

**CENTRO UNIVERSITÁRIO DAS  
FACULDADES METROPOLITANAS UNIDAS  
CURSO DE LETRAS – BACHARELADO EM TRADUÇÃO E INTERPRETAÇÃO  
(Inglês- Português)**

**LAURA MARIA FERNANDES DE CARVALHO**

**PROCESSO E CRIAÇÃO DE MÓDICO GLOSSÁRIO BILÍNGUE NA ÁREA DE  
CRIMINALÍSTICA COM BASE EM LINGUÍSTICA DE CORPUS**

**SÃO PAULO**

**2020**

LAURA MARIA FERNANDES DE CARVALHO

PROCESSO E CRIAÇÃO DE MÓDICO GLOSSÁRIO BILÍNGUE NA ÁREA DE  
CRIMINALÍSTICA COM BASE EM LINGUÍSTICA DE CORPUS

Projeto apresentado ao Programa Institucional de Iniciação Científica e Centro Universitário das Faculdades Metropolitanas Unidas como uma pesquisa na obtenção de efetivos resultados relacionados ao curso de Tradução e Interpretação (Inglês e Português) em conjunto ao arcabouço teórico de Investigação Forense e Perícia Criminal, exercendo foco na subdivisão Criminalística.

Orientadora: Prof.<sup>a</sup> Alessandra Cani Gonzalez Harmel.

SÃO PAULO

2020

## **AGRADECIMENTOS**

Sou grata à minha instituição de ensino por proporcionar a experiência de expandir meus conhecimentos por meio deste projeto que me iniciou no caminho rumo ao amadurecimento acadêmico.

Também à minha orientadora que esteve presente em todas as etapas, auxiliando diretamente na elaboração e contribuindo sempre com seus notáveis saberes, além de manter acesa a confiança em mim para concluir a pesquisa.

Sem hesitação, à minha família, que sempre investiu em meus estudos acreditando que através deles eu poderia chegar aonde quisesse, não importando o grau de dificuldade.

"Um dos maiores requisitos para possibilitar uma investigação oficial e trabalhar com precisão é o profundo conhecimento dos homens. Assim, tudo na vida pode ser útil: cada conversa, cada frase concreta, cada palavra dita por acaso, cada ação, cada desejo, cada traço de personalidade, cada ato da conduta, cada olhar ou cada gesto que observamos em outros".

- Dr. Hans Gross, considerado o pai da criminalística.

## **RESUMO**

Nesta produção, procuramos demonstrar e analisar significados terminológicos traçando um comparativo entre os termos utilizados na área de especialidade relacionada à Criminalística, nos idiomas inglês dos Estados Unidos e português do Brasil, considerando as especificidades e diferenças nos sistemas de ambos os países. Mantemos a atenção à norma culta das línguas, visando sempre a compreensão e facilitação do processo tradutório não somente para especialistas, mas qualquer pessoa que se interesse por essa área de conhecimento.

O processo se deu com busca canalizada em setores do campo criminalístico, para obtenção da coleta de termos esclarecidos, dentre esses: notáveis conceitos, materiais utilizados pela perícia durante a investigação, nomeação dos vestígios encontrados em cena de crime, exames posteriores, desenvolvimento forense sobre o percurso do crime até obtenção de resposta, entre outros.

Intentamos esclarecer com a elaboração de um breve glossário, termos não compreendidos por falantes da língua portuguesa e/ou inglesa, com o propósito de sanar quaisquer possíveis dúvidas acerca do sentido que as palavras dispostas nele buscam expressar.

Palavras-chave: criminalística, terminologia, linguística de corpus, glossário bilíngue.

## **ABSTRACT**

In this production, we aim at presenting and analyzing terminological meanings by drawing a comparison between the terms used in the area of speciality related to Criminalistics, in American English and Brazilian Portuguese, considering the specificities and differences in the systems of both countries. We try to prime at the educated norm of languages aiming at understanding and facilitating the translation process not only for specialists, but anyone who is interested in this area of knowledge.

In the process we emphatically researched relevant sectors of the criminal field in order to obtain the collection of terms to be studied and clarified together; among these are: remarkable concepts, materials used by the expert during the investigation, naming of the traces found at the crime scene, subsequent exams, forensic development on the route crime until a response is obtained, and others.

Our intent is clarification by elaborating a brief glossary, with terms unknown to Portuguese and / or English speakers, with the purpose of solving possible doubts about the meaning of the words and terms presented

Keywords: criminalistics, terminology, corpus linguistics, bilingual glossary.

## LISTA DE FIGURAS: Inglês (EUA)

Figura 1.....	The BootCat Frontend
Figura 2.....	Project Definition
Figura 3.....	Proceed
Figura 4.....	Search Engine
Figura 5.....	Seeds
Figura 6.....	Advanced options: Scielo
Figura 7.....	Queries
Figura 8.....	Word List
Figura 9.....	Keyword
Figura 10.....	Collocates
Figura 11.....	Concordance
Figura 12.....	Concordance Plot

## SUMÁRIO

1. INTRODUÇÃO.....	09
2. EMBASAMENTO TEÓRICO.....	10
2.1 Linguística de Corpus.....	10
2.2 Terminologia.....	11
2.3 Fraseologia.....	12
3. CRIMINALÍSTICA.....	13
3.1 Conceito.....	13
3.2 Breve contexto histórico.....	14
3.3 Aplicação na área de Perícia Forense.....	14
4. METODOLOGIA.....	15
4.1 Corpus.....	15
4.2 Escolha do tema.....	16
4.3 Preparação e elaboração.....	17
4.4 BootCat.....	17
4.5 AntConc.....	22
5. ANÁLISE.....	26
6. DIFICULDADES E APRENDIZADOS.....	26
7. CONSIDERAÇÕES FINAIS.....	27
8. REFERÊNCIAS BIBLIOGRÁFICAS.....	28
9. APÊNDICE: GLOSSÁRIO.....	29



## 1. INTRODUÇÃO

Por se tratar principalmente de área que consiste em amplo embasamento teórico e das humanidades, a natureza da pesquisa será de caráter explicativo, na qual definições serão propostas por meio da elucidação de termos possivelmente desconhecidos na área até o momento.

O delineamento do projeto apresentará teor bibliográfico, onde foram utilizadas ferramentas computacionais que auxiliaram diretamente na análise de terminologia, que se mantém como foco da pesquisa. Dentre estes, cito como principal, o software AntConc desenvolvido por Laurence Anthony, que possui notável prestígio no que se desenvolve na área de ato tradutório.

Necessário ressaltar que, de forma divergente, o glossário aqui apresentado não funcionará como um dicionário, mas sim como um modelo que destaca os termos técnicos recorrentes obtidos durante a pesquisa e coleta de textos disponíveis na temática abordada.

Quanto à perspectiva de tratamento dos dados, seguirá de forma qualitativa, unindo informações com suas notáveis funções explicativas.

Com a intenção de instigar os leitores do projeto a conhecerem melhor um pouco da área de Criminalística, e aprofundar também nossos conhecimentos a respeito dos dois campos de estudo, sendo estes a Tradução e Interpretação interligada à Perícia Forense, realizaremos então esta Iniciação Científica.

## 2. EMBASAMENTO TEÓRICO

Para início de pesquisa se fez necessária escolha de base para elaboração, desenvolvimento e criação do glossário. Esta foi dividida em três tópicos, os quais serão apresentados a seguir.

### 2.1 Linguística de Corpus

Como principal ferramenta metodológica, o do modelo aqui criado foi a Linguística de Corpus (LC), que Berber Sardinha define da seguinte forma:

[...] 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 por computador. (2004)

Tendo em vista que a palavra *corpus* descende do sentido de corpo, abrangendo composição relacionada a conjunto de documentos, é possível definir que o termo data período anterior ao da invenção computacional. Documentos comprovam na verdade, que até mesmo na Idade Média, citações da Bíblia cristã rendiam o que hoje nomeamos como *corpora*.

