

INFORMAÇÕES GERAIS DO TRABALHO

Título do Trabalho:

CALIENT: Corpus de aprendizes da Língua Inglesa do Ensino Médio Técnico

Autor (es):

Shirlene Bemfica de Oliveira
Coordenadora do Projeto, Doutora em Estudos Linguísticos

Leticia Rafaela Alves Palheiros
Bolsista, PIBIC Jr., Discente do Ensino Médio Técnico de Mineração

Marcela Neves Del'Gaudio
Bolsista PIBIC Jr., Discente do Ensino Médio Técnico de Automação Industrial

Palavras-chave:

corpus de aprendizes, produção escrita, produção oral, língua inglesa, interação

Campus: Ouro Preto

Área do Conhecimento (CNPq): Linguística Aplicada; Linguística de Corpus

RESUMO

O CALIENT - Corpus de Aprendizes da Língua Inglesa do Ensino Médio Técnico é um banco de dados composto por amostras da produção linguística de alunos do ensino médio técnico. O corpus encontra-se em construção e vem sendo coletado e compilado a partir de outros projetos de pesquisa desenvolvidos na área de ensino e aprendizagem de línguas estrangeiras do IFMG campus Ouro Preto. O CALIENT futuramente será um banco de textos orais e escritos, produzidos por alunos de língua inglesa como língua adicional e língua estrangeira, escritos individualmente, em pares e em grupos e será disponibilizado a comunidade científica no formato eletrônico. Em sala de aula entre os anos de 2010 e 2017, as coletas tiveram o foco instrucional da língua inglesa foi dado no desenvolvimento das habilidades integradas, nos multiletramentos e as produções textuais orais e escritas de diversos gêneros textuais demonstram o pensamento crítico dos alunos. Nesses anos, os projetos passaram pelo Comitê Nacional de Ética na Pesquisa e todas as produções foram autorizadas pelos pais dos alunos por meio de termos de assentimento e de consentimento. As pesquisas tiveram grande impacto acadêmico, social, econômico e tecnológico e o CALIENT, resultado dessas pesquisas tem grande potencial inovador, uma vez que não existe no Brasil nenhum corpus eletrônico com amostras de alunos do ensino médio técnico.

INTRODUÇÃO:

Este trabalho tem como objetivo criar e organizar o Corpus de Aprendizes da Língua Inglesa do Ensino Médio Técnico – CALIENT - para disponibilizar para a comunidade científica em formato eletrônico. O corpus será compilado e organizado com base nos projetos de pesquisa desenvolvidos pelo autor entre os anos de 2010 e 2017 no âmbito da escola regular técnica. As pesquisas sobre corpora de aprendizes são muito recentes e o caráter inovador deste estudo se deve a “uma grande carência de estudos sobre a interlíngua de aprendizes brasileiros” e de “compilações de corpus de aprendizes no Estado de Minas Gerais” (DUTRA, 2010, p. 03). Esta pesquisa pretende contribuir com as abordagens de ensino de inglês como língua adicional e estrangeira no contexto de escola regular, no que tange ao processo de ensino e aprendizagem da língua com foco na ativação dos conhecimentos prévios dos alunos, da interação entre

eles para a construção de conhecimentos que se materializam em uma diversidade de gêneros textuais orais e escritos.

Em estudos anteriores, pesquisadores aliaram o uso das tarefas de produção escrita individual, colaborativa com ênfase no processo e análises foram feitas a partir da produção escrita de diversos gêneros, da análise discursiva das interações e do processo de colaboração dos alunos participantes tanto em salas de aula como em laboratório. Os dados desses estudos já publicados são fonte rica que podem propiciar outras investigações nas áreas de Educação, Linguística Aplicada, Linguística de Corpus, uma vez que os dados propiciam a compreensão de como alunos do ensino médio técnico empregam a linguagem nesses gêneros (BERBER SARDINHA, 2004; ALUISIO et al., 2006; GONZALES, 2007). Além disso, podem auxiliar na criação de novas metodologias de investigação e de ensino de línguas no âmbito do IFMG.

