



Título do projeto: Compilação de um corpus anônimo de textos e excertos produzidos no curso ProFIS.

Bolsista: Henrique Nunes Marinho

Orientador: Inês Signorini

Local de execução: Instituto de Estudos da Linguagem – Unicamp

Vigência: 08/2019 – 08/2020

1. Introdução

Este projeto teve o intuito de dar continuidade à compilação de textos e excertos de alunos do curso ProFIS, matriculados em disciplinas obrigatórias do 1º semestre, as quais estão desde 2011 sob responsabilidade do Departamento de Linguística Aplicada do Instituto de Estudos da Linguagem (IEL). A compilação se iniciou em 2018 e contou com a participação de outros bolsistas de iniciação científica empenhados na constituição do *corpus*. Coube especificamente a esse projeto, realizado entre os anos de 2019 e 2020, o trabalho com os documentos produzidos no 1º semestre de 2017, na disciplina LA083 – Leitura e Produção de Textos Acadêmicos I, ministrada pela profa. Inês Signorini.

A relevância do material compilado se deve ao fato de os sujeitos que o produziram serem todos oriundos de escolas públicas da cidade de Campinas, ingressando na universidade através do ENEM e sem passar pela seleção do vestibular regular. Assim, a produção escrita que esses indivíduos realizam ao chegar no contexto acadêmico é uma relevante base de dados para pesquisas relativas ao ensino do letramento acadêmico e suas relações com a educação básica.

Além da compilação, este projeto teve como objetivo secundário a análise da primeira atividade realizada pelos alunos de 2017 na disciplina. Nela, entre outras questões de diagnóstico, os estudantes tiveram que responder ao seguinte comando: “Quais são suas dificuldades no campo da leitura e da escrita? Especifique”. A análise das respostas se deteve somente às dificuldades relativas ao campo da escrita, as quais serão esmiuçadas mais adiante.

2. Metodologia

2.1 Compilação dos textos e excertos

A compilação dos textos e excertos seguiu as seguintes etapas: a) levantamento e organização dos documentos físicos (impressos e manuscritos); b) digitalização dos documentos; c) transcrição dos documentos; d) inserção dos dados no *corpus* via *Google Drive*. Nesse processo, todos as produções - arquivos originalmente feitos em computador, documentos impressos e manuscritos - foram anonimizadas seguindo uma nomenclatura que o exemplo a seguir demonstra: S10_M18_20171_P2; no qual “S10” é um número arbitrário dado ao autor do texto para deixá-lo anônimo; “M18” se refere ao

sexo (M = masculino e F = feminino) seguido da idade do sujeito; “20171” designa o ano acrescido do semestre (1º ou 2º) no qual o texto foi produzido; e por último, “P2” é referente ao tipo de documento inserido (P1 = 1ª prova, P2 = 2ª prova, E = exame, A1 = 1ª atividade *etc*).

Os textos manuscritos e impressos - dos quais não havia o arquivo original digital - foram escaneados em .pdf e transcritos posteriormente em um arquivo .txt com a mesma nomenclatura de sua digitalização. Com relação aos arquivos entregues pelos alunos já em .word, bastou apenas a conversão para .txt e a retirada de qualquer dado que os identificasse.

Além disso, a transcrição dos textos manuscritos, que são maioria do *corpus*, seguiu algumas notações acordadas pelo grupo de pesquisa para destacar características dos originais. São elas as seguintes: <riscado>, usada quando o autor riscava um termo ou palavra; <acrescido>, empregada nos casos em que uma palavra ou frase foi inserida através de setas ou acima de outros termos; <incompreensível>, utilizada quando houve incompreensão de alguma frase ou palavra do original.

2.2 Levantamento da atividade diagnóstica para análise

Para a análise da atividade realizada pelos alunos no seu primeiro dia da disciplina na qual eles descreveram suas dificuldades de escrita, foi realizado um levantamento quantitativo das respostas dos estudantes. Naquelas em que os autores relatavam apenas uma dificuldade (exemplo 1), foi realizada a cópia integral da frase do aluno. Já em outras em que o estudante afirmava ter mais de uma dificuldade (exemplo 2), essas eram separadas em duas ocorrências distintas. Desse modo, dos 153 documentos utilizados obteve-se 233 ocorrências de dificuldades. Esses resultados foram posteriormente sistematizados em quatro grupos descritos na seção seguinte.