Levando também em consideração os tempos antigos, se torna evidente o fato de que para coleta de arquivos não havia auxílios tecnológicos ou ferramentas que de alguma forma pudessem substituir o trabalho humano, sendo então necessário o processo totalmente manual.

Nota-se que neste período a relevância destes documentos se dava no âmbito de ensino da língua e de seus idiomas.

Em 1950, a *Syntactic structures* surgiu como material produzido por Chomsky, nome de grande prestígio dentro da área de Linguística, inovando no que consiste em algo que anteriormente já constava como verdade absoluta. Temos aqui a sustentação de documentos com base em *corpora* com como na linguística gerativa. Leva-se então em consideração as probabilidades, no que tange aos termos e a quantidade de vezes nas quais ele pode aparecer no uso da língua, realizando-se uma análise de padrão nas expressões.

Já em 1980, a LC foi fortalecida com o desenvolvimento do computador e popularização dos *corpora*, que modernizou pesquisas anteriormente muitas vezes equivocadas, visto que os seres humanos são passíveis de erros e pela carência de consistência nos dados, algo que naturalmente poderia comprometer o trabalho. Muitos recursos de processamento desta forma foram disponibilizados

de fácil acesso, enriquecendo também conhecimentos e ações acerca desta área de estudo.

A tecnologia para esta área foi amplamente benéfica, facilitando a rotina de profissionais e encurtando o tempo gasto na conclusão de todo o percurso, desde a coleta e limpeza até a escolha final do material a ser utilizado chegando então finalmente nas produções que hoje conhecemos.

Atualmente, LC apresenta papel fundamental na área da linguagem e exploração dentro de variações e expressões linguísticas, além de ser uma frequente abordagem utilizada para o ensino e prática de idiomas a serem aprendidos. Evidentemente, é tido como principal auxiliar também à elaboração de glossários, dicionários e conteúdos que de alguma forma são informativos ao público.

Mantém-se em expansão técnica e aprimoramento sempre na enfática busca por maiores e melhores usos, tendo inclusive grande relevância nos Estados Unidos da América, onde existem notáveis centros de pesquisas voltados a essa questão. Grandes corporações bem equipadas fazem uso desta modalidade e tornam-se melhor creditados por seus conhecimentos, principalmente pelo fato de que hoje em dia ser tão valorizado o saber técnico de determinadas perspectivas virtuais baseadas em banco de dados e textos produzidos para diversas áreas do comércio. Como exemplo, citamos aqui a própria Microsoft que se utiliza desse método para organização e aplicações dentro de seus meios de atuação.

## **2.2 Terminologia**

Pode ser definida como um campo de estudo e pesquisa que pode, dentro de suas possibilidades, gerar produtos terminológicos que resultam em compartilhamento de conhecimentos, dentre esses sendo possível citar dicionários técnicos e científicos, bancos de dados, glossários específicos de determinadas áreas de especialidade e conclusões de processos que de alguma forma visem expandir o arcabouço teórico daqueles que se interessarem pela busca de determinado assunto.

Funciona de forma a proporcionar estudo sistemático de termos e conceitos específicos baseados nas mais diversas áreas de estudo. Auxilia principalmente profissionais que se encaixam na área como terminólogos e estudantes da língua, além dos próprios tradutores, que por seu uso obtém a possibilidade de integrar conteúdo externo ao seu material já adquirido, para produzir cada vez melhores

substituições entre as línguas, que sejam pertencentes a quaisquer idiomas já existentes.

Importante ressaltar o nome de maior influência dentro deste assunto que hoje se mostra de grande importância para, por exemplo, o próprio meio acadêmico do qual fazemos uso para concluir esta parte de nossa escrita: Eugene Wüster.

Eugen Wüster (1898-1977), engenheiro austríaco responsável por fundamentar campo de investigação cada vez mais dinâmico na linguística contemporânea, executou como especialidade nos domínios da lexicologia e lexicografia, a Teoria Geral da Terminologia, ou TGT, que como seus princípios apresenta a normalização de conceitos e termos em escala nacional e internacional. Realizou seus feitos por sentir a carência na normalização de termos de sua área, buscando então garantir êxito na comunicação profissional e de alguma forma transmitir os conhecimentos entre os profissionais do meio.

Deve também ser compreendido que para abordar este tópico, se faz necessário o conhecimento de sua base: o termo. Este pode ser nomeado também como unidade terminológica, o termo trata de ser um vocábulo responsável por “designar qualquer conceito específico dentro de um domínio de especialidade”, como afirmou (BARROS, 2004).

Com a intenção de simplificar, seria possível definir brevemente a palavra como sendo produto da língua comum e o termo algo pertencente a língua específica. Um exemplo claro seria a palavra “identificação”, que no dicionário léxico apresenta definições como um processo pelo qual duas coisas ou mais são consideradas idênticas ou similares. Além de: “reconhecimento da identidade de algo ou alguém”.

Dentro da Criminalística temos a seguinte conclusão: “Afirmação de que duas impressões digitais foram produzidas por um mesmo dedo. Quando, pelo exame de pontos característicos, o perito se convenceu de que se trata da mesma impressão papilar” (ARAÚJO, 1960, p.225).

Dentro desse estudo, também deve ser perceptível a importância de ressaltar a contextualização de termos, algo que (Cabré, 1995) definiu como TCT, ou Teoria Comunicativa da Terminologia. Podendo ser entendido brevemente como a ressignificação da simples palavra perante uma nova comunicação especializada, surge a necessidade de considerar todas as adversidades e circunstâncias presentes na construção de novos discursos e fundamentação para o tópico abordado a seguir: a fraseologia.

## **2.3 Fraseologia**

Situada na área de estudos lexicais e podendo ser considerada uma subdisciplina da lexicologia, a fraseologia se ocupa das combinações das unidades léxicas, constituídas por mais de duas palavras gráficas. Tem-se limite superior, sendo este denominado como frase.

Em glossários e dicionários, se mostram de suma importância para o profissional que traduz textos, visto que permite a este verificar de que maneira o termo interliga-se no contexto frasal. Evidentemente, também apresenta ao tradutor as possibilidades dentro de sua área específica de atuação naquele determinado momento.

É certo que as unidades fraseológicas se caracterizam pela polilexicalidade. Não existe consenso geral quando questionado o âmbito de estudo da área em questão, mas sim muitas definições que podem ser encontradas tanto em obras clássicas quanto contemporâneas.

Segundo Marçalo, podemos considerar que:

“Os avanços no domínio da investigação fraseológica serão de enorme utilidade no ensino de português como língua materna e em especial não-materna. Contribuirão ainda para uma feitura de dicionários da língua portuguesa mais cuidada e consentânea com a realidade da língua. (2009)

## **3. CRIMINALÍSTICA**

Para melhor entendimento da produção criada a partir deste projeto, segue abaixo a descrição sobre este que é campo da perícia imprescindível para elucidação de crimes na contemporaneidade.

### **3.1 Conceito**

O termo oficialmente foi utilizado pela primeira vez no ano de 1893 por Hans Gross. De descendência alemã, respondia aos cargos de jurista e magistrado sendo professor de Direito Penal responsável por lecionar na Universidade de Graz. Lá, publicou obra de notável importância para o meio de estudos em questão, intitulada *Criminal Investigation: A Practical Handbook*. Segundo ele, temos que: “É a análise de vestígios materiais extrínsecos relativos ao local periciado, relacionando modus operandi com a dinâmica descritiva”.

Podendo ser atribuída a uma das divisões da Ciência Forense, a Criminalística apresenta sua pesquisa voltada para a análise de vestígios e materiais diretamente relacionados ao local de crime. Isso, mantendo intrínseca relação entre o modus operandi e a dinâmica descritiva, oferece assim fundamentação material à instrução penal.