A compilação e organização do CALIEMT é embasada teórico e metodologicamente pela Linguística de Corpus (LC) que é a área do conhecimento que estuda a linguagem por meio da utilização do computador (GONZÁLES, 2007, p. 8). Ela é definida como uma maneira de se chegar à linguagem por meio da análise dos padrões probabilísticos que se constroem nos contextos em que os falantes os empregam (BERBER-SARDINHA, 2004). A principal característica da LC é, segundo Gonzáles (2007, p. 8), a observação de dados empíricos armazenados em bancos de dados que compõem um corpus, com a utilização de ferramentas eletrônicas que auxiliam na análise de dados verificando os fenômenos da língua em uso. Para este tipo de análise, recorreremos a um corpus que é entendido como

um conjunto de dados linguísticos (pertencentes ao uso escrito da língua), sistematizado segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso de algum de seu âmbitos, dispostos de tal modo que possam se processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise (SANCHEZ, 1995, p.8-9).

De acordo com Resende e Maverick (2016, p. 8), o corpus pode ser falado ou escrito. O corpus falado é aquele composto por porções de falas que foram transcritas e preparadas para serem armazenadas em um banco de dados, no caso desse estudo, as transcrições de falas das interações para as produções textuais gravadas presencialmente com o uso de filmadoras. Os dados escritos do CALIEMT serão textos escritos individualmente, em pares ou grupos na língua inglesa por alunos do ensino médio técnico. Esses materiais são digitalizados e como todos os corpora, armazenados em um banco de dados.

Além de serem falados ou escritos, os corpora também podem ser considerados sincrônicos, diacrônicos ou históricos. Os corpora sincrônicos são formados por textos com uma linguagem contemporânea, os corpora diacrônicos retratam diferentes períodos de tempo e os corpora históricos retratam um tempo passado (RESENDE; MAVERICK, 2016, p. 8). O CALIEMT será um corpus sincrônico composto por textos de temáticas contemporâneas e cunho social. Segundo os autores, há ainda os corpora de amostragem e os monitores.

Os corpora de amostragem são compostos por porções de textos ou variedades textuais finitas e são planejados para uma amostra finita da linguagem como um todo. Eles são sempre estáticos, ou seja, depois de compilados, não é mais possível acrescentar e nem diminuir dados. Os corpora monitores são compostos por dados constantemente reciclados de forma que possam refletir o estado atual de uma língua e por essa característica são também chamados de corpora dinâmicos ou corpora orgânicos, caso do

CALIENT (RESENDE; MAVERICK, 2016, p. 8). Quando os corpora são compostos por um número semelhante de textos ou de dados, ou seja, se possuem uma quantidade de dados equilibrada, como por exemplo, o mesmo número de textos ou o mesmo número de gêneros ou registros, eles recebem o nome de corpora equilibrados (RESENDE; MAVERICK, 2016, p. 9). Os corpora podem ser especializados, ou seja, específico de uma área ou paralelos que são também compostos por textos em dois idiomas diferentes, mas, diferente dos comparáveis, os textos são formados por traduções e podem ou não ser alinhadas. Quando os corpora são paralelos e alinhados isto quer dizer que uma linha do texto original fica alinhada com a linha do texto traduzido e pode-se ler o texto fonte e o texto traduzido um ao lado do outro ou um acima do outro. Por último, temos os corpora de treinamento que também são chamados de corpora de testes, eles são desenhados para permitir o desenvolvimento de aplicações e ferramentas de análise linguística (RESENDE; MAVERICK, 2016, p. 10).

