



O léxico da formação continuada de professores dos anos iniciais: tecendo caminhos para o reconhecimento de uma terminologia

Leicijane da Silva Barros^{1*} e Ana Claudia Castiglioni²

Secretaria Municipal da Educação, Av. Bernardo Sayão, 499, 77818-340, Araguaína, Tocantins, Brasil. ²Universidade Federal do Norte do Tocantins, Araguaína, Tocantins, Brasil. *Autor para correspondência. E-mail: leicijaneb@gmail.com

RESUMO. O presente artigo refere-se a um recorte de uma pesquisa desenvolvida em nível de doutoramento, em que se buscou identificar e analisar os termos da Formação Continuada de Professores dos Anos Iniciais do Ensino Fundamental. O objetivo deste trabalho é apresentar os parâmetros considerados no processo de reconhecimento do léxico especializado do domínio estudado e a sua organização por meio de sistemas de conceitos. O aporte teórico fundamentou-se nos pressupostos das Ciências do Léxico (Cabré, 1993, 1999; Biderman, 2001, 2006; Barros, 2004; Krieger & Finatto, 2020) e da Linguística de *Corpus* (Berber Sardinha, 2000, 2004). Para o processamento do corpus, extração e análise dos termos foram empregadas duas ferramentas computacionais: *BootCaT* e *AntConc*. Com a contribuição de especialistas da área, foram validados 2.851 termos para o domínio da Formação Continuada de Professores dos Anos Iniciais, os quais foram distribuídos em três subdomínios para facilitar o acesso de seus consultantes, a saber: conhecimentos linguísticos, conhecimentos matemáticos e conhecimentos didático-educacionais. A relevância dos dados apresentados neste trabalho reside na identificação da terminologia da formação continuada de professores dos anos iniciais, estabelecendo-se possíveis fronteiras com outras áreas similares, o que esperamos que possa contribuir para o desenvolvimento de estudos e objetos terminográficos futuros.

Palavras-chave: léxico; reconhecimento terminológico; formação continuada.

The lexicon of continuing education for early grade teachers: weaving paths towards the recognition of a terminology

ABSTRACT. The present article refers to a section of a doctoral research project that sought to identify and analyze the terms of the Continuing Education of Teachers of the Early Years of Primary School. The purpose of this work is to present the parameters considered in the process of recognizing the specialized lexicon of the domain studied and your organization by means of concepts systems. The theoretical support was based on the assumptions of Lexicon Sciences (Cabré, 1993, 1999; Biderman, 2001, 2006; Barros, 2004; Krieger & Finatto, 2020) and Corpus Linguistics (Berber Sardinha, 2000, 2004). For corpus processing, extraction and analysis of the terms two computational tools were employed: *BootCaT* and *AntConc*. With the contribution of specialists in the field, 2,851 terms were validated for the domain of Continuing Education for Teachers of Early Years, which were distributed into three subdomains to facilitate the access of its consultants, namely: linguistic knowledge, mathematical knowledge and didactic-educational knowledge. The relevance of the data presented in this work lies in the identification of the terminology of continuing education for early grade teachers, establishing possible boundaries with other similar areas, which we hope will contribute to the development of future studies and terminographic objects.

Keywords: lexicon; terminology recognition; continuing education.

Received on April 30, 2023.
Accepted on December 8, 2023.

Introdução

A formação docente tem sido tema constante nas discussões acadêmicas nas últimas décadas. No que se refere à dos profissionais que atuam nos Anos Iniciais do Ensino Fundamental, principal público deste estudo, os debates têm se intensificado, especialmente, no que se tange à formação inicial, considerada, por muitos

estudiosos da área da educação, como insuficiente e insatisfatória diante da complexidade do fazer docente nesta etapa escolar.

Ante o exposto, torna-se cada vez mais evidente a necessidade urgente de reestruturação curricular dos cursos preparatórios para professores, bem como repensar políticas públicas educacionais mais eficazes, voltadas à formação continuada dos educadores, como mecanismos capazes de dirimir lacunas deixadas pela formação inicial.

Nesse sentido é que, motivadas pela escuta dos profissionais que atuam na primeira etapa do ensino fundamental durante os treze anos dedicados à realização de encontros formativos na rede de ensino em que uma das autoras atua e pela constatação de que não há obras na área da Terminologia (por meio de consulta à Biblioteca Digital Brasileira de Teses e Dissertações - BDTD e a outros sites e bancos de pesquisa acadêmica) que apresentem os termos do domínio especializado da Formação Continuada de Professores dos Anos Iniciais do Ensino Fundamental, propomo-nos, neste trabalho, a apresentar o conjunto terminológico que estrutura esse domínio de especialidade, por meio da demonstração de árvores do domínio e dos parâmetros de reconhecimento dos termos da área, o que esperamos que possa contribuir para o desenvolvimento de estudos e objetos terminográficos futuros.

Portanto, o estudo exposto se insere na área das Ciências do Léxico, em especial, na da Terminologia. Na próxima seção, buscamos realizar uma breve exposição acerca da formação inicial e continuada de professores, a fim de contextualizarmos o objeto de pesquisa apresentado neste artigo – a terminologia da formação continuada dos professores dos anos iniciais.

Formação inicial e continuada de professores

O ser humano vive em permanente processo de aprendizagem. Paulo Freire (1989) tão bem pontua que todos nós aprendemos sempre, pois, embora saibamos alguma coisa, sempre ignoramos outras. No tocante à formação docente, a necessidade dessa consciência de que sempre temos algo a aprender é de extrema relevância, pois a aquisição do conhecimento não se finda ao término da graduação, encerra-se ali apenas uma etapa de um processo que se perpetuará por toda a vida.

Nesse contexto, a formação continuada de professores, como o próprio nome remete, deve ser contínua, regulada por um processo constante de ação-reflexão-ação. Negrine (1998) pontua que a formação inicial consiste na habilitação do indivíduo para atuar em determinada área do conhecimento, podendo esta ser adquirida com o término do curso de licenciatura ou bacharelado, constituindo-se, assim, como um pré-requisito para a atuação profissional e, no âmbito educacional, simboliza o princípio de uma trajetória balizada pela reflexão constante entre teoria e prática.

O art. 62, da Lei de Diretrizes e Bases da Educação Nacional (LDB), nº 9.394/1996, aponta a formação docente (inicial) necessária para atuação na educação básica:

A formação de docentes para atuar na educação básica far-se-á em nível superior, em curso de licenciatura plena, admitida, como formação mínima para o exercício do magistério na educação infantil e nos cinco primeiros anos do ensino fundamental, a oferecida em nível médio, na modalidade normal (Brasil, 2019, p. 42).