Como objetivo principal, interpreta através da aplicação de diversas ciências o que popularmente é conhecido como provas, visando sempre esclarecer de maneira técnica os problemas criminais relativos à existência de delito: qualificação deste, identificação do responsável pela execução dos fatos, legalização e perpetuação das provas.

### **3.2 Breve histórico**

É perceptível que desde os tempos mais remotos, a humanidade busca encontrar diferentes maneiras de esclarecer delitos, sendo estes de qualquer natureza. De fato, há muita controvérsia quando questionado o início oficial da investigação forense e, de maneira geral, é possível afirmar que na verdade a científica aplicada às questões legais, teve início na medicina legal. O médico nesses períodos realizava exame no local de crime, restringindo todas suas possibilidades ao próprio cadáver encontrado.

Considerando o não esclarecimento total dos fatos que este procedimento incitava, surge então a necessidade de esclarecer o motivo da morte, levando o profissional da área a pesquisar outros elementos que tentassem de alguma forma explicar a dinâmica dos fatos de maneira mais aprimorada, visto que, na época citada, a necropsia ainda não era uma válida opção.

### **3.3 Aplicação na área da Perícia Forense**

É certo que para ilustrar de maneira a torná-la mais compreensível, deve-se consciência de que a Perícia Forense apresenta suas divisões distribuídas sob diferentes ramificações, sendo que, todas, de alguma forma, se completam. Como principais, pode-se apontar a Química, a Biologia, a Física, a Medicina, a Balística, a Fonética, a Informática e a Genética. Todas auxiliam diretamente no processo que se encaminha para obtenção de respostas.

É de extrema importância considerar as adversidades que envolvem a área de Criminalística, parte essencial da Perícia, para que os resultados sejam de fato conclusivos. A preservação e isolamento do local, por exemplo, deve ser sempre

tópico relevante. Além disso, o conhecimento pleno de termos utilizados para a designação dos exames e instrumentos utilizados durante todo o processo é condição *sine qual non*.

## **4. METODOLOGIA**

### **4.1 Linguística de Corpus**

Sendo naturalmente conhecido como um conjunto de textos ou documentos voltados para determinado tipo de pesquisa na língua, é necessário o entendimento de que o termo abrange muitas outras ramificações que auxiliam a completar o todo deste que é instrumento de grande importância acadêmica.

Novamente, considerando Berber Sardinha e sua relevância no que consiste ao tópico apresentado, temos a seguinte afirmação, descrita em sua obra intitulada *Linguística de Corpus*:

(...) o conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou ambos), sistematizados de acordo com determinados critérios, suficientemente extensos em amplitude ou profundidade, de maneira que sejam representativos da totalidade do uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por um computador com a finalidade de propiciar resultados vários e úteis para descrição e análise. (2004, p.18)

Desta forma, é possível analisar que trata-se então de algo que representa a linguagem independente do idioma ou da variedade à qual pertence. Isto é, levando sempre em consideração o tema e a forma pela qual a pesquisa deverá ser realizada. Fatores interferem diretamente na conclusão de um projeto, então, a atenção deve estar principalmente evidente quanto aos textos previamente coletados, já que estes estabelecem a base necessária para todo o percurso a ser seguido.

Seleção, conteúdo, disponibilidade, autoria e finalidade das escolhas serão de suma importância. Berber Sardinha, em sua própria produção, levanta notáveis tópicos a serem considerados: Origem dos arquivos, sempre autênticos; composição e limpeza de conteúdo para que o foco não seja perdido; propósito bem definido, sendo objeto de estudo linguístico; representatividade, de variedade ou língua; extensão, quanto maior, melhor encaixado nos padrões que atendem a pesquisa representativa; formatação, ponto crucial na entrega de todo o projeto, deve ser desenvolvido de forma computacional e respeitando as necessidades do trabalho pelo qual está realizando a pesquisa.

Como no próprio campo de Perícia Forense, não se pode simplesmente fazer de tudo para obter seu material, visto que isso não seria tecnicamente aceito. Nesses casos, os meios justificam os fins; tanto que se alguém obteve prova por meios ilegais, ela não será aceita perante tribunal, assim como qualquer conteúdo obtido através de qualquer simples pesquisa virtual não produzirá resultados considerados realmente satisfatórios.

Importante reiterar que é o intuito da pesquisa e que deve definir objeto e os tipos de textos a serem colhidos. Disponibilizando uma ilustração clara, é interessante pensar da seguinte forma: Se o caráter da pesquisa for direcionado para elaboração de um glossário, e como este foi criado, as produções coletadas deverão ser plenamente escritas.

Nota-se que, não é porque se trata de um conjunto de dados que este se concretiza como corpus. É preciso ter a consciência de que apresenta linguagem natural e usual, não sendo previamente produzido com intenções exteriores, para uso em pesquisas científicas da natureza proposta.

Idealmente, nenhuma etapa do processo deve ser escolhida de forma aleatória ou de maneira apressada. O processo de compilação dos textos que constituem o corpus e a sua aplicação deve ser de um trabalho de considerável técnica.

## **4.2 Escolha do tema**

Como tópico primordial para início da pesquisa, se faz necessário o tema para o qual iremos acrescentar conteúdo. Desenvolver um glossário requer um objeto de estudo para que os textos da área sejam coletados, limpos e processados de forma computacional, utilizando-se a metodologia proposta, que é a Linguística de Corpus.

A elaboradora deste projeto sentiu-se já motivada quando iniciou os estudos como tecnóloga de Investigação Forense e Perícia Criminal no ano de 2018 e a de Tradutor e Intérprete em 2019. Por conseguinte, foi ideal a união dos dois neste ano de 2020. Um início de ouro para projetos e futuro profissional.

Para aqueles que de alguma forma apresentam contato com a área de estudos criminalísticos, é perceptível que o material estrangeiro apresenta sempre grandes conteúdos que só teriam a acrescentar nos documentos nacionais, mas, muitas vezes, não são aqui oferecidos por escassez de técnicos a contribuírem com essa proposta. Seguimos então na busca e desejo por fazer parte deste campo, que por mais específico que seja, apresenta incontáveis conhecimentos e novas descobertas a serem ainda realizadas.



De fato, a Perícia Forense é área bastante vasta, dotada de inúmeras ramificações que buscam sempre atingir o objetivo de solucionar novos casos ainda não compreendidos ou carentes de justiça. Mas, considerando todos os processos e a maneira como se interligam com o fato consumado, a Criminalística tem sua porcentagem no ápice dentro das preferências da elaboradora. Eis então o motivo de explorar o máximo possível as diretrizes pelas quais esta nomeia seus materiais e conteúdos disponíveis em seu arcabouço técnico. Buscando textos que alcancem o máximo de dados, para formar representatividade teórica e possibilitar a organização de termos através de critérios, tentaremos realizar nesta pesquisa a junção necessária para criação de corpus específico.

### **4.3 Preparação e Elaboração**

Posteriormente à definição do tema, o primeiro tópico a ser considerado relevante trata-se da coleta de textos, estes sendo originais e estando disponíveis nos dois idiomas de análise: inglês dos Estados Unidos e português do Brasil. Considerando sempre os critérios previamente esclarecidos neste documento, foi seguido um padrão de tópicos para que as produções dispostas fossem adequadas ao nosso assunto e intenção de pesquisa.

Resolvemos então, através de dez textos escolhidos em cada idioma, extrair cerca de 20 à 25 termos para darmos início aos significados de cada um deles e definir também suas respectivas fraseologias. Reiteramos também, que todas as etapas do processo foram levadas em consideração a partir da necessidade de nosso glossário, desde o programa responsável por auxiliar na coleta direcionada à artigos acadêmicos até número, data de publicação, autoria e formato que consistia cada um de nossos achados.