METODOLOGIA:

Esta pesquisa, de natureza empírica, está sendo desenvolvida no Laboratório de Ensino, Pesquisa e Extensão em Línguas Estrangeiras IFMG - OP com a participação de 2 bolsistas do ensino médio técnico e a pesquisadora. O processo de compilação e organização do CALIENT está ancorado pela literatura da área de Linguística de Corpus que sugere que para a compilação de um corpus, os pesquisadores devem seguir alguns procedimentos: primeiramente, a elaboração do projeto do corpus, que inclui a seleção ou organização dos textos considerando a autenticidade, representatividade, balanceamento e diversidade (BIBER et al., 1998; BERBER SARDINHA, 2004; SINCLAIR, 2005). O processo de compilação está sendo feito com base em (RESENDE; MAVERICK, 2016) que consiste na:

1. seleção de textos (a partir da Web e a partir de textos impressos) e pedidos de permissão de uso;
2. manipulação e limpeza do corpus;
3. etiquetagem (anotação estrutural, cabeçalhos) e nomeação dos arquivos;
4. exclusão das etiquetas XML2 para processamento do corpus e contagem de palavras (RESENDE; MAVERICK, 2016, p. 11).

No caso do CALIENT, alguns dos textos já foram produzidos por alunos do ensino médio técnico dos níveis básicos em salas de aula e no laboratório de línguas estrangeiras por alunos dos níveis intermediário e avançado que já conseguem manter uma conversação na língua inglesa. Em sala de aula, os alunos tiveram encontros semanais de 1h e 40 min. e utilizam materiais didáticos aprovados pelo PNLD. Durante as aulas foram desenvolvidas atividades que contemplavam as habilidades de compreensão e produção oral e escrita (*listening, speaking, reading and writing*), além de pronúncia, gramática e vocabulário.

Além disso, a língua tem sido ensinada como uma língua global (CRISTAL, 2003) com base na valorização dos processos de interculturalidade e ênfase nos conhecimentos e experiências prévias dos participantes. Houve também uma preocupação em promover o empoderamento dos jovens por meio das interações, ativismo e no engajamento de causas sociais para promover um processo de reconstrução de suas identidades (RAJAGOPALAN, 2003). Os alunos convidados a coletarem dados no laboratório também participaram de um projeto de extensão de autoria do mesmo pesquisador e também tiveram uma abordagem de ensino para o desenvolvimento linguístico, para a livre expressão da imaginação e da criatividade. O foco no projeto era nos Letramentos Sociais, com propostas de tarefas em que os

participantes eram desafiados a trabalhar em equipe, a questionar, resolver problemas (STREET, 2014); Esses dados orais e escritos foram anteriormente analisados com foco na produção textual e no processo de construção colaborativa. Estas análises foram feitas com o auxílio de concordanciadores eletrônicos apresentando as construções lexicais mais frequentes e seus colocados. A discussão foi pautada pelos pressupostos teóricos de aquisição e de letramento em LE.

O levantamento bibliográfico para as análises está sendo feito ao longo do processo pela pesquisadora e o(s) bolsista(s). A triangulação dos dados coletados através dos diferentes instrumentos de análises e tarefas pretende comparar as diferentes perspectivas do mesmo assunto. Para o período de 2018, o foco do projeto está na sistematização do processo de compilação, manipulação, nomeação dos arquivos de textos e arquivamento digital dos pedidos de permissão de uso das produções e para esse trabalho contamos com a participação de dois bolsistas: um para a compilação do corpus escrito e outro para o corpus oral.

Os arquivos selecionados são salvos em um servidor local já criado pela TI no Laboratório de Pesquisa e Extensão em Línguas Estrangeiras em Terminologia (CALIEMT). As imagens dos mapas conceituais são digitalizados utilizando o *Cmptools* para serem integradas e linkadas aos textos escritos. Para a limpeza e manipulação do corpus e para que os textos, em seus formatos originais de disponibilização, pudessem ser corretamente processados por ferramentas computacionais, será necessário converter manualmente todos os arquivos de texto de seus formatos originais (*Microsoft Word* de extensão “.doc”, para um único formato padrão, no caso o Bloco de Notas de extensão “.txt”, pois este não possui códigos de formatação, mas apenas caracteres do teclado (letras, números e símbolos ortográficos). Além disso, são excluídas possíveis tabelas, gráficos, fórmulas, cálculos, fotos e toda informação que não esteja em forma de texto para a limpeza do corpus.