Em seu art. 61, parágrafo único e respectivos incisos, a LDB (Brasil, 2019) determina os fundamentos da formação de profissionais da educação, a qual seria respaldada por uma sólida formação básica (formação inicial), orientada por conhecimentos científicos pertinentes ao fazer docente, por saberes advindos de uma associação teórico-prática, obtidos em estágios ou cursos de qualificação em serviço (após a graduação), além das experiências profissionais adquiridas ao longo da vida.

Negrine (1998) deslinda que, quanto à formação continuada, esta deve, como o próprio nome diz, ser compreendida como um processo contínuo, englobando todas as aprendizagens oriundas de atualização profissional permanente, ocorrida após a formação inicial, agregadas às experiências vivenciadas, sendo estas relacionadas ou não a cursos de atualização em nível *lato* ou *stricto sensu* (pós-graduação).

De acordo com a LDB, art. 44, inciso III, são compreendidos como pós-graduação os “[...] programas de mestrado e doutorado, cursos de especialização, aperfeiçoamento e outros, abertos a candidatos diplomados em cursos de graduação e que atendam às exigências das instituições de ensino” (Brasil, 2019, p. 33).

Em consonância com Negrine (1998), Duarte (2017, p. 1) explica que a formação continuada de professores consiste, de igual modo, em um processo constante e “[...] permanente de aperfeiçoamento dos saberes necessários à atividade docente, realizado ao longo da vida profissional, com o objetivo de assegurar uma ação docente efetiva que promova aprendizagens significativas”. Souza e Lima (2017) complementam afirmando que

A dinamicidade das informações requer que o professor se atualize constantemente para saber lidar melhor com os conhecimentos construídos. Visto que a formação inicial não garante a preparação suficiente para exercer a sua profissão, a formação continuada torna-se um meio essencial para as exigências atuais (Souza & Lima, 2017, p. 1).

Quanto à natureza, são considerados cursos de formação continuada ou educação continuada, todos aqueles realizados após a graduação, que vão desde cursos de aperfeiçoamento profissional (ofertados em serviço), até os de pós-graduação *lato sensu* (especialização) e *stricto sensu* (mestrado e doutorado), mas eles não se confundem, pois apresentam propostas e exigências diferentes entre si.

Insta esclarecermos que os termos ‘formação continuada’ e ‘pós-graduação’ não são semelhantes nem sinônimos, mesmo se tratando de processos formativos posteriores à graduação. A eventual confusão em relação ao conceito dos dois é causada porque os cursos de pós-graduação também seriam cursos de formação continuada, pois buscam a qualificação do profissional após a graduação. Todavia, nem todos os cursos de formação continuada poderiam ser enquadrados como cursos de pós-graduação.

A formação continuada diz respeito a um processo permanente, contínuo, que se dá após a formação inicial, com o objetivo de preencher as lacunas deixadas por esta e atualizar o profissional quanto às inovações de sua área de atuação; já a pós-graduação consiste em especializações de longa duração, também subsequentes à graduação, de caráter acadêmico.

Os cursos de pós-graduação, *lato* e *stricto sensu* têm validade acadêmica, devendo, pois, ser ofertados por instituições de ensino superior certificadas pelo Ministério da Educação (MEC), com carga horária mínima de 360 horas e exigência de apresentação de monografia, para pós-graduação *lato sensu* ou dissertações e teses, no caso de *stricto sensu*, modalidade que envolve também o desenvolvimento de pesquisa.

Já os cursos de capacitação no âmbito das formações continuadas em serviço, por sua vez, são cursos de curta a média duração (carga horária inferior a 360 horas), não exigem diploma de graduação para o ingresso, tampouco pesquisa e trabalhos acadêmicos como dissertação e tese. Para exemplificar, na Tabela 1 a seguir demonstramos os principais programas voltados para a formação de professores alfabetizadores implantados no Brasil nas últimas décadas e cujos documentos orientadores serviram de *corpus* para a pesquisa que, em parte, apresentamos neste artigo:

Tabela 1. Programas voltados à formação de professores alfabetizadores no Brasil.

Programa	Ano de lançamento
PCN em Ação – Alfabetização	1999
Gestar I – Programa Gestão da Aprendizagem Escolar	2001
PROFA – Programa de Formação de Professores Alfabetizadores	2001
PRALER – Programa de Apoio à Leitura e a Escrita	2003
Pró-Letramento – Programa de Formação Continuada de Professores dos Anos/Séries Iniciais do Ensino Fundamental	2005
PNAIC – Programa Pacto Nacional pela Alfabetização na Idade Certa	2012
Programa Tempo de Aprender	2020

Fonte: Elaborado por Barros, 2022 (adaptado de Brito e Viédes, 2015).

Para fins do estudo que foi realizado, concebemos apenas os cursos de formação continuada de professores dos Anos Iniciais do Ensino Fundamental, desenvolvidos (em serviço) na última década, no âmbito da rede educacional pesquisada.

Estudos do léxico: terminologia em foco

No campo da ciência, há algumas disciplinas tradicionais que se ocupam do estudo do léxico, tais como a Lexicologia, a Lexicografia e a Terminologia, sendo esta última, dado o recorte deste trabalho, a que nos interessa.

Biderman (2001, p. 19) explica que a Terminologia “[...] se ocupa de um subconjunto do léxico de uma língua, a saber, cada área específica do conhecimento humano”. Assim, temos, por um lado, a Lexicologia que estuda o componente lexical geral das línguas e, por outro, a Terminologia, que se dedica aos termos de uma área especializada.

Os termos são instituídos, conforme apregoa Krieger (1998, p. 1), “[...] por força de especificidades conceituais dos diferentes campos do conhecimento”. A este respeito, Benveniste (1989) assevera que a constituição de uma terminologia própria

[...] marca, em toda ciência, o advento ou o desenvolvimento de uma conceitualização nova, assinalando, assim, um momento decisivo de sua história. Poder-se-ia mesmo dizer que a história particular de uma ciência se resume na de seus termos específicos (Benveniste, 1989, p. 252).

No que concerne à natureza dessa área de estudos, Cabré explica que a Terminologia é um campo inter e transdisciplinar:

[...] é interdisciplinar porque não pode descrever seu objeto sem o auxílio de diversas disciplinas (ciências da linguagem, ciências cognitivas e ciências da comunicação). É transdisciplinar na medida em que participa de todas as disciplinas especializadas, uma vez que todas possuem e utilizam terminologia para representar seus conhecimentos (Cabré, 1999, p. 97, tradução nossa).¹

Numa vertente descritiva, baseada no comportamento dos termos nos textos especializados, compreendendo-os como unidades linguístico-comunicacionais, segundo uma orientação pragmática da linguagem, que considera sua complexidade e dinamismo, surge a Teoria Comunicativa da Terminologia (TCT), a partir dos estudos da pesquisadora Maria Tereza Cabré.