De maneira padronizada, foi criado dentro do programa escolhido, um processo sistemático que facilitou e concluiu nossa corpora de criminalística com grande êxito, finalizando então as etapas mais teóricas e dando início a prática de produção de nosso módico glossário.

Importante também ressaltar a necessidade de limpeza de cada texto, através da conversão de formato PDF para TXT, não permitindo, dessa forma, que nenhum arquivo ou elemento indesejado (como exemplo tabelas e numerais), estivesse presente na próxima etapa de realização.

## 4.4 BootCaT

Trata-se de um extrator automático de corpus e de termos. O *Bootstrapping Corpora and Terms*, disponível no atual momento somente em língua inglesa, propõe a montagem do corpus através de uma maneira interativa com o pesquisador a partir de textos obtidos em websites.

Composto por várias ferramentas projetadas para executar pequenas partes do processo de montagem do corpus, ele é responsável por realizar quatro etapas que no total concluem a formação de corpora.

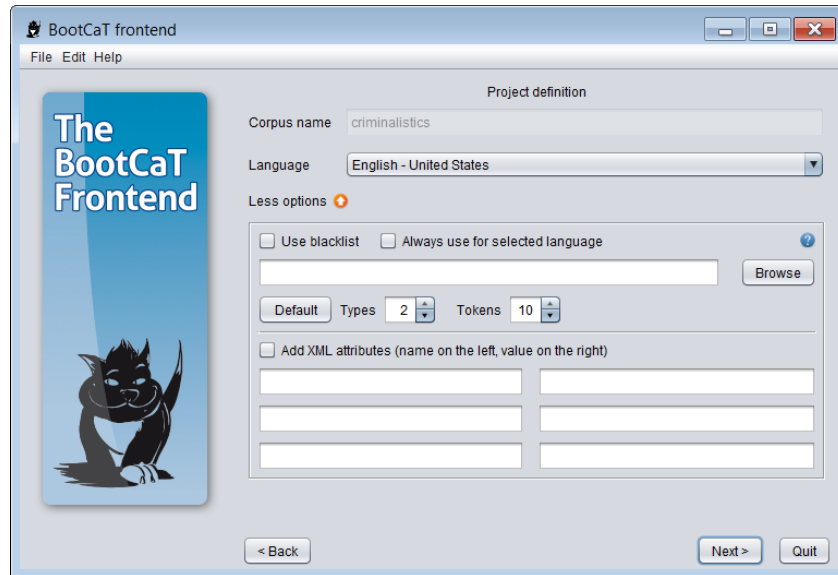
É notável que o primeiro processo se constrói automaticamente, a partir de buscas no Google, que utilizam determinados conjuntos de itens léxicos denominados como sementes. Estas, o próprio pesquisador tem como objetivo disponibilizar no programa, direcionando sua pesquisa para a área de atuação.

Na sequência, ele extrai novas sementes desse corpus e posteriormente utiliza essas para novas buscas na plataforma de pesquisa, cujos textos recuperados serão adicionados ao corpus, de forma a aumentá-lo. Ele também extrai novas sementes desse corpus complementando-o, e assim por diante.

O programa contém um tutorial de apresentação e instruções para auxiliar o pesquisador na obtenção de seu novo corpus. A princípio ele pede para você nomear seu novo projeto e selecionar o idioma pelo qual você deseja obter material. Em nosso caso, temos o seguinte:



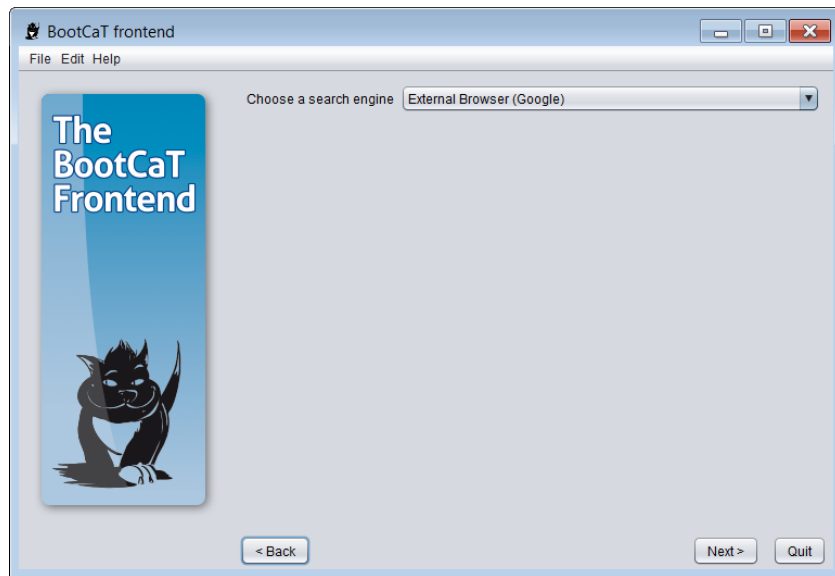
- I. Introdução do programa, responsável por apresentar e guiar o criador no início do trajeto de pesquisa e formação de corpora.



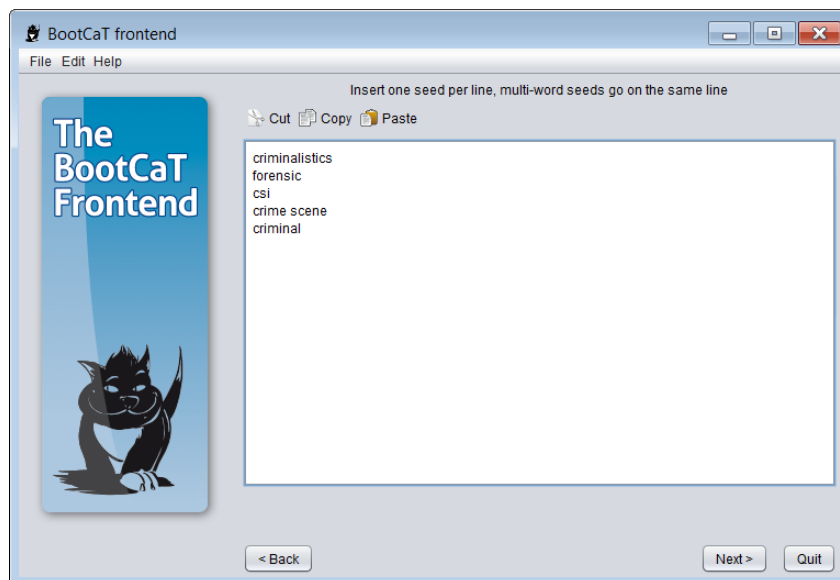
- II. Posteriormente, temos a opção de nomear nosso projeto e selecionar o idioma no qual devemos trabalhar. Importante reiterar que o processo se repetiu em inglês e português, para que o glossário bilíngue tivesse material disponível em ambas as linguagens. Neste momento, também selecionamos a quantidade de sementes a se complementarem na posterior busca online.



