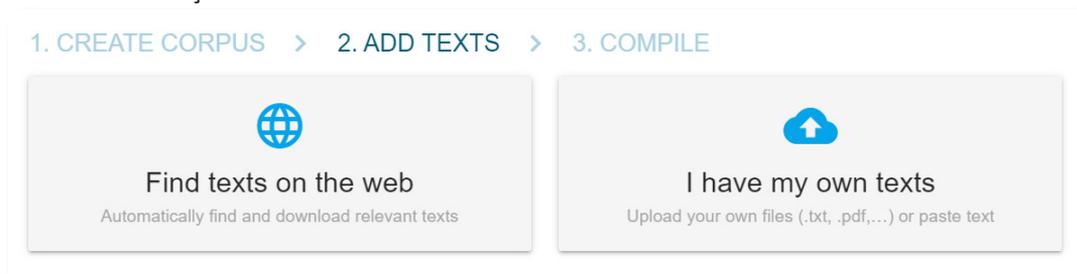
Name	Manifesto
Corpus type	<input checked="" type="radio"/> Single language corpus <input type="radio"/> Multilingual corpus
Language	English
Description	Unabomber's manifesto.

Storage used: 34,136 of 1,000,000 words (3%)

FONTE: As Autoras (2020)

Na FIG. 1, depois encontrado o corpus online e salvo no computador no formato PDF, o corpus é criado para ser processado pelo programa do *Sketch Engine*. Para isso, ele foi nomeado como Manifesto. A segunda opção, *Corpus type*, deixamos como *Single language corpus*, pois se tratava de um corpus em uma única língua. Em *Language* (idioma), selecionou-se o idioma inglês, língua que estava escrito o corpus. Por último, em *Description* (descrição) insere-se (opcional) uma descrição para o corpus criado.

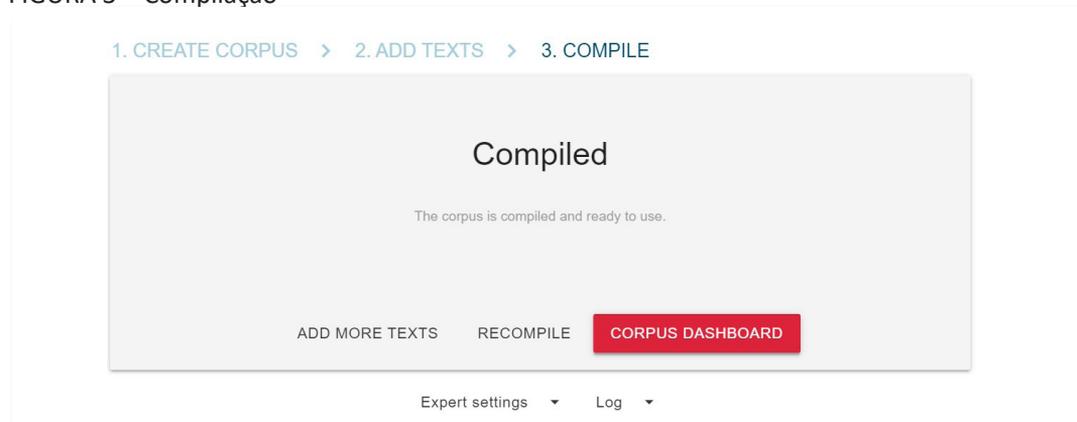
FIGURA 2 – Seleção de Textos



FONTE: As Autoras (2020)

O segundo passo do programa é adicionar o arquivo ou arquivos que irão compor o corpus. Existe a possibilidade de buscar o arquivo pela internet ou buscar dentro do computador que está sendo utilizado para acessar o programa.