O trabalho posterior segundo Resende e Maverick (2016), é a anotação estrutural, geração de cabeçalhos e nomeação de arquivos¹. Por meio da versão adaptada do Editor de Cabeçalhos do Projeto *Lacio-Web*, será possível realizar a anotação estrutural dos textos, adicionando-se em cada texto um cabeçalho e uma nomeação específica e padronizada. A anotação estrutural, de acordo com Aluísio e Almeida (2006), compreende a marcação de dados externos e internos dos textos. Os dados externos são formados pela documentação do corpus na forma de um cabeçalho que inclui os metadados textuais, isto é, tamanho do arquivo em palavras, autoria, a tipologia textual e informação sobre a distribuição do corpus. Os dados internos é a anotação de segmentação do texto cru, que envolve:

- a) marcação da estrutura geral – capítulos, parágrafos, títulos e subtítulos;
- b) marcação da estrutura de subparágrafos – elementos que são de interesse linguístico, tais como sentenças, citações, palavras, abreviações, nomes, referências, datas e ênfases tipográficas do tipo negrito, itálico, sublinhado, etc (ALUÍSIO; ALMEIDA, 2006, p. 14).

De acordo com os autores, a anotação estrutural do texto é um processo pelo qual parcelas de texto que constituem informações diferenciadas por sua relevância (ou pela falta de relevância), e são marcadas por etiquetas por meio do Editor de Cabeçalhos. Aluísio e Almeida (2006) instruem que são necessárias para a geração dos cabeçalhos: as informações: título, subtítulo, ano e local da produção, comentários (informações adicionais sobre o processo de produção), autoria (Individual, par, grupo, ano de instrução,

¹ Este trabalho será feito em 2019.

sexo do autor), Gênero Textual, Tipo de Texto. Essas informações mostram-se relevantes para o CALIEMT bem como para outros projetos que pretendam reutilizar o corpus. Reproduzem-se, abaixo, a título de exemplo, todas as informações que podem ser recuperadas do texto:

**** separa os textos
 * separa as variáveis sexo 1 masculino, 2 feminino
 ins ano no ensino médio técnico
 cur curso no ensino médio técnico
 MIN Mineração
 Exemplo: **** *ano_2011 *n_001 *sex_11 *ins_1 *cur_MIN

Com a finalização do corpus, será preciso contar o número de palavras que ele contém. Para isso, segundo Aluisio e Almeida (2006) não é recomendável considerar as etiquetas XML inseridas pelo Editor de Cabeçalhos, pois o programa contaria todas as palavras que estão entre as etiquetas como sendo palavras do corpus. Dessa forma, é preciso remover tais etiquetas a fim de deixar o corpus enxuto. Uma outra razão para a exclusão das etiquetas é o fato de que muitos programas de PLN não processam com facilidade (e alguns simplesmente não processam) corpus com etiquetas XML.

RESULTADOS E DISCUSSÕES:

O trabalho de novas coletas, organização e compilação foi iniciado em 2018 e a pesquisa apresenta dados parciais que serão apresentados a seguir. O corpus será especializado com foco no emprego da língua inglesa por alunos do ensino médio técnico, conforme mostra a tabela abaixo:

Tabela 1: Composição do CALIEMT

CRITÉRIOS	CORPUS DE LINGUA ESTRANGEIRA	
Abrangência	Regional: região sudeste do Brasil	
Meio	Escrito e oral	
Tempo	Contemporâneo	
Renovação	Dinâmico Anual	
Línguas	Inglês	
Produtores	Aprendizes de língua inglesa do Ensino Médio Técnico	
Emprego	Estudo	
Estratégia	2011: Texto produzido a partir de duas versões 2012-2014: Texto produzido a partir de mapas conceituais e produção escrita colaborativa 2015-2017: Escrita colaborativa e gravação das interações	
Gêneros Registros	Texto expositivo de definição (2011) Mapas Conceituais (2014-2015) Textos argumentativos (2012-2017) Poemas, Narrativas, Depoimentos, gravações de interações (2015 – 2017)	- Distúrbios alimentares - Bullying - Violência na sociedade moderna - Arte e sociedade - Temas propostos pelo ICLE

Fonte: Dados da Pesquisa

O CALIEMT é um corpora especializado, pois é específico de uma área: de língua inglesa. Até o momento, ele tem dados representativos de alunos da região sudeste e por isso é denominado corpus de aprendiz. Ele possui textos de diversos registros em dois idiomas: inglês e português.