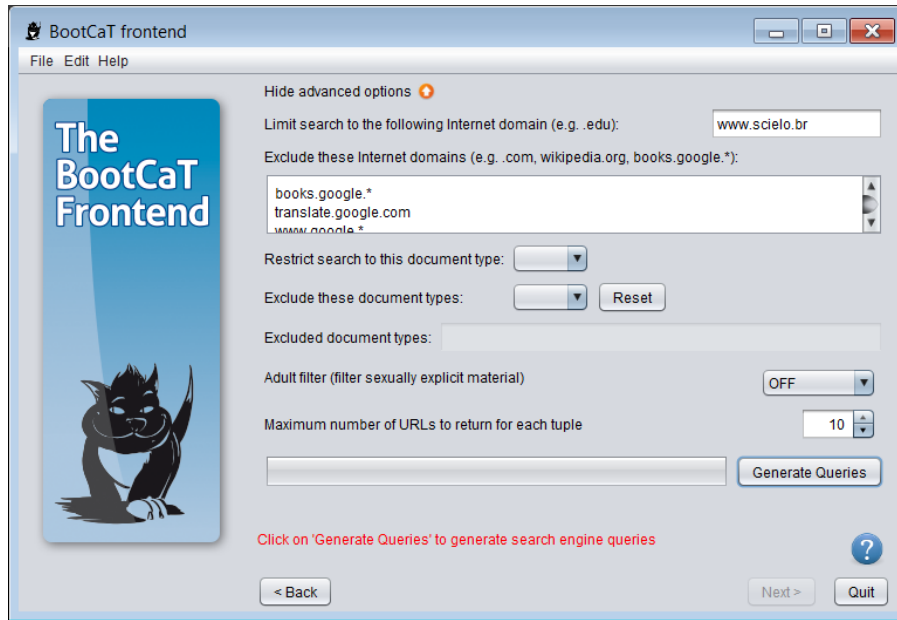
- III. O processo também pode ser definido através de diferentes alternativas, em nosso caso, foi escolhida a primeira opção visando facilitar a sequência para obtenção de resultados significativos.



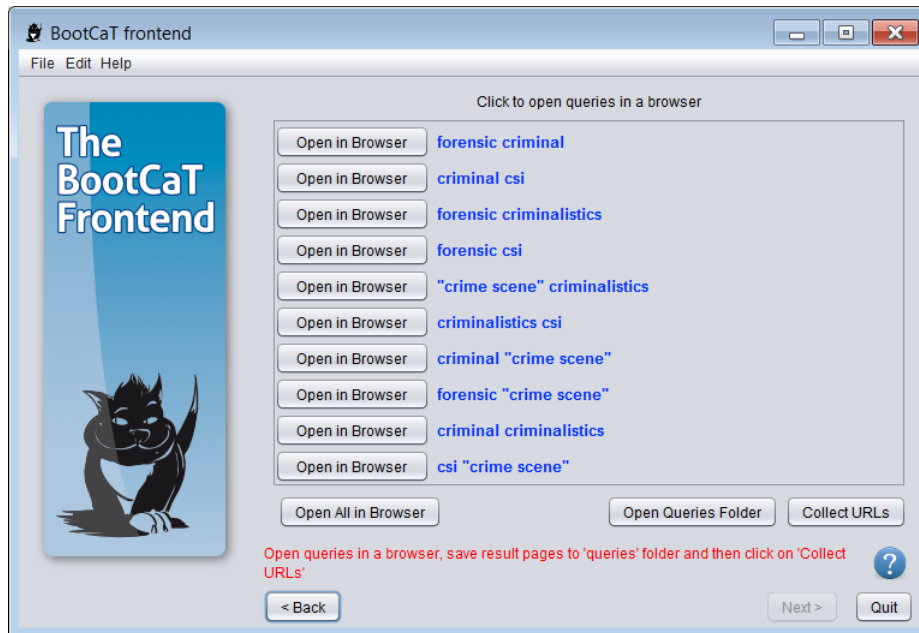
IV. Escolhemos, a princípio, o Google como site padrão.



V. Depois, definimos os cinco termos apresentados na imagem acima, para coleta direcionada em nossa área de interesse.



- VI. Por se tratar de pesquisa científica, escolhemos aqui afinilar nosso conteúdo através da escolha do website SCIELO, que em sua base apresenta conteúdo principalmente acadêmico, nos auxiliando desta forma a unir material de relevância estritamente universitária.



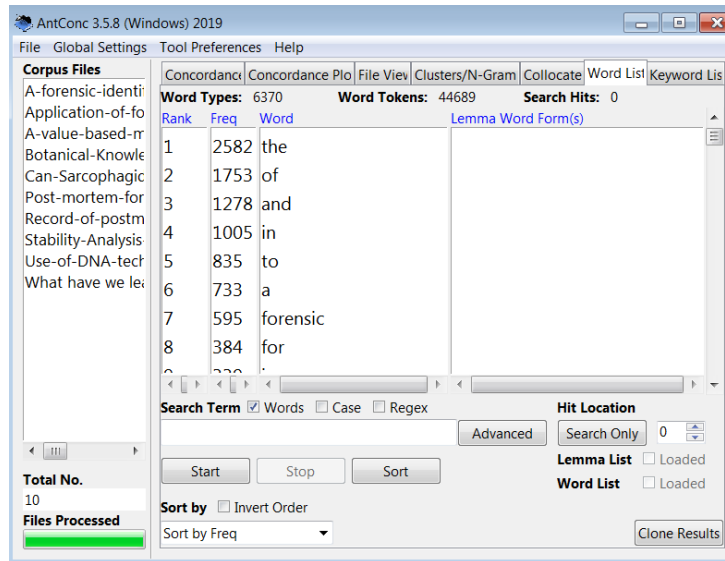
- VII. Por último, direcionamos nossas sementes para prática pesquisa na própria web, salvando todas as conclusões geradas a partir das sementes e salvando os resultados em nossos arquivos computacionais (separando-os por idioma) para dar início a próxima etapa, que consiste em incluir os textos aqui obtidos na plataforma de concordância “AntConc”, sob a qual estabelecemos maiores explicações no seguinte tópico.

#### **4.5 AntConc**

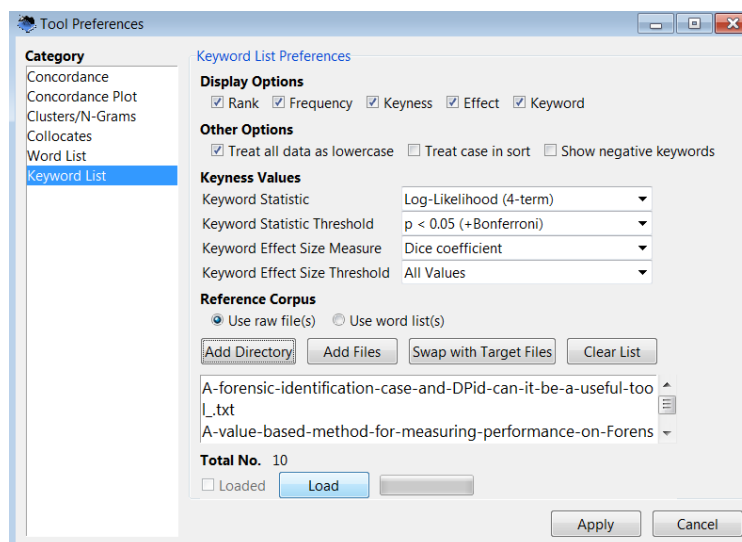
Como escolhida a Linguística de Corpus para elaboração do projeto, foi necessária a utilização de um concordanciador para auxiliar no processo que consiste em listar a ocorrência de determinados termos e suas respectivas ramificações perante os textos escolhidos. A frequência de determinada palavra foi de suma importância para obtermos resultados satisfatórios na questão de coletar as palavras a serem posteriormente esclarecidas.

O *software* foi desenvolvido por Laurence Anthony, da *Faculty of Science and Engineering – Waseda University* e o uso dessa plataforma foi necessário para facilitar nossa compreensão sobre associações entre itens lexicais e gramaticais nas mais diversas conotações, nos auxiliando diretamente na identificação do padrão que foi criado a partir da corpora criminalística, direcionando nossos segmentos para início efetivo do glossário.