FIGURA 3 – Compilação

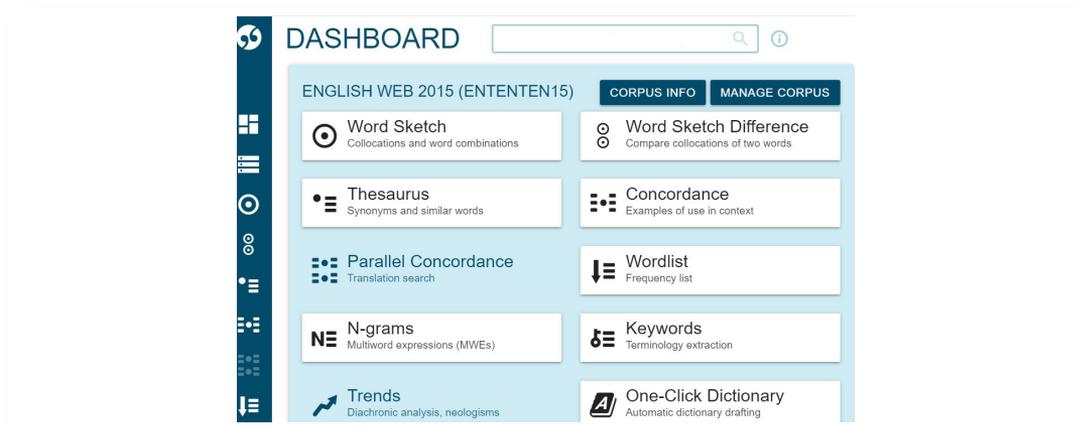


FONTE: As Autoras (2020)

Por último, o usuário deve compilar o arquivo, sendo assim, após a espera da compilação, o corpus estará pronto para ser utilizado. Tais passos foram seguidos três vezes por existirem três corpus.

O primeiro corpus, o Manifesto, *Industrial Society and Its Future*. O segundo corpus, textos online, 20 textos encontrados na internet, escritos pelo Unabomber após sua prisão. E o terceiro corpus, *The Communiques of Freedom Club*, uma coleta de 14 cartas escritas pelo Unabomber sob o pseudônimo de FC, endereçados para diversos destinatários.

FIGURA 4 – Dashboard



FONTE: As Autoras (2020)

Na tela de início, em *Dashboard*, existem as ferramentas de análise: *Keywords* (palavras-chaves); *Thesaurus* (sinônimos e palavras semelhantes); *Concordance* (exemplos de uso dentro de um contexto); *Wordlist* (lista de palavras por ordem de frequência), entre outras funções.

3 RESULTADOS E DISCUSSÕES

Após as coletas de dados do FBI, foi feita uma lista com os termos que chamaram a atenção juntamente com os termos que foram considerados mais adequados. Por exemplo, que Unabomber trocou “*Willfully*” por “*Wilfully*”, logo ambas formas foram colocadas na lista, tal seria: *analyse, analyze, licence, license, wilfully, willfully, instalment, installment, surrogate, God, tautology, cake*. O motivo de colocar as duas formas é para o caso de o termo aparecer de outra maneira, mas ainda fazer presença nos corpus. Vale ressaltar que as palavras escritas de forma “correta” serão consideradas irrelevantes a depender do contexto. A primeira tabela apresenta o corpus “Manifesto”; a segunda, o corpus “Textos Online” e a terceira, “The Communiques of Freedom Club”.

FIGURA 5 – Tabelas dos corpus

Corpus 1 - Manifesto			Corpus 2 - Textos			Corpus 3 - Freedom Club		
Word	↓ Frequency ?		Word	↓ Frequency ?		Word	↓ Frequency ?	
1 surrogate	40	...	1 god	2	...	1 instalment	3	...
2 license	3	...	2 tautology	1	...			
3 god	3	...	3 surrogate	1	...			
4 willfully	1	...	4 analyze	1	...			
5 tautology	1	...						
6 cake	1	...						

FONTE: As Autoras (2020)

Como visto na tabela, o termo “*surrogate*” é o mais frequente de todos da lista de termos, aparecendo 40 vezes e é um termo acompanhado por “*activity*” ou “*activities*” em todas as vezes. Kaczynski aborda o termo a busca de um propósito artificial, o qual move as pessoas para que trabalhem e cumpram as metas até chegar ao seu objetivo. Ele comenta sobre o gasto de energia mental e física em busca de uma meta específica, sendo a tal uma “*surrogate activity*” caso a pessoa não se sinta privada de seus objetivos ao usar sua energia para outras atividades. Apesar do termo aparecer 40 vezes no manifesto, ela é repetida diversas vezes em um curto espaço, dando a sensação de repetição e até mesmo “voltas” ao redor do mesmo assunto, de forma insistente.