Os dados produzidos para o CALIEMT foram coletados anualmente, o que demonstra uma constante dinâmica e reflete algumas características da interlíngua dos aprendizes. Ele não será equilibrado, pois está sendo construído com diversos gêneros e registros por aprendizes de níveis diferentes.

O exemplo que apresentamos a seguir advém da organização dos dados de 2010 que está em andamento. O objetivo da coleta na ocasião era analisar a ocorrência de pacotes lexicais em produções escritas de alunos do ensino médio técnico. O estudo, numa perspectiva empírica, foi desenvolvido com a participação de 230 alunos do ensino médio técnico de um Instituto Federal Minas Gerais. Os dados foram coletados em aulas de inglês, nas quais foram desenvolvidas as habilidades de compreensão e produção (leitura, audição, escrita e fala). A coleta de dados foi realizada em um procedimento em três fases, conforme apresentado na tabela 1:

Tabela 1: Coleta de textos de 2010

Fase	Tarefa	Objetivo
Corpus 1	Distúrbios Alimentares	- Verificar a ocorrência de pacotes lexicais.
Fase 2	Atividades com foco linguístico	- Dar <i>feedback</i> positivo - Aumentar a percepção dos alunos sobre a produção deles
Fase 3 Corpus 2	Reescritura do texto	- Melhorar a produção escrita.
Análise dos pacotes lexicais	Contraste da produção 1 e 2	- Contrastar o texto dos alunos antes e depois das tarefas com foco linguístico e mapear os pacotes lexicais. -

Fonte: Dados da Pesquisa

Na ocasião, os textos foram digitados em arquivos este ano, foram separados e salvos no CALIEMT: o primeiro corpus com 41 textos e o segundo com 40 sobre distúrbios alimentares. Esses oitenta textos foram analisados na ocasião e geraram algumas publicações que podem ser conferidas no anexo 1. A tabela 2 abaixo apresenta o tamanho de cada corpora comparando o total de palavras (12290 tokens de palavras) e o número de palavras usadas (tipos de palavras de 1925) em 804 sentenças, bem como a riqueza lexical da riqueza. De acordo com Finatto et. al. (2011), “corresponde à medida dada pela razão entre o número de palavras diferentes no corpus (formulários) com o número total de palavras” (p. 219).

Tabela 2: Composição dos corpora

Text file	File size	Tokens	Types	Type/token	Sentences
Overall	74311	12290	1925	15,7	804
corpus1.txt	35536	5903	1375	23,3	395
corpus2.txt	38775	6387	1456	22,9	409

Fonte: Dados da pesquisa calculado pelo *Wordsmith Tools*

CONCLUSÕES:

A pesquisa de organização e compilação do CALIEMT ainda é inicial, mas tem potencial inovador por não existir no Brasil um corpus dessa natureza. Os resultados da compilação e organização do banco de dados podem contribuir fortemente para as pesquisas nas áreas de Educação, Linguística de Corpus e Linguística Aplicada. O produto parcial a ser apresentado em dezembro de 2018 será um banco de dados eletrônicos com amostras de alunos dos níveis básicos, intermediário e avançado. Futuramente, com financiamento específico, pretende-se transformá-lo em um corpus em servidor próprio a ser divulgado a comunidade científica externa. O CALIEMT tem grande potencial social e impacto tecnológico: para os bolsistas, a participação trará o desenvolvimento de habilidades no processo de criação de bancos de dados e no uso de softwares. O docente enquanto pesquisador de sua própria prática pode refletir e tomar consciência do nível de abstração e generalidade presentes na profissão e das possibilidades de inclusão de novos saberes, metodologias e construções discursivas que beneficiem o aprendizado dos alunos e sua própria práxis. Aos alunos e bolsistas é dada a oportunidade de se tornarem mais ativos e conscientes de seu papel enquanto cidadãos e de serem pessoas capazes de transformar a educação por meio da produção de conhecimento científico. Além disso, o CALIEMT tem impacto econômico na medida em que pode contribuir com outros pesquisadores que poderão utilizar o banco de dados para outras pesquisas otimizando o processo científico