Exemplo 1: “Elaborar boas ideias para textos de uma forma que seja pouco redundante” (S606_F20_20171_D1).

Exemplo 2: “(1)Tenho dificuldade principalmente com gramática e em relação a (2) agrupar ou separar ideias ao escrevê-las” (S577_F19_20171_D1)

3. Resultados

A compilação dos textos e excertos resultou na inserção no *corpus* de 973 documentos produzidos pelos alunos, dentre os quais 766 foram manuscritos e 207 foram arquivos digitais no formato .word ou impressos. Esses documentos são relativos às diversas atividades aplicadas na disciplina, desde avaliações, exercícios em sala de aula e relatórios produzidos no final do semestre.

Com relação à análise da atividade sobre as dificuldades de escrita, a sistematização descrita anteriormente levou aos seguintes resultados divididos em quatro grupos:

Dificuldade	Ocorrência	Porcentagem
1) Dificuldade textual no/nas:		
A) Nível lexical	46	19,74%
B) Nível da frase / período	10	4,29%

C) Nível do parágrafo	2	0,85%
D) Nível do texto	92	39,48%
E) Convenções gráficas	47	20,17%
2) “Dificuldades gramaticais” ou na “transposição mental” do texto	26	11,15%
3) Caligrafia	4	1,71%
4) Nenhuma dificuldade ou nenhuma resposta	6	2,57%

O primeiro grupo (84,53%) é formado pelas respostas que citaram algum aspecto linguístico-textual específico nas cinco subcategorias listadas: nível lexical, da frase/período, do parágrafo, do texto e das convenções gráficas. O segundo grupo (11,15%) se refere as respostas nas quais houve algum comentário vago sobre dificuldade com a gramática ou com a “transposição mental” das ideias de elaboração do texto para o papel. Já o terceiro grupo (1,71%) é composto por apenas algumas respostas que citaram a caligrafia como uma dificuldade. O quarto e último grupo (2,57%) é composto por respostas de alunos que afirmaram não ter nenhuma dificuldade ou que não responderam à pergunta.

4. Discussão dos resultados

Ao final da análise das respostas, observou-se que a maioria dos estudantes conseguiu citar dificuldades de escrita que envolvessem algum aspecto linguístico-textual (84,53%). Essas respostas não utilizaram em sua maioria o vocabulário gramatical ensinado na escola. Como exemplo disso, pode-se observar que entre os doze alunos que afirmaram ter alguma dificuldade de ortografia, apenas quatro de fato utilizaram o termo. Tendo em vista que o ensino gramatical nas escolas, apesar das progressivas mudanças nas últimas décadas, ainda é muito pautado por uma abordagem tradicional que foca na transmissão de nomenclaturas gramaticais aos estudantes sem associá-las a reflexões linguísticas ou a produção de textos (MENDONÇA, 2006; GERALDI, 2011), talvez não seja surpreendente que os alunos não utilizem essa nomenclatura quando precisam fazer um autodiagnóstico que poderia ser mais específico e melhor elaborado se utilizassem os termos gramaticais que são relevantes para uma reflexão metalinguística.

Contudo, é positivo o fato de que a maioria das respostas concentradas no grupo 1 se refiram a algum aspecto linguístico-textual, enquanto as respostas do grupo 2 que citam vagamente dificuldades gramaticais ou na “transmissão mental” das ideias de texto para o papel somam apenas 11,15% das ocorrências. Esse número é ainda menos significativo quando nota-se que 70% dos alunos que dizem ter “dificuldades gramaticais” sem especificá-las também citam outros problemas em suas respostas. Já 77,7% dos estudantes com dificuldades de “transmissão” de suas ideias não relataram outra dificuldade, evidenciando um problema em, de fato, especificar suas inabilidades. Esses alunos, que representam apenas 5,3% do número total de estudantes, demonstram uma concepção de escrita que não a entende como uma atividade sociocognitiva que estabelece o sentido no texto através da interação entre locutores e interlocutores, mas sim como simples instrumento de transmissão de uma mensagem centrada na consciência do autor (KOCH, ELIAS; 2011).