“*License*”, aparece três vezes no manifesto e pede por um pouco mais de atenção do leitor. Turchie (2003) apontou que Unabomber teria invertido os termos de “*License*” para “*Licence*”, entretanto, no corpus disponível o termo estava “*license*” como pode-se ver na FIG. 5. O termo está inserido dentro do contexto de liberdade do ser humano, abordando que a presença de carros não tira a liberdade de uma pessoa escolher ou não de ter um veículo, entretanto, restringiu a liberdade de locomoção das pessoas. Foi necessário regulamento do uso de automóveis, incluindo o ritmo do veículo, pois existem leis de trânsito. Inclusive, existem obrigações como prova prática de direção, carteira de motorista, entre outros elementos. A respeito do homem moderno, no sentido de sistema industrial e grandes organizações, que acreditam ter poder em relação à natureza. Logo, o homem moderno necessita de licenças para atuar e exercer sua função dentro de supervisão e controle.

O terceiro termo, “*God*”, de acordo com o programa, aparece três vezes, entretanto, o programa não captou uma quarta citação, pois o termo foi escrito como “*God’s*”, o termo se apresenta quatro vezes. No parágrafo 122 ele aborda a cura de doenças que não podem ser erradicadas, por exemplo, a diabetes. Em caso de uma cura para a diabetes, a seleção natural irá parar, o genes irá se espalhar e segue que já é uma hipótese, pois existe o controle pelo uso da insulina. Conclui o parágrafo afirmando que a eugenia, ciência que visa a reprodução da espécie com os melhores genes, é uma alternativa para eliminação de doenças, porém, a depender da religião do leitor, o humano não será mais uma criação de Deus, mas um objeto de fábrica. Depois aborda como é irresistível a presença da tecnologia que facilita o dia a dia das pessoas, por exemplo, o telefone. Depois, ele faz comparação a não resistir à engenharia genética que pode eliminar doenças hereditárias, insistindo mais uma vez que o humano se tornará um produto de fábrica e não uma criação divina. Por último, que a presença da tecnologia ainda fará o humano ser útil como mão de obra e as máquinas fazendo trabalhos menores; porém, que haverá um cenário que o trabalho humano será menos relevante e que os humanos despreparados para essa realidade podem suprir suas necessidades com

uso de drogas, cultos ou grupos de ódio. Depois, fala da possibilidade de sobrevivência do sistema tecnológico-industrial e que os “burgueses” integrados serão dependentes dessas organizações, citando Deus como método ou motivo dessa situação. Na última citação ele aborda sobre a natureza selvagem incluindo os humanos, sendo resultado do acaso, do livre arbítrio ou de Deus. Nas quatro vezes, Kaczynski aparenta ser um ateu que respeita a religião alheia, não exatamente um homem religioso.

“*Willfully*”, de acordo com Turchie (2003), foi alterado para “*Wilfully*”, entretanto, apesar do termo aparecer no corpus, não está escrito de acordo com o apontamento. “*But we are suggesting neither that the human race would voluntarily turn power over to the machines nor that the machines would willfully seize power.*” (KACZYNSKI, 1995, p. 22)

No capítulo “*Some Principles of History*”, o autor aborda cinco princípios de cunho político. O primeiro aborda, em sua opinião, que a construção da histórica é somada de um componente errático e um componente que comporta tendências históricas ao longo prazo; usando como exemplo um movimento para acabar com a corrupção política e, após o êxito, esse movimento irá relaxar e a corrupção irá retornar. Sendo assim, a corrupção política de uma nação tem tendência em se manter presente em uma nação, mudando somente com uma evolução dessa sociedade. Em seguida, o autor fala que esse princípio é praticamente uma “*tautology*”, pois é necessária a relação da uma tendência maior com as pequenas mudanças, ou irá apenas vagar sem propósitos, deixando de ser uma grande tendência.

101. The first principle is almost a tautology. If a trend were not stable with respect to small changes, it would wander at random rather than following a definite direction; in other words it would not be a long-term trend at all (KACZYNSKI, 1995, p. 12).