Considerar a natureza linguística do termo é admitir que ele está sujeito a toda complexidade e grandiosidade que isso implica, é entender o lugar social e cultural das línguas de especialidade, é reconhecer que os termos só existem dentro de um contexto comunicativo e estão a serviço da transmissão do conhecimento especializado.

Nesse sentido, conforme assinala Cabré (1999) na proposta da TCT, os termos não pertencem a um único domínio, mas são usados nele, pressupondo-se, dessa maneira, que podem, também, ser utilizados em outros domínios. É o caso do termo ‘vírus’, que no âmbito da Medicina, corresponde a um ser pequeno, parasita, capaz de infectar organismos vivos e transmitir doenças; já na esfera da informática, o *vírus* é um programa desenvolvido com o objetivo de causar sérios problemas ao usuário de um computador ou empresa.

Assim, conforme elucida Cabré (1999, p. 123, tradução nossa), “[...] os termos são unidades lexicais ativadas singularmente por suas condições pragmáticas de adequação a um tipo de comunicação [...]”², configurando-se como termos se forem definidos e utilizados em textos especializados.

Segundo a ISO 1087 (International Organization for Standardization [ISO], 1990, p. 5, tradução nossa), “[...] um termo pode consistir em uma ou mais palavras (simples ou complexo) ou mesmo conter símbolos [...]”³, sendo que o termo simples é aquele que “[...] consiste em apenas um radical, com ou sem afixos [...]”, já o complexo, possui “[...] dois ou mais radicais com ou sem outros elementos do termo” (International Organization for Standardization [ISO], 1990, p. 7, tradução nossa)⁴.

Para exemplificar, apresentamos alguns termos extraídos de nosso *corpus*, tais como: ‘alfabetização’, ‘avaliação’, ‘currículo’ e ‘descritores’ (simples); e, ‘ambiente alfabetizador’, ‘direitos de aprendizagem’, ‘eixos temáticos’ e ‘ensino híbrido’ (complexos).

Além dos termos simples e complexos, Barros (2004) acrescenta os termos compostos, também constituídos por unidades lexicais formadas por dois ou mais radicais. No entanto, segundo a autora, distinguem-se dos termos complexos

[...] pelo alto grau de lexicalização e pelo conjunto de morfemas lexicais e/ou gramaticais que os constitui, em situação de não-autonomia representada graficamente pela utilização do hífen, como em mão-de-obra, pé-de-cabra, pá-de-cavalo. Cumpre ressaltar que consideramos as unidades lexicais compostas por aglutinação (como fidalgo, embora etc.) e pela justaposição sem hífen de dois ou mais radicais como termos simples (Barros, 2004, p. 100).

Cabré e Estopà (2005) complementam que as unidades sintagmáticas podem corresponder a três estruturas:

a) as formadas por um núcleo nominal complementado por um adjetivo ou sintagma adjetivo; b) as formadas por um núcleo nominal complementado por um sintagma preposicional; c) as formadas por um núcleo nominal complementado por um sintagma nominal (Cabré & Estopà, 2005, p. 14, tradução nossa)⁵.

Para melhor ilustrar as estruturas elencadas, expomos, a seguir, alguns exemplos, tomando por base o *corpus* da pesquisa apresentada neste artigo. No que tange ao primeiro caso (núcleo nominal + adjetivo ou sintagma

¹ “Es interdisciplinar porque no puede describir su objeto sin el concurso de disciplinas diversas (las ciencias del lenguaje, las ciencias cognitivas y las ciencias de la comunicación). Es transdisciplinar por cuanto participa de todas las materias especializadas, ya que todas ellas poseen y usan terminología para representar sus conocimientos”.

² “Los términos son unidades léxicas, activada singularmente por sus condiciones pragmáticas de adecuación a un tipo de comunicación”.

³ “A term may consist of one or more words (simple term, or complex term) or even contain symbols”.

⁴ “Simple term: Term consisting of only one stem with or without affixes. Complex term: Term consisting of two or more stems with or without other term element”.

⁵ “a) las formadas por un núcleo nominal complementado por un adjetivo o sintagma adjetivo b) las formadas por un núcleo nominal complementado por un sintagma preposicional c) las formadas por un núcleo nominal complementado por un sintagma nominal”.

adjetivo), podemos citar: 'leitura deleite', 'projeto didático', 'sequência didática'. Quanto ao segundo tipo (núcleo nominal + sintagma preposicional), temos: 'roda de leitura', 'orientador de estudos', 'direitos de aprendizagem'. E, por último (núcleo nominal + sintagma nominal), temos: 'Prova Brasil', 'livre-docência', 'hora-aula'.

É imperioso ressaltar que, dada a complexidade do assunto, identificar o termo ou sintagmas terminológicos dentro de um contexto especializado não é uma tarefa fácil, ao contrário, representa um dos maiores desafios para quem trabalha com estudos terminológicos, pois as fronteiras são muito tênues no que se refere à distinção entre termos e palavras da língua comum, bem como entre os termos de áreas próximas.

Concluída esta explanação, passamos a discorrer acerca dos procedimentos metodológicos que serviram de norte para o estabelecimento da terminologia que apresentamos neste trabalho.

Tecendo caminhos para o reconhecimento de uma terminologia

Para chegarmos ao reconhecimento da terminologia do domínio da formação de professores dos anos iniciais que apresentamos neste artigo, seguimos algumas etapas metodológicas: a primeira consistiu na definição da temática, do público-alvo e dos objetivos a serem alcançados; na segunda etapa, realizou-se um levantamento bibliográfico acerca das Ciências do Léxico, com ênfase na Terminologia; na terceira fase foram levantados os textos/documentos e/ou materiais da área de formação continuada de professores dos anos iniciais que constituiriam o *corpus*; a quarta fase consistiu na escolha das ferramentas computacionais a serem utilizadas no tratamento do *corpus*; na quinta fase, com o auxílio do programa *BootCaT*⁶, foi realizado o processamento do *corpus*; a sexta fase consistiu na extração, análise e seleção do conjunto terminológico da área estudada, por meio do programa *AntConc*⁷ e adoção de critérios de reconhecimento terminológico; a sétima etapa correspondeu à validação dos termos do domínio pelas especialistas das áreas do conhecimento envolvidas.

Delimitação do domínio estudado

Em função da experiência profissional de uma das autoras deste trabalho, que há 13 anos tem atuado com formação continuada de professores em uma rede municipal de ensino, além da motivação resultante da escuta dos educadores em contextos formativos, em sua maioria pedagogos, os quais demonstraram, por meio de diversos relatos, a dificuldade na compreensão de termos presentes nos documentos orientadores dos programas de formação implantados na rede e nas matrizes de referência das avaliações externas, é que decidimos pela delimitação do domínio especializado da Formação Continuada de Professores dos Anos Iniciais do Ensino Fundamental.