REFERÊNCIAS BIBLIOGRÁFICAS:

- ALUÍSIO, S.M.; ALMEIDA, G. M. B. .O que é e como se constrói um Corpus? Lições aprendidas na compilação de vários corpora para pesquisa lingüística. Calidoscópio (UNISINOS), v. 4, p. 156- 178, 2006. Disponível em: http://www.unisinos.br/publicacoes_cientificas/images/stories/pdfs_calidoscopio/vol4n3/art04_aluisio.pdf
- BERBER SARDINHA, A. P. Linguística de Corpus: Histórico e Problemática. In: D.E.L.T.A. v.16, n. 2, 2000, p. 323-367.
- BERBER SARDINHA, A. P. Linguística de Corpus. Barueri-SP. Manole, 2004.
- BIBER, D. S.; CONRAD; REPPEN, R. Corpus Linguistics: Investigating language structure and use. Cambridge: Cambridge University Press, 1998.
- CRISTAL, D. English as a global language. Cambridge: Cambridge University Press, 2003.
- DUTRA, D. P. Agrupamentos lexicais na escrita de aprendizes brasileiros de inglês: um estudo baseado em corpus. Plano de trabalho apresentado ao Programa Pesquisador Mineiro. Edital FAPEMIG 03/2010.
- GONZÁLES, Z. M. G. Lingüística de Corpus na análise do Internetês. Dissertação (Mestrado em Linguística Aplicada e Estudos da Linguagem) - Faculdade de Letras. Pontifícia Universidade Católica de São Paulo, São Paulo, 2007.
- RESENDE, S. V.; MAVERICK, R. Planejamento, Compilação e Organização de Corpora. Blucher Social Sciences Proceedings, v.2, n. 3, Março de 2016. Disponível em: http://pdf.blucher.com.br/s3-sa-east-1.amazonaws.com/socialsciencesproceedings/viii-ebic-xiiiclc/06_artigo_03.pdf Acesso em 28/01/2018.

Anexo 1 Participação em Congressos, publicações e/ou pedidos de proteção intelectual:

As publicações e participações em eventos apresentados a seguir foram realizados com base nos textos que estão sendo compilados para o CALIEMT.

Apresentação Oral no 11º Congresso Brasileiro de Linguística Aplicada, 2015. (Congresso). O uso de mapas conceituais na produção escrita em língua inglesa.

Apresentação de Poster / Paineis na Semana de Ciência e Tecnologia do IFMG, 2015. (Feira). Produção Escrita Colaborativa: corpus de aprendizes de língua inglesa.

Apresentação de Poster / Paineis na Semana de Ciência e Tecnologia do IFMG, 2015. (Feira). O uso de mapas conceituais na produção escrita em língua inglesa.

Apresentação de Poster / Paineis na Semana de Ciência e Tecnologia do IFMG, 2015. (Feira). O uso de mapas conceituais na produção escrita em língua inglesa.

X Simpósio de Formação e Profissão Docente SIMPOED, 2015. (Simpósio). Aprender com o outro: análise da produção escrita colaborativa em língua inglesa.

Apresentação Oral no I Seminário Educação, Trabalho e Sociedade, 2014. (Seminário) Ampliando conhecimentos com mapas conceituais no ensino de língua inglesa.

OLIVEIRA, S. B. Iniciação científica júnior: experiências de pesquisas com foco no ensino e aprendizagem de línguas adicionais com uso de corpora In: 10º Congresso Brasileiro de Linguística Aplicada, 2013, Rio de Janeiro. Anais do 10º Congresso Brasileiro de Linguística Aplicada (CBLA). Rio de Janeiro: Associação de Linguística do Brasil, 2013. v.2. p.1 – 30.