O termo “*Cake*” aparece na frase invertida que tanto Unabomber quanto Kaczynski utilizaram. O termo está presente no parágrafo 185, o qual diz o seguinte: “185. *As for the negative consequences of eliminating industrial society — well, you can’t eat your cake and have it too. To gain one thing you have to sacrifice another*” (KACZYNSKI, 1995, p. 24). O sentido do ditado é a impossibilidade de fazer duas coisas ao mesmo tempo, por exemplo “Assobiar e chupar cana”. Entretanto, o ditado correto é “*You can’t have your cake and eat it too*”, sendo o verbo “*have*” anterior ao verbo “*eat*”, o qual Kaczynski inverteu.

Já em “Textos Online escritos por Ted Kaczynski”, o termo “*God*” foi o mais frequente dentro do corpus e ambas as vezes aparecem no mesmo texto, no qual o autor critica o anarcoprimitivismo e fala sobre o cristianismo. Ele faz uma comparação na qual a revolução dos anarcoprimitivistas assemelha-se ao Dia do Julgamento, dia de destruição na qual Babilônia caiu e que o retorno de uma utopia primitiva é como

o retorno do Reino de Deus. Comenta sobre ativistas que acorrentam nas estradas para impedir a passagem de caminhões que transportam árvores. Assim, ele faz uma comparação desses ativistas com verdadeiros mártires cristãos. Cita: *“were beheaded for the witness of Jesus, and for the word of God”* (Revelation 20:4, apud KACZYNSKI, 2008, p. 37).

“Tautology” é encontrado em um texto no qual o autor afirma que o sistema tecnológico irá se auto destruir. Fala sobre o sistema que promove sua própria sobrevivência e propagação, abordando sobre um sistema social que apresenta como hierarquia em animais, assim como seres humanos; abordando que a seleção natural está presente em sistemas que se auto propagam; comparando que sistemas auto propagantes com organismos, que possuem melhor características para o meio, logo sobrevivem à seleção natural. Por fim, alega: *This of course is an obvious tautology, so it tells us nothing new. But it can serve to call our attention to factors that we might otherwise overlook* (KACZYNSKI, 2011, p. 3).

“Surrogate” e *“analyze”* aparecem no mesmo texto, no qual o nome do autor está como *“Último Reducto”*. O primeiro termo está presente em uma nota de rodapé, que foi um dos termos frequentes no manifesto. Entretanto, no decorrer do texto, existe o termo *“analyze”*, que foi listada como um dos termos dentro do manifesto que foram escritas de forma diferente; neste caso, o autor fala a respeito de analisar a luta anti-tecnologia.

“Instalment” foi um dos termos que chamou atenção durante a investigação e nesse corpus aparece três vezes em uma carta endereçada à Warren Hoge, um jornalista que trabalhou no New York Times de 1978 até 2008. Na carta, o termo aparece três vezes na mesma frase: *If material is serialized, first instalment becomes public property six months after appearance of first instalment, second instalment, etc.* (KAZYNSKI, 1995, p. 10).

Inicialmente esperava-se encontrar no manifesto todos os termos que a investigação encontrou, entretanto, as três versões encontradas na internet não dispunham tais dados, o que levou à conclusão que não existem garantias que um texto na internet não possa ter sido alterado no decorrer de 25 anos. Por outro lado, a frase que tanto Unabomber quanto Kaczynski inverteu foi satisfatório encontrar, pois foi a maior pista linguística.

Levando a análise do corpus *“Manifesto”*, a FIG. 6 apresenta uma listagem diferente, que é composta pelas palavras que co-ocorrem com mais frequência, entretanto levar em consideração que a pesquisa feita dentro do software é *“keywords”* ao invés de *“wordlist”*.