Vale destacar que, no âmbito dos anos iniciais, os programas de formação continuada lançados pelo MEC têm foco no desenvolvimento das habilidades que são essenciais para essa etapa educacional, quais sejam: a leitura, escrita e alfabetização matemática. Contudo, além dos termos relacionados à Linguagem e Matemática, apresentamos também unidades lexicais voltadas aos conhecimentos didático-educacionais gerais, indispensáveis aos profissionais que atuam nessa fase escolar, cuja natureza do ensino é, essencialmente, interdisciplinar.

Definição dos critérios de seleção dos textos e constituição do *corpus*

Após a delimitação do objeto de pesquisa, outra etapa importante é a definição dos critérios de seleção dos textos que constituem o *corpus* do trabalho, especialmente, quando desses textos resulta(rá) um produto lexicográfico/terminográfico.

Nesse sentido, os critérios de seleção dos textos implicaram diretamente na qualidade e credibilidade do produto gerado pelo *corpus*. A esse respeito, Cabré (1993) alerta que a definição do *corpus* deve ser realizada de forma muito cautelosa, uma vez que servirá de base para a detecção das unidades terminológicas, influenciando diretamente nos resultados obtidos.

Sob essa perspectiva, é que a autora orienta que uma série de condições sejam seguidas no processo de seleção do *corpus* da pesquisa terminológica, sem as quais não seria possível garantir confiabilidade nos resultados obtidos.

Tendo em vista que este trabalho se pauta nos pressupostos teóricos de Cabré, valemo-nos das condições elencadas pela pesquisadora para definirmos os critérios de seleção de nosso *corpus*, quais sejam:

⁶ *BootCaT* é uma interface que serve como guia na criação de corpora simples a partir da *web*. O programa automatiza o processo de encontrar textos de referência na *web* e reuni-los em um único *corpus*. Foi lançado em 2004, por Marco Baroni (Universidade de Trento) e Silvia Bernardini (Universidade de Bolonha).

⁷ *AntConc* é um *software* de análise textual e linguística de *corpus*. Foi desenvolvido por Laurence Anthony, Professor na Faculdade de Ciência e Engenharia da Universidade Waseda, no Japão. É considerado um dos programas mais acessíveis e úteis para realização desse tipo de análise.

- deve ser relevante, ou seja, representativo do campo em que atua e, se possível, escrito por um autor de qualidade;
- deve ser completo, portanto deve incluir todos os aspectos relacionados ao assunto do trabalho;
- deve ser atual, para que a lista de termos que podem ser extraídos reflita a realidade linguística atual do campo em questão;
- deve ser original, ou seja, deve ser expresso na língua em que é utilizado (Cabr , 1993, p. 298, tradu  o nossa)⁸.

Quanto ao primeiro crit rio, o *corpus* que foi selecionado trouxe, em sua maioria, materiais oficiais dos programas de forma  o continuada implantados pelo MEC nos  ltimos anos, produzidos em parceria com Institui  es de Ensino Superior renomadas, em fun   o da relev ncia dos trabalhos desenvolvidos por seus pesquisadores.

No caso da produ   o dos materiais formativos, os autores s o estudiosos com ampla experi ncia em pesquisas sobre as tem ticas neles abordados. Os demais textos corresponderam a artigos, disserta  es e teses sobre a  rea, extra dos de sites de pesquisa acad mica, cuja credibilidade   reconhecida.

Assim, al m da qualidade dos autores dos materiais e textos produzidos, o *corpus* foi expressivo e relevante para  rea, pois envolve um consider vel n mero de pesquisadores experientes, que t m trazido em seus materiais atualiza  es te rico-metodol gicas no que se refere aos temas voltados   forma  o continuada dos professores dos Anos Iniciais do Ensino Fundamental.

No que se refere ao segundo crit rio, o *corpus* foi composto por um total de 157 textos/materiais, com os quais buscamos abarcar os aspectos mais diversificados e representativos do campo em quest o, o que garante, em tese, uma representatividade dos mais diversos aspectos do dom nio. O programa *AntConc* gerou, por ordem de frequ ncia, as unidades lexicais candidatas a termos do dom nio de especialidade, totalizando 89.205 unidades lexicais diferentes.

Em rela  o ao terceiro crit rio, considerando que a pesquisa buscou analisar os documentos dos programas de forma  o de professores da primeira fase do Ensino Fundamental, implantados nos  ltimos anos, e que a data de publica  o desses materiais compreende o per odo entre 2005 e 2021,   poss vel julgarmos que o *corpus*   atual. Por conseguinte, os candidatos a termos dele retirados refletir o as mudan as e atualiza  es te rico-metodol gicas do campo em foco. Por  ltimo, o *corpus* pode ser considerado original, uma vez que todos os materiais/textos selecionados foram escritos em L ngua Portuguesa.

No que se refere   delimita  o do quantitativo de textos que comp em o *corpus*, de modo que este possa ser representativo para uma dada  rea, vale ressaltar que, segundo aponta Berber Sardinha (2000, p. 342), “[...] na sua ess ncia, um *corpus*, seja de que tipo for,   tido como representativo da linguagem, de um idioma, ou de uma variedade dele”. Sinclair (1991), por sua vez, avalia que o crit rio da representatividade est  diretamente relacionado ao tamanho do *corpus*. Nesse sentido, quanto maior for sua extens o, maior ser  sua representatividade.

A despeito disso, Berber Sardinha (2000) pondera que, embora a extens o seja um crit rio primordial para a determina  o da representatividade de um *corpus*, a quantidade m nima de dados necess rios para sua forma  o ainda n o foi estimada, sendo, pois, um crit rio subjetivo, o que dificulta sua mensura  o.

Assim, um dos grandes entraves para os estudiosos da  rea   a aus ncia de pesquisas que busquem estabelecer crit rios que definam a extens o de um *corpus*. Berber Sardinha (2000) pontua que o tamanho de um *corpus* comporta tr s dimens es: n mero de palavras, de textos e de g neros, registros ou tipos textuais.

Postas todas as informa  es supramencionadas acerca da representatividade de um *corpus* e analisando a extens o do *corpus* que formamos, o n mero de textos (157), n mero de palavras (89.205) e variedades de registros (materiais formativos do MEC, artigos, teses e disserta  es) que o comp em, ponderamos que ele pode ser, em sua ess ncia, considerado expressivo, significativo e representativo para o dom nio estudado.