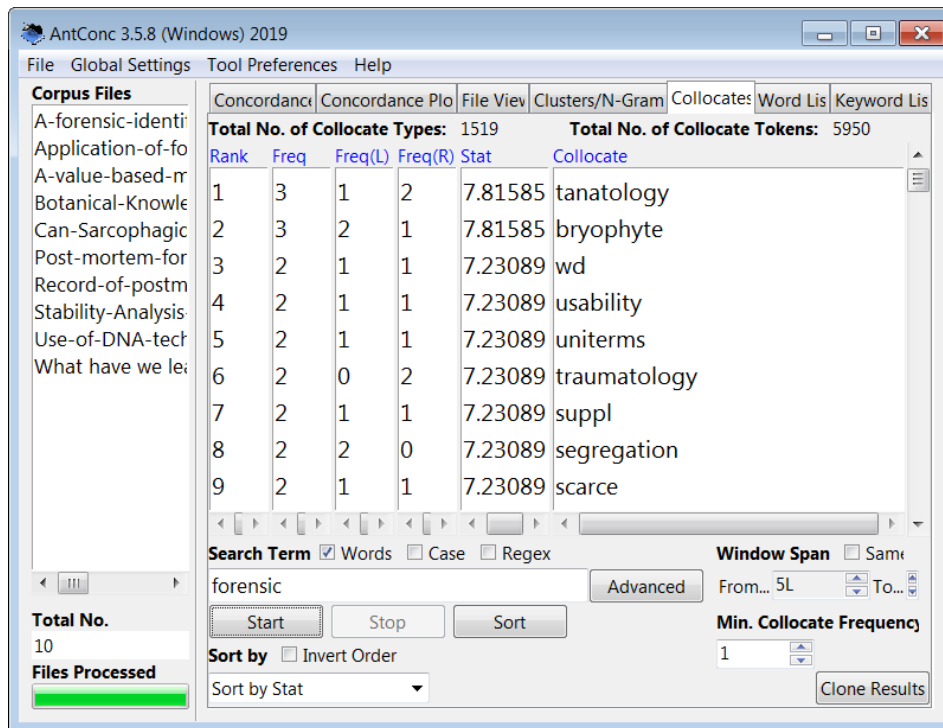
O primeiro passo foi instalar o programa AntConc no computador, sendo isso possível através de acesso ao site <http://www.laurenceanthony.net/software/antconc/>. Posteriormente, iniciamos na ferramenta a leitura dos arquivos coletados até então, através da opção “*Open file, Word List e Start*”, respectivamente .



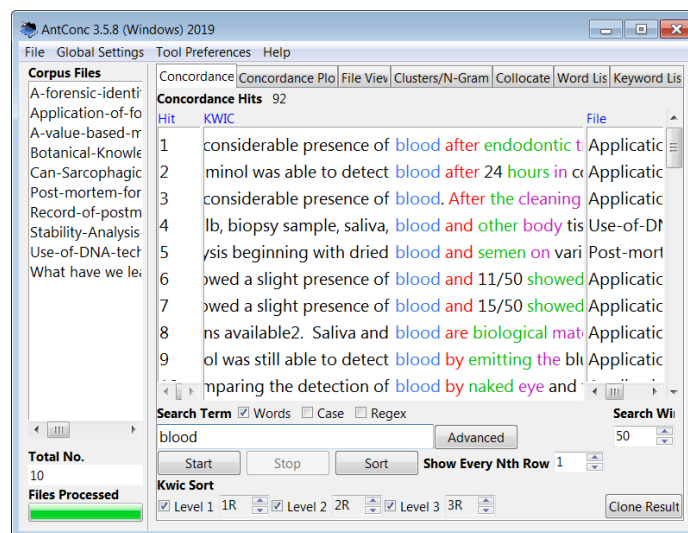
Depois de absorver os textos, o programa ilustra todas as palavras presentes no corpus, sendo possível a partir de então a interação do pesquisador na escolha de formato sob o qual deseja visualizar o conteúdo. Por se tratar de glossário, seguimos para *Keyword List*, onde tivemos a chance de fato escolher os termos para compor nossa pesquisa.



Cr terios foram adotados ap s a listagem de palavras-chave, isso por meio da pr pria ferramenta de concord ncia, que dentro de sua formata o disponibiliza incont veis formas de explorar o conte do disponibilizado. Importante reiterar que tamb m foi utilizada a aba de *collocates* para j  escolher os poss veis candidatos  s fraseologias, que no gloss rio se encaixa na aba de contextos onde os termos podem ser utilizados.

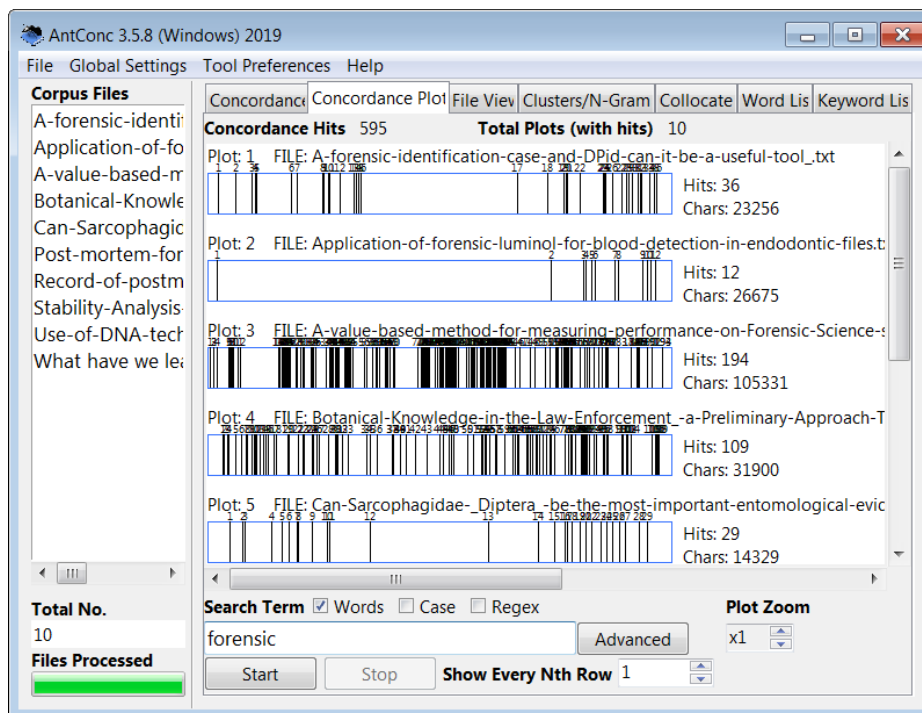


Também devemos considerar interessantes as cores que evidenciam as principais ligações presentes ao termo escolhido, dando destaque para que possamos melhor visualizar cada contexto presente na corpora.



Certamente outros fatores também vão se apresentando no decorrer da utilização, o que nos possibilita enriquecer a atividade de maneira satisfatória quando relacionada a novos conhecimentos técnicos. Em *Concordance Plot*, por exemplo, foi possível visualizar graficamente a frequência de termos buscados nos vários textos do corpus.





É perceptível que sem a utilização do programa o processo seria muito mais trabalhoso, levando ainda mais tempo para ser finalizado. Aqui o processamento automático das palavras e as possibilidades de diferentes listagens permitem um fácil entendimento e organização de termos, o que torna o processo etapa fundamental do projeto.

Deve se ter a compreensão de que a seleção das palavras e das fraseologias aqui dispostas não é executada pelo programa, mas sim preparada para que o pesquisador utilize-a como mecanismos de listagem, podendo dessa forma escolher como aplicar seus resultados. Em nosso caso, a escolha foi definida baseada em alguns tópicos, entre eles o termo devia estar visível em pelo menos 5 textos, constar como palavra-chave, estando presente na *word list*, e ter frequência mínima de 3 vezes no corpus de estudo.

No fim foram coletados 10 textos de cada idioma, e é primordial ressaltar que após a identificação dos termos em língua inglesa, seguimos para a busca por equivalência na língua portuguesa, isso principalmente por saberes relacionados ao tema e padrões de pesquisa, além de consultas diversas.