FIGURA 6 – Compilado de corpus “Keywords”

Word	Word	Word
1 power process	1 techno-industrial system	1 identifying number
2 industrial system	2 global self-prop	2 industrial system
3 surrogate activity	3 technoindustrial system	3 industrial society
4 industrial society	4 self-prop system	4 technological progress
5 technological system	5 modern society	5 basic sequence
6 modern man	6 industrial society	6 random sequence
7 human behavior	7 technological system	7 present form
8 genetic engineering	8 modern man	8 confidential note
9 modern leftism	9 revolutionary movement	9 industrial-technological system
10 primitive man	10 industrial system	10 book form

FONTE: As Autoras (2020)

A FIG. 6 apresenta a tabela dos três corpora apresentando seus 10 termos que mais co-ocorrem, a ferramenta utilizada foi “*keywords*”, que extrai os dados automaticamente. Utilizamos o corpus “Manifesto” para comparar quais 10 termos frequentes também aparecem na mesma listagem dos outros dois corpus.

Em comparação das *keywords* do corpus “Manifesto” com o corpus “Textos Online”, existe grande semelhança, pois coincidem os termos: “*industrial society*”, “*technological system*”, “*modern man*” e “*industrial system*”. Apesar da falta dos demais termos na tabela, o restante dos termos aparecem na pesquisa dentro do grupo de 355 termos frequentes do corpus “Textos Online de Ted Kaczynski”. Exceto *power process* e *modern leftism* que não apareceram na listagem dos 1.000 termos que mais co-ocorrem no “Manifesto”.

O terceiro corpus apresenta 308 termos que co-ocorrem e têm uma frequência menor de termos quando comparado com o corpus “Manifesto”, logicamente isso acontece devido ao tamanho reduzido do corpus. Os termos semelhantes são apenas duas: *industrial system* e *industrial society*. Os termos *power process*, *surrogate activity*, *technological system*, *modern man*, *modern leftism* e *primitive man* não aparecem na tabela de 308 termos que co-ocorrem.

Levando em conta esse levantamento de dados e comparação de termos que se repetem entre os três corpora, afirma-se que são constantes as repetições, assim percebe-se que as chances de se tratar do mesmo autor são grandes, pois são termos muito específicos. Na maioria dos casos os termos em destaque abordam sobre a tecnologia e o comportamento do humano em sociedade, temas que estão presentes em *Industrial Society and Its Future*. Outro detalhe pertinente nos três corpus é

uma sensação de que o autor está “dando voltas” em um assunto, algo que poderia exemplificar em quatro linhas ou menos, mas usa uma quantidade maior para explicar suas ideias, o que acaba deixando a leitura mais cansativa. Adendo que existe grande crítica sobre a presença da tecnologia não apenas destruindo o ecossistema, mas também tirando a liberdade das pessoas de alguma maneira. Os corpora também foram escritos com uma linguagem informal, sem se importar com excesso de repetição de termos e o autor passa segurança no que está dizendo. Essas insistências nos três corpora fluíram em uma análise linguística que apresenta grandes possibilidades de que foram escritas pelo mesmo autor.

Vale a pena abordar que a discrepância de informações foi incômoda durante a pesquisa, pois encontrar uma lista de termos que a investigação utilizou, mas não encontrar alguns desses termos no corpus do manifesto, fez questionar a possibilidade de uma edição no corpus online. Portanto, não se garante que nenhuma alteração tenha ocorrido no texto após 25 anos da prisão de Uabomber. Entretanto, encontrar a frase “*You can’t have your cake and eat it too*” foi satisfatório, e possivelmente a maior pista linguística.

Uma sugestão de pesquisa, entretanto, na área de psicologia, seria a análise do perfil de Kaczynski. Durante a pesquisa encontramos artigos e sites que afirmavam a personalidade retraída, porém simpática do criminoso; mas questionam a possibilidade de ter tido seu instinto de maldade aguçado após os traumas que passou na época de universitário, cujo maior problema é solucionar a questão se, caso ele não tivesse passado por tais experiências, o Unabomber existiria.

Também compreendemos que, para identificar um autor, não basta apenas extrair as palavras de maior uso, mas sim conhecer sua forma de expor as palavras, de apresentar ideias e como se sente em relação do que está abordando. Logo, para chegar a alguma conclusão, deve-se conhecer a escrita de quem está sendo investigado, além da visão de mundo e suas experiências.