Para constitui  o do *corpus*, seguimos, assim, os postulados de Cabr  (1999) e da Lingu stica de *Corpus*, que elencam quatro pr -requisitos essenciais a ser considerados no processo de compila  o de um *corpus*: a) ‘autenticidade’ – deve ser composto de textos aut nticos, em linguagem natural; b) ‘naturalidade’ – os textos devem ter sido produzidos por falantes nativos; c) ‘criteriosidade’ – o conte do do *corpus* deve ser escolhido criteriosamente, seguindo, acima de tudo, os princ pios da autenticidade e naturalidade; e, d) ‘representatividade’ – os textos devem ser representativos de uma l ngua, variedade lingu stica ou dom nio em an lise (Berber Sardinha, 2004).

⁸ “- debe ser pertinente, esto es, representativo del campo en que se trabaja y, a ser posible, redactado por un autor de calidad;

- debe ser completo, por lo que debe incluir todos los aspectos relacionados con el tema de trabajo;

- debe ser actual, para que la lista de t rminos que pueda extraerse refleje la realidad ling stica presente del  mbito en cuesti n;

- debe ser original, es decir, debe estar expresado en la lengua en que se trabaja”.

Segundo os requisitos expostos, alguns já discutidos nesta seção, pode-se dizer que o *corpus* desta pesquisa é ‘autêntico’ e ‘original’ (Cabr ,1993), na medida em que todos os textos foram escritos por falantes nativos e em L ngua Portuguesa. Ademais, os textos foram ‘criteriosamente’ escolhidos, tendo em vista os objetivos da pesquisa e uma maior ‘representatividade’ do dom nio em an lise.

Vale ressaltar que os princ pios da autenticidade, naturalidade e representatividade do dom nio foram os crit rios mais relevantes para a determina  o do *corpus* da pesquisa.

Apresentamos, a seguir, o processo de organiza  o e processamento do *corpus* estudado.

Organiza  o e processamento do *corpus*

Na fase de organiza  o e processamento do *corpus*, Berber Sardinha (2004) alerta que tr s aspectos essenciais devem ser considerados: a ‘formata  o’ – os textos devem ser leg veis para computador; a ‘extens o’ – o *corpus* deve ser vasto, a fim de que seja, de fato, representativo; e o ‘prop sito’ – a compila  o do *corpus* deve atender aos objetivos da investiga  o ling stica, a partir de crit rios de composi  o dos *corpora* predefinidos.

Diante do exposto,   que, segundo os pressupostos da Lingu stica de *Corpus* e, baseados em uma colet nea textual pr via e criteriosamente selecionada, representativa do dom nio especializado em estudo, organizamos/selecionamos o *corpus*, partindo, inicialmente, dos documentos oficiais do MEC, voltados para o desenvolvimento das pol ticas de forma  o continuada nos Anos Iniciais do Ensino Fundamental, seguidos pela busca de textos sobre a tem tica na Biblioteca Digital Brasileira de Teses e Disserta  es (BDTD), no *Google Acad mico* e *SciELO*.

No que concerne ao tratamento/processamento do *corpus*, bem como   extra  o dos candidatos a termos e an lise dos dados, utilizamos dois programas computacionais: o *BootCaT* e o *AntConc*.

Vale ressaltar que, embora o *software BootCaT* tenha se mostrado extremamente pr tico e  til no que se refere   organiza  o e tratamento do *corpus*, observamos que ainda precisa passar por algumas atualiza  es de modo a melhor atender demandas de pesquisa, uma vez que no processo de convers o para o formato TXT, a quantidade de arquivos convertidos   sempre menor do que a de URLs submetidas   convers o. Nesses casos, foi necess rio conferir, individualmente, os textos n o convertidos e utilizar programas *on-line* gratuitos pr prios para convers o em TXT dos materiais n o convertidos pelo *BootCaT*.

Na se  o seguinte, apontamos os procedimentos de extra  o e categoriza  o das unidades lexicais candidatas a termos.

Extra  o e categoriza  o dos candidatos a termos

Uma vez constitu do o *corpus*, procedemos com a extra  o semiautom tica dos candidatos a termos do dom nio da Forma  o Continuada de Professores dos Anos Iniciais do Ensino Fundamental, figurando, neste ponto, o programa *AntConc*, ferramenta computacional por n s escolhida para realizar esse processo.

O *software* oferece uma variedade de funcionalidades. Por exemplo, produz lista de palavras-chave, localiza-as em determinados contextos, apresenta  ndice de frequ ncia de palavras, dentre outras. Todavia, uma de suas maiores vantagens,   a capacidade de analisar grandes quantidades de documentos de uma  nica vez, possibilitando ao pesquisador, op  es avan adas de manuseio das informa  es de seu *corpus*.

Em fun  o do tamanho do *corpus*, o processo de an lise dos candidatos a termos foi relativamente demorado. Durou cerca de tr s meses e foi realizado em etapas. Ap s a primeira an lise, reduzimos de 89.205 para 5.092 unidades lexicais e optamos por utilizar o *Excel* por facilitar a contagem dos futuros termos.

  relevante destacar que nessa primeira etapa, a redu  o foi grande devido   quantidade de termos em ingl s extra dos dos *abstracts* dos artigos, disserta  es e teses, das palavras incompletas, das repetidas na forma singular e plural, dos caracteres soltos n o inclu dos nos corpos de exclus o,  l m das palavras n o representativas para a  rea.

A segunda an lise resultou em uma nova redu  o, totalizando 3.498 candidatos a termos e, antes de com armos a pr xima fase, decidimos iniciar o processo de categoriza  o dos candidatos a termos, distribuindo-os nos tr s subdom nios inicialmente elencados: conhecimentos ling sticos, matem ticos e did tico-educacionais, conforme apresentado, a seguir, na proposta de sistema de conceitos, organizada por meio de  rvores de dom nio⁹.

Na concep  o da proposta do nosso sistema de conceitos, partimos das rela  es hier rquicas entre os n veis de abstra  o, seguindo um diagrama composto por termos-chave da  rea estudada, como em um

⁹ Diagrama hier rquico composto por termos-chave de uma especialidade, semelhante a um organograma (Krieger & Finatto, 2020).

organograma, considerando o foco das formações dos anos iniciais (Linguagem/Matemática) e outros saberes gerais ligados à Educação.

Assim, optamos por não propor uma organização do conjunto terminológico apenas pelo critério da ordem alfabética, mas também por áreas temáticas, o que não significa que buscamos segmentar o conhecimento, tampouco ‘dividi-lo em caixinhas’, mas organizá-lo de modo a facilitar a compreensão e acesso dos futuros consulentes.

Vale destacar que a proposta de sistema de conceitos, exibida na Figura 1, correspondeu a uma propositura inicial, pois à medida em que a pesquisa foi avançando, outras subcategorias foram criadas, uma vez que a divisão considerando apenas os três subdomínios expostos, por si só, não foi capaz de abarcar toda a variedade terminológica.