BRAGA, I. M. S.; OLIVEIRA, S. B.; REIS, B. M. O uso de mapas conceituais como ferramenta para a aprendizagem de língua inglesa: do conhecimento prévio à leitura e produção escrita. Semana de Ciência e Tecnologia, 2013.

BRAGA, I. M. S.; OLIVEIRA, S. B.; REIS, B. M. O conhecimento prévio em língua inglesa: um alicerce para a aprendizagem significativa de uma língua estrangeira moderna. Semana de Ciência e Tecnologia, 2013.

OLIVEIRA, S. B.; ROSSI, A. M. O.; LEITE, G. M. F.; CARMO S, K. O.; OLIVEIRA, T. M. Pacotes Lexicais em corpus de aprendizes do ensino médio In: X Encontro de Linguística de Corpus : aspectos metodológicos dos estudos de corpora, 2012, BELO HORIZONTE. ANAIS DO X ENCONTRO DE LINGUÍSTICA DE CORPUS ASPECTOS METODOLÓGICOS DOS ESTUDOS DE CORPORA. BELO HORIZONTE: Faculdade de Letras - UFMG, 2012. p. 337 – 363.

OLIVEIRA, S. B.; FORTES, A. R.; SOL, M. T. A. Formulaic sentences versus lexical bundles: Aspectos Lingüísticos da Interlíngua em Corpus de Aprendizes de LE, 2010. (Outra, Apresentação de Trabalho) Referências adicionais : Brasil/Português. Meio de divulgação: Impresso; Local: OURO PRETO; Cidade: MINAS GERAIS; Evento: SEMANA DE CIENCIA E TECNOLOGIA 2010; Inst. promotora/financiadora: Instituto Federal Minas Gerais - Campus Ouro Preto.

17 AILA - The 18th World Congress. COLLABORATIVE WRITING AS AN EDUCATIONAL TOOL TO LEARN ENGLISH IN A PUBLIC SCHOOL. 2017.

I Seminário Educação, Trabalho e Sociedade. Ampliando conhecimentos com mapas conceituais no ensino de língua inglesa. 2014. (Seminário).

I Seminário Interinstitucional de Extensão em Formação de Professores de Línguas: a integração entre formação inicial e continuada. Movimentos de mudança conceitual na educação continuada. 2014. (Seminário).

XII Encontro de Linguística de Corpus. Mapas conceituais como ferramenta para construção de Corpus de Aprendizes de língua inglesa. 2014. (Encontro).

10º Congresso Brasileiro de Linguística Aplicada. Iniciação científica júnior: experiências de pesquisas com foco no ensino e aprendizagem de línguas adicionais com uso de corpora. 2013. (Congresso).

Semana de Ciência e Tecnologia. O conhecimento prévio em língua inglesa: um alicerce para a aprendizagem significativa de uma língua estrangeira moderna. 2013. (Outra).

ENCONTRO INTERNACIONAL SOBRE NOVOS LETRAMENTOS. 2012. (Encontro). XI SEMANA DE LETRAS DA UFOP. Práticas de ensino/aprendizagem de inglês por meio de projetos: propostas de um ensino significativo. 2010. (Oficina).

Anexo 2

Publicação em Periódico

OLIVEIRA, S. B. Produção escrita colaborativa: o olhar de aprendizes da língua inglesa sobre a arte. Fórum Linguístico (Online). , 2017. Qualis A2

Anexo 3

Premiações

- Premiação no Seminário Interno de pesquisas do IFMG Campus Ouro Preto em 2016. Segundo lugar na modalidade de comunicação oral e terceiro lugar na modalidade de painel: Produção escrita colaborativa: corpus de aprendizes de língua inglesa sobre arte.
- Premiação no Seminário Interno de pesquisas do IFMG Campus Ouro Preto em 2015. Segundo lugar na modalidade de comunicação oral e terceiro lugar na modalidade de painel: Produção escrita colaborativa: corpus de aprendizes de língua inglesa sobre arte.
- Melhor trabalho de Iniciação Científica apresentado no Seminário Interno de pesquisas do IFMG Campus Ouro Preto em 2014 com a comunicação oral O uso de mapas conceituais na produção escrita em língua inglesa.