Uma ciência forense, de qualquer linha, isolada é incompleta, é necessária uma “esfera” para a solução de um crime. Exemplificando, um linguista forense irá trabalhar com textos e áudios, porém, se houver busca de evidências dentro de aparelhos eletrônicos como computadores e celulares, será necessária ajuda de algum perito em computação forense. Outro adendo seria que as ciências forenses são eliminatórias e não classificatórias, que trabalham para ir reduzindo um grupo de suspeitos até a um único indivíduo. Um exemplo a respeito seria a quantidade de casos que foram arquivados por falta de evidência ou falta de atividade do criminoso.

A área da linguística visivelmente passa por um processo de pouco caso. A comunicação, seja ela verbal ou não, está presente no contexto humano em qualquer aspecto e foi graças a ela que a humanidade evoluiu e compartilhou conhecimento. Especificamente dentro de um contexto da criminalística, é comprovado que a Linguística Forense foi extremamente necessária na captura de Unabomber.

CONSIDERAÇÕES FINAIS

Apesar de algumas diferenças entre as fontes e os corpus, considerando que alguns termos não foram encontrados, conseguimos concluir que os textos foram escritos por Kaczynski. Analisando de outra perspectiva, como semelhança de termos que co-ocorrem, a forma de escrita e a abordagem dos assuntos que envolvem tecnologia, liberdade e homem moderno. Sendo assim, apesar desses elementos de termos ausentes, conseguimos seguir com a pesquisa à nossa maneira, obtendo resultados sem precisar repetir a investigação do FBI do começo ao fim. Isso foi graças à Linguística de Corpus que possibilitou a extração de outros termos e fazer sua comparação.

É inegável que a linguística forense não irá resolver algum caso sem auxílio de outras especialidades, além da necessidade que o criminoso cometa deslizes para que novas pistas surjam e o caso não seja arquivado. Unabomber cometeu deslizes de elementos linguísticos, por exemplo, escrever o manifesto e inverter os verbos em um ditado; após seu irmão, David Kaczynski ler o manifesto e reconhecer a escrita do irmão, decidiu fazer a denúncia ao invés de protegê-lo e um último elemento foi encontrar a máquina de escrever que Unabomber utilizou para criar o manifesto e os componentes que usava para criar suas bombas. Tais eventos sucessivos foram fundamentais para que a investigação avançasse e chegasse à sua conclusão. Provavelmente, se o terrorista não tivesse mantido os componentes das bombas em sua residência ou se não tivesse escrito o manifesto ou se sua família optasse por protegê-lo, Unabomber não teria sido preso.

A Linguística Forense necessita de mais avanço, ainda é uma área recente e pouco estudada, é uma área que levanta possibilidades do que um resultado objetivo. É fundamental que o linguista forense tenha familiaridade com a forma de escrita do suspeito, identificando sua maneira de expressar-se sobre os assuntos e a forma de posicionar as ideias, indo além das palavras frequentes e que co-ocorrem dentro dos textos comparados. Afinal, dois autores podem escrever textos sobre o mesmo assunto, de forma bem similar, porém o que vai diferenciar um do outro é a maneira que as ideias foram expressas.

Durante a leitura dos textos teóricos, concluímos que a forma de uma pessoa se comunicar corresponde como a uma “impressão digital linguística”, cuja formação adere vários elementos sociais que correspondem onde o indivíduo morou, suas experiências e conhecimento de mundo. A maioria das fontes teóricas repetiam a mesma informação, porém utilizando termos diferentes, o que enaltece a afirmação que a área da Linguística Forense e da Linguística de Corpus necessita de maior estudo e valorização.

Dentro de universidades, considera-se a sugestão de adicionar a área da Linguística Forense, pois até mesmo os próprios letrólogos em formação desconhecem a sua vertente em criminalística e havendo mais proveito dessa área em ambiente acadêmico, existem altas probabilidades que novos peritos em linguística irão surgir e descobrir novos avanços. É necessário que mais estudantes realizem pesquisas e que grupos de pesquisas sejam mais valorizados no meio acadêmico, pois a linguística foi capaz de ajudar a capturar um dos piores terroristas dos Estados Unidos.

REFERÊNCIAS

BHATIA, Tej K. **Forensic Linguist Tej Bhatia on the Hunt for the Unabomber**. 2017. Disponível em: <<https://surface.syr.edu/III/21>>. Acesso em: 10 maio 2020.