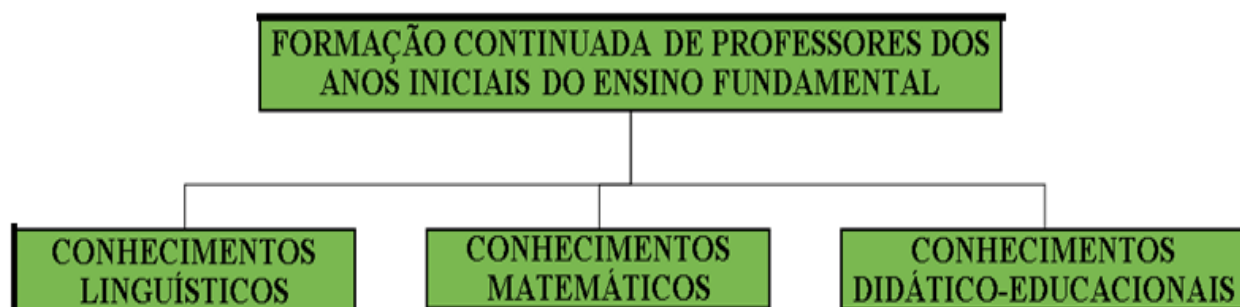


Figura 1. Proposta de sistema de conceitos - formação continuada de professores dos anos iniciais: domínio e subdomínios (Barros, 2022, p. 160).

Nas Figuras 2, 3 e 4, expomos as categorias criadas dentro de cada subdomínio, sendo que não apresentaremos as subcategorias de nosso sistema de conceitos, devido à limitação espacial dos diagramas:

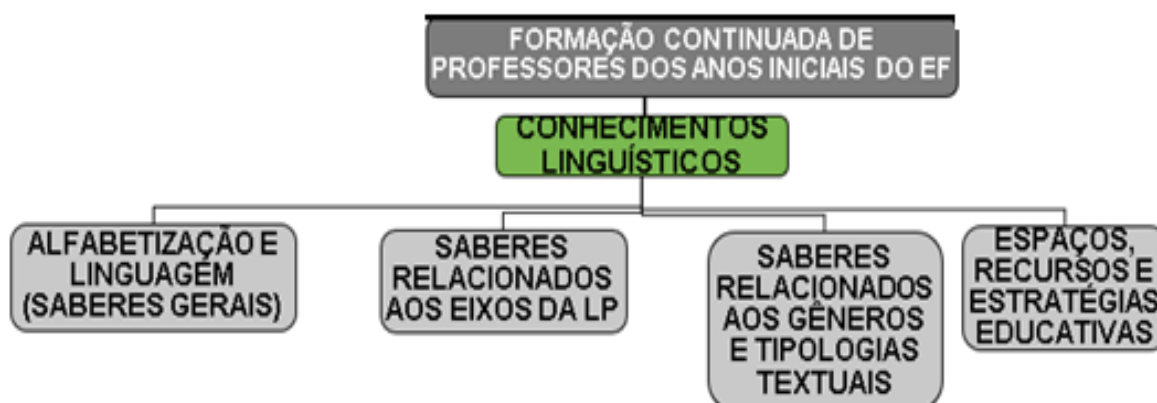


Figura 2. Proposta de sistema de conceitos: conhecimentos linguísticos (Barros, 2022, p. 161).

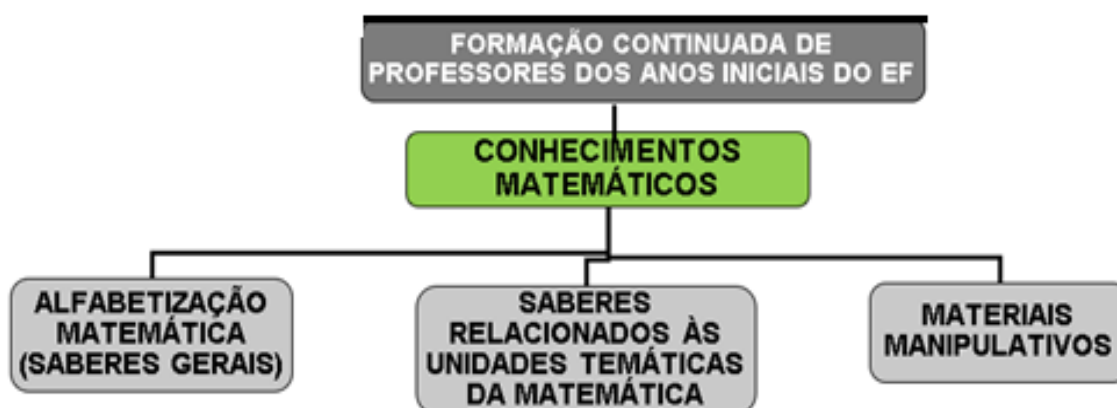


Figura 3. Proposta de sistema de conceitos: conhecimentos matemáticos (Barros, 2022, p. 161).

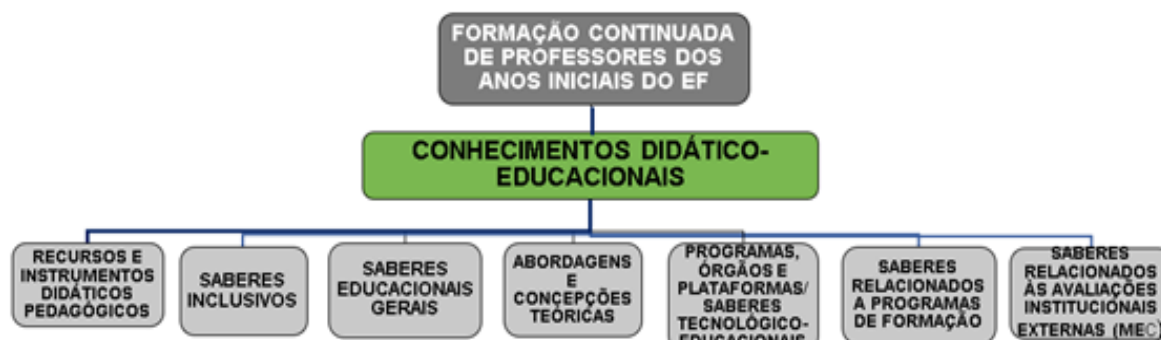


Figura 4. Proposta de sistema de conceitos: conhecimentos didático-educacionais (Barros, 2022, p. 161).

Expostos os subdomínios e suas ramificações, passamos a explicitar os critérios de reconhecimento adotados para definição do conjunto terminológico da área em estudo.