## 5. ANÁLISE

O glossário foi então elaborado de maneira a preencher os cinco tópicos de relevância em cada idioma. Sob o padrão estabelecido, fomos capazes de aferir 25 termos no total, capazes de compor nosso módico projeto. Em cada coluna traduzimos as palavras e contextualizamos para que o leitor possa se inserir na temática de maneira fácil e direta. Quanto à coluna de observação, escrevemos a frequência de determinado termo nos textos coletados.

## 6. DIFICULDADES E APRENDIZADOS

Durante a pesquisa, consideráveis dificuldades foram surgindo. Como início, é possível citar a que se refere à utilização das ferramentas escolhidas para auxílio tecnológico. Na prática, o uso dos programas demandaram notável preparação e estudo de funcionamento para que nenhuma parte do processo fosse prejudicada.

Outro tópico importante a ser mencionado deve ser a coleta dos textos em língua portuguesa. É perceptível que no Brasil, estudiosos da área tenham o conhecimento de que muito conteúdo de relevância acaba sendo importado e não traduzido por escassez de profissionais direcionados para este ramo especificamente. A pesquisa para elaboração deste glossário funcionou como uma prova bastante clara sobre esse ponto, visto que os textos em inglês abordaram maior número de termos buscados do que o português, que também acabou sendo de certa forma menos específico e mais abrangente em sua proposta.

Notamos também que a maior quantidade de termos relacionados à criminalística são referentes em grande parte aos institutos e laboratórios responsáveis por promover justiça, mas isso, somente como menção, não de forma aprofundada na ilustração de casos detalhados, como podemos ver em textos disponíveis em língua inglesa.

Interessante ressaltar também que alguns termos presentes neste glossário, como por exemplo "*research*", em sua essência fazem parte do vocabulário geral em ambos os idiomas, o que nos fez trabalhar com foco maior na fraseologia, direcionando a palavra na contextualização de perícia forense e tendo a consciência de que sua repetição nos textos é devido à suma importância que o significado deste apresenta na criminalística.

Conciliar cotidiano com a pesquisa acabou também se tornando ponto de dificuldade. Realizando o curso de tradução na manhã, estágio no período da tarde e o curso de perícia à noite, fez com que muitas pendências se evidenciassem ao mesmo tempo, exigindo grande atenção, de todos os lados.

Finalizar o projeto se tornou uma conquista pessoal e aprendizado enriquecedor no que consiste na organização de tarefas e tempo pessoal.

## **7. CONSIDERAÇÕES FINAIS**

Essa Iniciação Científica foi realizada como um contato mais aprofundado com o meio acadêmico, gerando sem dúvida alguma, novos aprendizados e desafios que influenciaram diretamente em nosso crescimento acadêmico e pessoal como participantes de um ambiente universitário.

De fato uma experiência enriquecedora, ainda mais considerando o inconveniente de uma pandemia mundial ter surgido durante a execução do projeto. Isso nos fez criar uma rotina diferente da qual nos programamos, tendo que manter contato somente online, através de troca de e-mails, mensagens e reuniões por vídeo conferência. Todo o processo, em suas pequenas etapas, foi de grande importância para ambas às partes, visto que nos acrescentou experiência e novos conhecimentos práticos e teóricos.

Diversas descobertas foram realizadas no percurso deste trabalho e isso foi possível em grande parte a tecnologia que nos é ofertada diretamente com seus incontáveis avanços. Importante reiterar isso, pois tanto a tradução quanto a perícia, diariamente precisam ser executadas através do constante estudo e aperfeiçoamento nas técnicas.

Sempre terá relevância em conteúdo exterior à nossa realidade, e o papel do tradutor, desde a bíblia, será requisitado na mesma constância. Por ser inevitável, a morte e a existência de pessoas destinadas a homicídio, seja por transtornos ou convívio social, a criminalística sempre terá seu papel de busca pela verdade como grande necessidade para acalantar os que permaneceram neste mundo. Unir as duas áreas foi algo de extrema satisfação e conhecimento sobre as peculiaridades de cada idioma.

## REFERÊNCIAS BIBLIOGRÁFICAS

**ANTHONY, Laurence.** AntConc. Disponível em:

<http://www.laurenceanthony.net/software/antconc/> >. Acesso em 28/1/2020.

**ARAÚJO, Álvaro Placeres de.** Manual de Dactiloscopia – 02. ed. 1960. Ed. Coletânea Acácio Nogueira, Escola de Polícia de São Paulo.

**HOUAISS** Dicionário. Disponível em: <<http://houaiss.uol.com.br/>>. Acesso em 28/1/2020.

**KRIEGER, M. G.; FINATTO, M.J.B.** Introdução à Terminologia: teoria e prática. São Paulo, Contexto, 2004.

**MARÇALO, Maria João.** E-Dicionário de Termos Literários. Fraseologia. 2009. Disponível em: <<https://edtl.fcsh.unl.pt/encyclopedia/fraseologia>>. Acesso em 28/1/2020.

**SANTIAGO, Elizeu.** Criminalística Comentada. Millennium. 1. ed. Campinas, SP: 2014, módulo 18, 28, 45.

**SARDINHA, Tony Berber.** Linguística de corpus. 01. ed. São Paulo, Editora Manole, 2004.

**SEER Unisinos.** Revista de Administração e Contabilidade. What is a corpus and how to build it? Disponível em: <<http://www.revistas.unisinos.br>>. Acesso em 28/1/2020.

**TOCCHETTO, Domingos.** STUMVOLL, Vitor. Criminalística. Millennium. 6. ed. Campinas, SP: 2014. Cap. 3.

## APÊNDICE: GLOSSÁRIO

Termo em inglês	Cat. gram.	Fraseologias	Contexto	Observações	Termo em português	Cat. gram.	Fraseologias	Contexto	Observações
Analysis	noun	Forensic analysis	Postmortem injuries are a source of misinterpretations in <b>forensic analysis</b> and therefore are subject matter of taphonomic interest.	106	Análise	sub fem	Análise de solicitação	Após <b>análise de solicitação</b> e mediante a constante escassez de dados que contribuem para a identificação dos restos mortais a serem exumados, outros dados complementares são gradativamente e requeridos.	37
Blood	noun	Saliva and Blood	<b>Saliva and blood</b> are biological materials with high potential of infectious transmission in dental environments.	92	Sangue	sub masc	Manchas de sangue	A extração de DNA pós-morte tem sido feita também em <b>manchas de sangue</b> , pelos, sêmen, dentes e ossos.	15

Body	noun	human body	Human identification is one of the major fields of study and research in forensic science because it deals with the <b>human body</b> and aims at establishing human identity.	33	Corpo	sub masc	Corpo como um "boneco"	Ao olhar o <b>corpo como um "boneco"</b> ou como um "objeto", o perito banaliza a realidade com a qual tem que conviver como forma de torná-la suportável.	27
Case	noun	present case	These criteria were followed in the <b>present case</b> , in which dental positive identification was achieved before DNA exam.	103	Caso	sub masc	Conhecimento do caso	De forma a direcionar a execução de ações, visando à investigação do maior número possível de elementos para o completo <b>conhecimento do caso</b> , buscaram-se, na literatura, roteiros que possam ser seguidos durante as perícias exumatórias	67

Conditions	noun	The body conditions	After analyzing <b>the body conditions</b> , the samples and the decaying process, the mPMI was estimated using the average temperature of the week before the body discovery.	33	Condições	sub fem	condições de fazer levantamento	Com o término do módulo, os peritos novatos estariam, em tese, em <b>condições de fazer levantamento</b> de local de crime	10
Crime	noun	Crime Scene	The <b>crime scene</b> examination by forensic is obligatory in all criminal infractions that leave traces, under penalty of nullity of the criminal trial, and cannot be replaced even by the confession of the accused.	84	Crime	sub masc	Autoria do crime	Nos demais casos, os policiais priorizam os boletins de ocorrência que já trazem elementos de prova necessários para a conclusão de um inquérito (informações sobre a <b>autoria do crime</b> , com filmagens, depoimentos, testemunhas, registros diversos).	170