CAREGNATO, Rita Catalina Aquino; MUTTIR, Regina. Pesquisa **qualitativa**: análise de discurso versus análise de conteúdo. **Texto & Contexto: Enfermagem**, Porto Alegre, v. 15, n. 4, p. 679-648, 2006. Disponível em: <<https://www.scielo.br/pdf/tce/v15n4/v15n4a17.pdf>>. Acesso em: 6 abr. 2020.

CHAMPOD, Christophe; MEUWLY, Didier. The inference of identity in forensic speaker recognition. **Elsevier**, Lausanne, v. 31, n. 2-3, p. 193-203, Sep. 1999. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0167639399000783>>. Acesso em: 15 jul. 2020.

CHRISTENSEN, Tanya Karoli; FORD, Juliane E. R.; LEONARD, Robert A. Forensic Linguistics: applying the science of linguistics to issues of the law. **Hofstra Law Review**, Hempstead, v. 45, n. 3, p. 880-897, Mar. 2017. Disponível em: <<https://scholarlycommons.law.hofstra.edu/cgi/viewcontent.cgi?article=2969&context=hlr>>. Acesso em: 2020.

KACZYNSKI, Ted. **The Communiques of Freedom Club**. Disponível em: <<https://theanarchistlibrary.org/library/ted-kaczynski-the-communiques-of-freedom-club-ted-kaczynski>>. Acesso em: 30 out. 2019.

KACZYNSKI, Theodore John. **Industrial Society and Its Future**. 1995. Disponível em: <<http://editions-hache.com/essais/pdf/kaczynski2.pdf>>. Acesso em: ago. 2019.

LEON, Manuel de. ‘Manhunt: Unabomber’: um matemático com quatro prisões perpétuas. **El País**, jan. 2018. Disponível em: <https://brasil.elpais.com/brasil/2018/01/17/actualidad/1516203713_482238.html>. Acesso em: 10 set. 2019.

OLIVEIRA L. P. **Linguística de corpus**: teoria, interfaces e aplicação. Rio de Janeiro: Matruga, 2009. Disponível em: <<https://www.e-publicacoes.uerj.br/index.php/matruga/article/view/27796/19917>>. Acesso em: 10 jan. 2020.

PAULON, Andréa; NASCIMENTO, Jarbas Vargas do; LARUCCIA, Mauro Maia.

Análise do discurso: fundamentos teórico-metodológicos. **Diálogos Interdisciplinares**, São Paulo, v. 3, n. 1, p. 25-45, 2014. Disponível em: <https://www.academia.edu/24696076/An%C3%A1lise_do_Discurso_Fundamentos_Te%C3%B3rico-Metodol%C3%B3gicos>. Acesso em: 12 out. 2019.

SAUSSURE, Ferdinand. **Curso de linguística geral**. 28. ed. São Paulo: Cultrix, 2012.

SILVA, Thais Cristóforo. **Fonética e Fonologia do Português**. 11 ed. São Paulo: Contexto, 2017.

SOUZA, PEDRO. **Análise do discurso**. Florianópolis: LLV/CCE/UFSC, 2011. Disponível em: <https://petletras.paginas.ufsc.br/files/2016/10/Livro-Texto_Analise-do-Discurso_UFSC.pdf>. 2011. Acesso em: 10 dez. 2019.

TIMBANA, Alexandre. A linguística forense: um desafio para a investigação criminal no século XXI. **Revista Científica do ISCTAC**, Moçambique, v. 3, n. 7, p. 30-42, jan./mar. 2016. Disponível em: <https://www.researchgate.net/publication/311767989_A_LINGUISTICA_FORENSE_UM_DESAFIO_PARA_A_INVESTIGACAO_CRIMINAL_NO_SECULO_XXI>. Acesso em: 20 dez. 2019.

TURCHIE, D. Terry. Affidavit of assistant special agent in charge. **Court TV Library**, Distrito de Montana, 2008. Disponível em: <<https://web.archive.org/web/20081218190755/http://www.courtstv.com/archive/casfiles/unabomber/documents/affidavit.html>>. Acesso em: 25 jun. 2020.