Critérios de reconhecimento terminológico

Dada a complexidade da área em estudo, sobretudo, por sua natureza inter- e multidisciplinar, cujas fronteiras são muito tênues com as de outras áreas, especialmente a da Educação, da Pedagogia, da Matemática e da Linguística, reconhecer sua terminologia constituiu-se em um grande desafio.

Quando se fala em formação continuada de professores, mais especificamente dos anos iniciais, na prática, não se abordam termos do campo sem que sejam mencionados os conteúdos escolares das disciplinas presentes no currículo da modalidade educacional em foco, por exemplo.

Diante do exposto, é que se torna relevante, a nosso ver, uma proposta que intente abarcar não somente o conjunto terminológico da área, mas que contemple também os termos relacionados a outras áreas e/ou disciplinas escolares tão presentes no cotidiano dos ambientes formativos dos educadores, uma vez que estes dialogam, inter cruzam-se e perpassam o domínio estudado. A esse respeito, Maciel esclarece que

[...] o fato de ser incluído na nomenclatura de um dicionário especializado não significa necessariamente que o termo se constitui, *stricto sensu*, em um item da terminologia repertoriada. Muitos dos termos incluídos apenas se relacionam com o domínio *lato sensu* (Maciel, 2001, p. 277).

Nesse contexto, quando se pensa em uma formação continuada de linguagem para os anos iniciais, por exemplo, é recorrente o estudo acerca de gêneros discursivos, sequências didáticas e, nesta última, comumente são apresentadas proposituras metodológicas de como se trabalhar os conteúdos que envolvem a leitura, escrita, oralidade e análise linguística, o que nos levou, portanto, a seleção/inclusão destas temáticas na definição de nosso conjunto terminológico.

Após essa rápida contextualização, passamos, então, a explicitar os critérios de reconhecimento terminológico empregados durante a pesquisa que neste trabalho apresentamos.

É importante destacar que, embora o estudo tenha utilizado programas/*softwares* especializados para realizar filtros e identificação lexical, por mais avançados que estes sejam, sempre executarão uma ação semiautomática, sendo necessária, neste ponto, a intervenção humana (de terminólogos, especialistas) para efetivar o reconhecimento das unidades lexicais com potencial terminológico para a área pesquisada.

Maciel (2001, p. 277) aponta que “[...] os traços que distinguem um termo de um não termo conferem o estatuto terminológico à unidade lexical em uma dada área”. Contudo, segundo a autora, nem sempre esses traços se apresentam de forma nítida, impondo-se, assim, a necessidade de se estabelecer critérios de reconhecimento terminológico.

Nesse sentido, mesmo com a utilização do *AntConc*, foi necessária a realização de uma análise minuciosa e criteriosa de cada unidade lexical, considerando-se, especialmente, os critérios da pertinência e da representatividade, além de parâmetros linguístico-textuais, baseados nos estudos de Cabré (1994), Maciel (2001), Krieger (2004) e Santiago (2011), a fim de que pudéssemos verificar se determinado candidato se constituía como termo da área de Formação Continuada de Professores dos Anos Iniciais do Ensino Fundamental, bem como examinar sua importância para área.

No que se refere à pertinência, Maciel (2001), baseada em Cabré (1994), subdivide-a em dois tipos: temática e pragmática. O primeiro consiste na “[...] propriedade de um termo pertencer a uma terminologia *stricto sensu* pelo fato de denominar um conceito que faz parte do campo cognitivo de domínio inventariado” (Maciel,

2001, p. 278). Em nosso estudo, termos como ‘Professor formador’, ‘Guia do formador’ e ‘Formador local’ são, indiscutivelmente, autênticos do domínio da Formação Continuada de Professores.

No entanto, segundo Maciel (2001), há os termos cuja pertinência temática não é nitidamente marcada, por serem oriundos de outras áreas conexas ou, ainda, de áreas que, de algum modo, colaboram para uma melhor compreensão dos conceitos englobados pela terminologia estudada. É necessário, assim, para definição dos termos, que se considere o contexto em que a comunicação se realiza.

Sob essa perspectiva, o segundo tipo de pertinência, a pragmática, trata-se da circunstância ou situação que possibilita que um termo integre uma dada terminologia, em razão do contexto em que ele se encontra inserido, abarcando conceitos de campos especializados que perpassam ou penetram um domínio principal, favorecendo a constituição de termos de aspectos mistos, pertencentes a mais de uma área (Santiago, 2011), como é o caso de ‘reciclagem’ (biologia), ‘matéria’ (química) e ‘currículo’ (mercado de trabalho/educação), presentes em nossa pesquisa.

Conforme já exposto, dada a natureza do domínio em estudo, que dialoga com outras áreas, no processo de reconhecimento dos termos, dentre esses dois critérios, o da pertinência pragmática foi o que mais sobressaiu na aplicação destes ao *corpus*.

Vale destacar que o duplo papel de umas das autoras deste artigo (pesquisadora/especialista) também foi de grande relevância, uma vez que sua experiência com a área em estudo contribuiu na difícil tarefa de reconhecimento dos termos com representatividade para o domínio. Representatividade esta que não está atrelada à frequência de aparição da unidade lexical em análise nos textos especializados, mas a sua capacidade de cobrir conceitos importantes da área.

No que se refere aos parâmetros linguístico-textuais, Krieger (2004) aponta que, na complexa tarefa de reconhecimento terminológico, deve-se levar em consideração “[...] os propósitos cognitivos e/ou pragmáticos da área temática [...]” e “[...] as particularidades textuais articuladoras da comunicação especializada” (Krieger, 2004, p. 333). Para Santiago (2011, p. 101), “[...] este parâmetro envolve, sobretudo, a observação das unidades lexicais presentes nos textos especializados, considerando seus contextos de ocorrência”.

No caso do estudo que realizamos, a análise das unidades lexicais em contextos de uso, nos textos especializados do *corpus* da área, foi de suma importância para a identificação dos candidatos a termos. Nessa etapa, as ferramentas do *AntConc* foram de grande valia, pois permitiram visualizar o contexto de uso da unidade lexical analisada em um pequeno recorte textual ou no texto integral de onde ela foi extraída no *corpus*.

Após o processamento do *corpus* pelo programa *AntConc* e submissão deste aos critérios de reconhecimento terminológico elencados (em especial, os de pertinência e de representatividade), entraram em cena especialistas da área da formação continuada de professores, com a missão de confirmar/reafirmar/validar ou refutar as unidades lexicais candidatas a termos do domínio pesquisado, sendo que a utilização dos verbos ‘confirmar’ e ‘reafirmar’ se deve ao fato já mencionado de que uma das autoras deste trabalho também é especialista da área. A validação foi concluída com a apreciação individual dos candidatos a termos por sete especialistas: seis delas, atuantes na área de formação continuada de professores; e uma em supervisão escolar.