Death	noun	Natural death	Byard (2005) also reported three cases in which postmortem injuries caused by ants resembled perimortem lesions in different cases involving suicide, homicide, and <b>natural death</b> , raising doubts on the times of the injuries.	34	Morte	sub fem	Morte violenta	Em se tratando da formação de novos peritos criminais, a socialização secundária requer a preparação do neófito para lidar com a <b>morte violenta</b> .	200
Dental	adj	Dental identification	The teeth represent an excellent source of DNA because this biological material may provide the necessary relation for identification of an individual in case of failure of conventional methods for <b>dental identification</b>	138	Dental	adj	Fratura dental	As lesões do tipo equimose são mais incidentes em mulheres, ao passo que <b>fratura dental</b> , ferida contusa e perfurocontusa, em homens.	13



DNA	noun	DNA profiles	DNA analysis has been used successfully in a number of large accidents to associate body parts and for the purposes of identification, by comparing victims' <b>DNA profiles</b> with those of relatives.	191	DNA	sub masc	DNA de restos mortais	Com o aperfeiçoamento de técnicas, mais recentemente, foi possível recuperar o <b>DNA de restos mortais</b> humanos exumados, surgindo assim a procura pela investigação de paternidade post mortem.	38
Entomology	noun	Forensic entomology case	First medicolegal <b>forensic entomology case</b> of central Amazon: a suicide by hanging with incomplete suspension.	46	Entomologia	sub fem	Entomologia forense	A <b>entomologia forense</b> é mais comumente associada a investigações de morte, ajudando a determinar local e tempo dos incidentes de acordo com a fauna encontrada no cadáver.	10
Evidence	noun	Material evidence	The Forensic Science Service is the agency in charge of producing the <b>material evidence</b> at the public safety and criminal justice system, applying scientific knowledge and technology.	24	Evidência	sub fem	Uma evidência	Aquilo que o perito diz ser <b>uma evidência</b> será tomado como verdade pela autoridade.	13

Forensic	adj	Forensic case	The aim of this study was to show DPid as an important tool of potential application to solve cases with dental prosthesis, such as the <b>forensic case</b> in which skull, denture, and dental records were received for analysis.	595	Forense	adj	Identificação forense	Ligando cada crime a seu ambiente e cada cadáver a seu cenário, as fotografias de <b>identificação forense</b> acentuam ainda mais o processo.	30
Fragments	noun	Body fragments	Forensic Dentistry is the specialty with the goal of investigating psychological, physical, chemical and biological phenomena that can reach human beings (alive, dead or <b>body fragments</b> ), comprehending aspects of human identification	25	Fragmento	sub masc	Fragmentos de DNA	Os <b>fragmentos de DNA</b> produzidos foram analisados por eletroforese em gel desnaturante de posliacrilamida , seguida de coloração com nitrato de prata.	13

Identification	noun	Human identification	Studies on molecular biology applied to <b>human identification</b> will probably further enhance DNA extraction with less material available and under increasingly adverse conditions.	170	Identificação	sub fem	Dificuldade na identificação	Uma das maiores dificuldades relacionava-se à coleta dos restos mortais solicitados pelo geneticista, quando haviam múltiplas inumações numa mesma sepultura, o que acarretava maior <b>dificuldade na identificação.</b>	71
Investigation	noun	Investigation and processing of the crime scene	During the <b>investigation and processing of the crime</b> scene, lies with the forensic expert the judgment of whether a trace is relevant to the casework.	48	Investigação	sub fem	Manuais de investigação	No que se refere especificament e às polícias, foram criados <b>manuais de investigação</b> , bem como introduzidos novos procedimentos operacionais para otimizar o desempenho dos investigadores	103

Justice	noun	Criminal justice	The Forensic Science Service is the agency in charge of producing the material evidence at the public safety and <b>criminal justice</b> system, applying scientific knowledge and technology	14	Justiça	sub fem	Para que a justiça seja posta em prática	Quando não há uma justificativa plausível (se é que para a violência ela existe), a população reage e exige o desvendamento da tragédia em tempo exíguo como forma de obter explicação e também <b>para que a justiça seja posta em prática.</b>	41
Luminol	noun	The luminol solution	It was concluded that <b>the luminol solution</b> is effective in detecting blood tissue in endodontic files as well as in validating the cleaning/sterilization process.	64	Luminol	sub masc	Substâncias presentes no Luminol	A reação química ocorre quando as <b>substâncias presentes no Luminol</b> entram em contato com o sangue, mais precisamente com as partículas de ferro existentes na hemoglobina.	7
Material	noun	Biological material	Blood is a <b>biological material</b> with high potential of infectious transmission in dental environments, including herpes simplex, hepatitis and AIDS.	51	Material	sub masc	Material a ser colhido	O geneticista fazia previamente a escolha do <b>material a ser colhido</b> para pesquisa de DNA, e o material obtido era identificado e acondicionado em caixa para transporte.	27

Method	noun	Luminol method	The <b>luminol method</b> is non-toxic and easy to use, being up to 20 times more sensitive to hemoglobin than other blood detection tests, including those using phenolphthalein (Kastle-Meyer test), leucomalachite green, benzidine reagent and fluorescein.	72	Método	sub masc	Método científico	Os peritos devem adotar os princípios do <b>método científico</b> , baseados nos conhecimentos técnico-científicos (KUSER, 2009, p. 2).	69
Police	noun	the police and justice	Therefore, the value of the service lies in its ability to help <b>the police and justice</b> to elucidate complex crimes, using science, respecting human rights.	52	Polícia	sub fem	Polícia Federal	Os dados coletados, a síntese situacional das competências dos gestores da criminalística da <b>Polícia Federal</b> e a análise de resultados respeitaram os limites de complexidade, espaço e tempo do estudo.	109
Postmortem	noun	The separation of post-mortem from perimortem	Even experienced crime scene professionals may have questions on <b>the separation of postmortem from perimortem</b> injuries.	47	Pós-morte	sub fem	Produzida em vida ou pós-morte	Podem ser levantadas hipóteses médico-legais, como distância e trajeto de disparo ou se uma lesão foi <b>produzida em vida ou pós-morte</b> .	7

Research	noun	Area of research	A potential <b>area of research</b> in dental science involving human identification is the study of the bite marks.	68	Pesquisa	sub fem	O universo da pesquisa é constituído pela criminalística	<b>O universo da pesquisa é constituído pela criminalística</b> do DPF, ou seja, os peritos criminais federais responsáveis pela realização das perícias de natureza criminal requeridas na apuração de crimes da competência da União.	120
Results	noun	The results provided evidence	<b>The results provided evidence</b> that the storage temperature of determined materials does not exert a significant influence.	52	Resultados	sub masc	Os resultados apontam ênfase em determinados aspectos	<b>Os resultados apontam ênfase em determinados aspectos</b> presentes na atuação dos peritos, a saber: responsabilidade, dedicação, impossibilidade e de cometer erros.	54
Scene	noun	Crime scene	The expert's judgment is based on the observation of all the elements in a <b>crime scene</b> , aiming to establish its dynamics under the light of the hypothetico-deductive method	69	Cena	sub fem	Cena do crime	O laudo permite vincular o autor à <b>cena do crime</b> (ou inocentar alguém erroneamente acusado) ou elucidá-lo.	23

Science	noun	Science of identification	The <b>science of identification</b> of criminal cases was based only on serological analyses of protein polymorphism, blood groups and some genetic markers.	141	Ciência	sub fem	Ênfase na ciência	A morte violenta será desconstruída pela <b>ênfase na ciência</b> , na técnica.	21
---------	------	---------------------------	---	-----	---------	---------	-------------------	---	----