Vale ainda ressaltar que entre as respostas do grupo 1, as do nível do texto possuem a maior porcentagem (39,48%), número bastante expressivo se comparado com as dificuldades relativas ao nível da frase e do parágrafo, que somam apenas 4,29% e 0,85% respectivamente. Dentre as respostas do grupo 1, destacam-se a dificuldade na organização textual (28,26%), na argumentação (13%) e na introdução ou finalização do texto (10,86%). Com relação as queixas sobre dificuldade de argumentação, pode-se associá-la com a predominância do ensino de redações dissertativo-argumentativas nos anos do Ensino Médio, seguindo o modelo avaliativo do ENEM.

Entre as dificuldades mais relatadas considerando todas as ocorrências, observam-se: a pontuação (11,58%), a organização textual (11,1%), a acentuação (8,5%) e pouco repertório lexical (6,43%), além das menos expressivas dificuldades com ortografia (5,15%) e problemas na argumentação (5,15%).

5. Conclusões e perspectivas futuras de trabalho

O letramento acadêmico para os ingressantes do ProFIS se constitui como um desafio por apresentar outros gêneros textuais e práticas de escrita distintas daquelas que eles tiveram contato até o ensino médio. Considerando que são oriundos do sistema público de ensino – que apresenta diversos problemas em letrar os alunos em práticas efetivas de escrita – a situação se torna mais complexa para muitos deles.

Além da diferença entre os gêneros da escola básica e da universidade, há também um aumento no grau de exigência dos alunos e a necessidade que esses desenvolvam mais a autonomia do seu processo de aprendizagem. Nesse sentido, a habilidade de realizarem um autodiagnóstico das dificuldades de escrita é significativa para que possam ter um entendimento do que é necessário ser aprendido nesse novo percurso educativo.

Os resultados do levantamento evidenciaram que há uma diversidade de dificuldades dentro e fora dos diferentes níveis linguísticos-textuais, se concentrando majoritariamente no nível textual. Contudo, fora desse nível, a pontuação (11,58%) aparece como a dificuldade mais citada entre todas, seguida pela organização textual (11,1%) e a acentuação (8,5%). Além disso, poucas são as ocorrências em que se afirmou não ter nenhuma dificuldade ou não houve resposta (grupo 4 – 2,57%), ou que apresentaram definições pouco específicas (grupo 2 – 11,15%).

A partir desses dados, pretende-se a continuidade deste trabalho por uma comparação das dificuldades relatadas com os desvios realmente existentes nos textos dos alunos. Uma análise inicial de 40 textos, dentro dos mais de 150 estudantes matriculados na disciplina em 2017, mostrou que o emprego equivocado de pontuação é mais frequente do que outros desvios, aspecto esse que foi o mais citado pelos estudantes em suas respostas (11,58%). Porém, ainda é preciso um trabalho mais minucioso e com maior volume de textos. Essa análise e de outros pesquisadores interessados em questões relativas ao letramento acadêmico será possível pela compilação já realizada dos 973 documentos do ano de 2017, parte bastante significativa do trabalho empreendido nesta iniciação.

6. Bibliografia

GERALDI, J. W. (Orgs.) **O texto na sala de aula**. São Paulo: Ática, 2011 (1984).

KOCH, I. G. V.; ELIAS, V. M. **Ler e escrever: estratégias de produção textual**. 2. ed. São Paulo, SP: Contexto, 2011. 220 p., il. ISBN 9788572444231 (broch.).

MENDONÇA, M. Análise Linguística no ensino médio: um novo olhar, um outro objeto. In: BUNZEN, C.; MENDONÇA, M. (Orgs.) **Português no ensino médio e formação do professor**. São Paulo: Parábola, 2006, pp. 199-226.