Finalizamos esta etapa com um total de 2.851 termos validados para o domínio da Formação Continuada de Professores dos Anos Iniciais do Ensino Fundamental, que englobam saberes relacionados à Linguagem, à Matemática e à Educação em geral.

Considerações finais

Buscamos apresentar neste artigo o percurso para identificação da terminologia da área da formação continuada de professores dos anos iniciais, por meio da apresentação de árvores de domínio desta especialidade, estabelecidas a partir da determinação de parâmetros de reconhecimento terminológico, baseados na Teoria Comunicativa da Terminologia – TCT (Cabré, 1999) e na Linguística de *Corpus* (Berber Sardinha, 2000).

Nesse sentido, pautadas nos pressupostos teóricos da TCT (Cabré, 1999), que considera os aspectos comunicativos da linguagem especializada, é que selecionamos 157 textos especializados para a constituição do *corpus*, o qual, com auxílio das ferramentas computacionais *BootCaT* e *AntConc*, passou por processamento, extração e análise.

O resultado obtido foi um sistema de conceitos formado por 2.851 termos¹⁰, distribuídos nos três subdomínios estabelecidos pela pesquisa (conhecimentos linguísticos, matemáticos e didático-educacionais), que, por não terem se mostrado suficientes para atender a diversificada demanda de categorização terminológica, suscitaram a criação de outras categorias e subcategorias.

Todavia, no âmbito deste trabalho, dadas suas delimitações, foram apresentadas apenas as árvores de domínio do conjunto terminológico analisado (para acessar o sistema de conceitos em sua totalidade, ver Barros, 2022).

Compreendemos que há outras possibilidades de determinação de parâmetros linguísticos de reconhecimento terminológico, que intentem abarcar todo o universo da formação continuada de professores dos Anos Iniciais do Ensino Fundamental. Todavia, não pretendemos esgotar a questão. Ao contrário, apresentamos uma propositura de parâmetros para a identificação da terminologia da área em análise, que acreditamos que possa contribuir com estudos e produções terminográficas futuras, podendo servir de referência para uma área relativamente nova no que se refere ao seu reconhecimento terminológico.

Referências

- Barros, L. A. (2004). *Curso básico de terminologia*. São Paulo, SP: Editora da Universidade de São Paulo.
- Barros, L. S. (2022). *Formação continuada de professores dos anos iniciais: a tessitura de uma terminologia*. (Tese de Doutorado). Universidade Federal do Tocantins, Araguaína.
- Benveniste, É. (1989). *Problemas de Linguística Geral II*. Campinas, SP: Pontes.
- Berber Sardinha, T. (2000). *Linguística de corpus: histórico e problemática*. D.E.L.T.A., 16(2), 323-367.
- Berber Sardinha, T. (2004). *Linguística de corpus*. Barueri, SP: Manole.
- Biderman, M. T. C. (2001). As ciências do léxico. In A. M. P. P. Oliveira, & A. N. Isquierdo (Orgs.), *As ciências do léxico: lexicologia, lexicografia, terminologia* (p. 13-22). Campo Grande, MS: UFMS.
- Biderman, M. T. C. (2006). O conhecimento, a terminologia e o dicionário. *Ciência e Cultura*, 52(2), 35-37.
- Brasil. (2019). *Lei de diretrizes e bases da educação nacional. Lei nº 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional*. Brasília, DF: Senado Federal.
- Brito, V. M., & Viédes, S. C. A. (2015). A política educacional de alfabetização da criança consubstanciada pelos programas, projetos e pactos: o novo ou tudo de novo? *Revista HISTEDBR*, 63(1), 147-171.
- Cabré, M. T., & Estopà, R. (2005). Unidades de conocimiento especializado: caracterización y tipología. *Coneixement, Llenguatge i Discurs Especialitzat*, 7(1), 1-40. Recuperado de https://scholar.google.ca/citations?view_op=view_citation&hl=en&user=kDt9u4QAAAAJ&citation_for_view=kDt9u4QAAAAJ:u-x6o8ySG0sC
- Cabré, M. T. (1993). *La terminología: teoría, metodología, aplicaciones*. Barceleona, ES: Editorial Antártida.
- Cabré, M. T. (1994). Terminologie et dictionnaires. *Meta*, 4(39), 590-597.
- Cabré, M. T. (1999). *Terminología: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos*. Barcelona, ES: Universitat Pompeu Fabra/IULA.
- Duarte, A. R. C. (2017). A formação continuada como processo de aperfeiçoamento da atividade docente. *Revista Ciência Plural*, 3(2), 1-3.
- Freire, P. (1989). *A importância do ato de ler: em três artigos que se completam*. São Paulo, SP: Autores Associados; Editora Cortez.
- International Organization for Standardization [ISO]. (1990). *ISO 1087: terminology – vocabulary*. Genebra, CH: ISO.
- International Organization for Standardization [ISO]. (2000). *ISO 704: terminology work: principles and methods*. Recuperado de https://edisciplinas.usp.br/pluginfile.php/490906/mod_resource/content/1/ISO%20704.pdf.
- Krieger, M. G. (1998). Terminologia em contextos integradores: funcionalidade e fundamentos. *Organon*, 12(26), 1-7. DOI: <https://doi.org/10.22456/2238-8915.29556>
- Krieger, M. G. (2004). Do reconhecimento de terminologias: entre o linguístico e o textual. In A. N. Isquierdo, & M. G. Krieger (Orgs.), *As ciências do léxico: lexicologia, lexicografia, terminologia* (Vol. II, p. 327-339). São Paulo, SP: UFMS.

¹⁰ Neste artigo foram apresentadas apenas as árvores de domínio do conjunto terminológico analisado. Para acessar o sistema de conceitos em sua totalidade, ver Barros (2022).

- Krieger, M. G., & Finatto, M. J. B. (2020). *Introdução à terminologia: teoria & prática*. São Paulo, SP: Contexto.
- Maciel, A. M. B. (2001). Pertinência pragmática e nomenclatura de um dicionário terminológico. In M. G. Krieger, & A. M. B. Maciel (Orgs.), *Temas de terminologia* (p. 275-284). UFRGS/Humanitas/USP.
- Negrine, A. (1998). *Terapias corporais: a formação pessoal do adulto*. Vitória, ES: Editora Edita.
- Santiago, M. S. (2011). O léxico da educação a distância: parâmetros para o reconhecimento terminológico de uma área em desenvolvimento. *Domínios de Linguagem*, 5(2), 90-106.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.
- Souza, L. B., & Lima, Â. V. A. (2017). Formação continuada de professores: história e desafios. In *Anais do IV Simpósio Nacional de Linguagens e Gêneros Textuais* (p. 1-11). Recuperado de <https://editorarealize.com.br/artigo/visualizar/27353>