

VOCABULÁRIO, COMPLEXIDADE TEXTUAL E COMPREENSÃO DE LEITURA EM AMBIENTES DIGITAIS DE ENSINO: UMA INVESTIGAÇÃO INICIAL COM ALUNOS DO ENSINO MÉDIO

VOCABULARY, TEXTUAL COMPLEXITY AND READING COMPREHENSION IN DIGITAL LEARNING ENVIRONMENTS: AN INITIAL INVESTIGATION WITH HIGH SCHOOL STUDENTS

Maria José Bocorny Finatto

Universidade Federal do Rio Grande do Sul
mariafinatto@gmail.com

Monica Stefani

Universidade Federal do Rio Grande do Sul
mokistefani@yahoo.com.br

Aline Evers

Universidade Federal do Rio Grande do Sul
aline.evers@gmail.com

Bianca Franco Pasqualini

Universidade Federal do Rio Grande do Sul
bianca.pasqualini@gmail.com

RESUMO: Neste artigo, relata-se uma investigação inicial que pretendeu qualificar a elaboração e a facilidade de uso de recursos didáticos para Educação a Distância (EAD) na área de Letras/Língua Portuguesa e Leitura. Apresentam-se o planejamento do recurso, a seleção de materiais e as noções teóricas envolvidas e o desenho inicial da atividade, que consistiu em ler e avaliar a complexidade de um conjunto de textos curtos. A experiência revelou sucesso apenas para um pequeno grupo controlado de alunos e insucesso em grande grupo, sem controle. Para a melhoria do recurso criado e da sua implementação didática, indica-se a necessidade de realização de atividades prévias presenciais de estudo com os grupos envolvidos e de avaliação de resultados com os alunos respondentes após a tarefa realizada.

PALAVRAS-CHAVE: ambientes digitais; complexidade textual; recursos de EAD; vocabulário.

ABSTRACT: In this paper, we describe an initial investigation that intended to qualify the elaboration and usability of didactic resources for Distance Learning (DL) in the field of Languages/Portuguese Language and Reading. We present the planning of the resource, the selection of materials and the theoretical notions involved, and the initial design of the activity, which consisted in reading and evaluating the complexity of a set of short texts. The experience was successful only for a small controlled group of students and unsuccessful for the large uncontrolled group. In order to improve the devised resource and implement it didactically, there is the need to perform previous presential learning activities with the involved groups and proceed with the student's evaluation of the results after the task is accomplished.

KEYWORDS: digital environments; textual complexity; distance learning resources;

vocabulary.

1 Introdução

A pesquisa sobre recursos em Educação a Distância (EAD) tem um importante nicho a ocupar no âmbito dos Programas de Graduação e de Pós-Graduação no Brasil. Afinal, recursos em EAD são cada vez mais utilizados para mediar encontros entre professores, estudantes e pesquisadores em ambientes digitais. Nesses ambientes, salientam-se os desafios de se enfrentar um novo formato e de dar conta de um novo tipo de interação, além do bom atendimento das necessidades dos professores e de cativar os interesses de jovens leitores já bastante familiarizados com a rapidez de informações do mundo digital. Muitos recursos e atividades escolares de EAD, ainda que, em tese, cuidadosamente planejados, podem esbarrar em dificuldades de uso, proporcionadas justamente pelas novas dinamicidades de uma leitura digital e pelas demandas de estudantes que já convivem, dos seus modos e muito intensamente, com o acesso à Internet. Nesse contexto, a boa interação do estudante com textos escritos dependerá da sua disposição para com eles, considerando-se que o processo de compreensão de leitura exige toda a mobilização de uma série de processos cognitivos complexos, como apontam estudos recentes sobre complexidade textual e processamento da leitura (por exemplo, MCNAMARA, 2007; PASQUALINI, 2012; FINATTO et al., 2015). Além disso, dada a velocidade de informação e o hábito de relacionamento dos jovens internautas com textos curtos apoiados em muitas imagens, especialmente como os que vemos em redes sociais, encarar um texto longo, tal como um exercício de leitura dirigida, pode tornar-se algo bastante enfadonho. Esse enfado dos estudantes, muitas vezes associado a um sentimento de não adequação das atividades que se apresentam a eles, pode gerar resultados de aprendizagem frustrantes. A pouca adesão dos alunos às atividades tende inclusive a desmotivar os professores ao uso de recursos de EAD.

Assim, uma boa interação com tarefas *online* dependerá do nível de atratividade e de facilidade de uso dos recursos empregados. Além disso, deve-se considerar que o processo de compreensão de leitura, também em ambiente digital, não é algo trivial, pois exige, da mesma forma, a mobilização de vários processos cognitivos complexos (cf. FULGÊNCIO; LIBERATO, 1998, p. 13-32). Por isso, construir um conhecimento mais aprofundado sobre alternativas didáticas, sobre como se conseguir boa interação entre alunos e recursos em ambientes EAD, sobre engajamentos em tarefas EAD e sobre leitura em ambientes digitais torna-se fundamental.

Nesse sentido, trazemos aqui um experimento inicial que envolveu a modelagem de uma tarefa implicando compreensão de leitura em ambiente digital e reconhecimento de perfis de vocabulário em diferentes tipologias de textos escritos por parte de alunos de uma escola pública federal do Ensino Médio. Criamos um exercício-piloto que foi conduzido, a modo de teste inicial, com alguns estudantes do Ensino Médio do Colégio de Aplicação da UFRGS¹ no ano de 2015. O objetivo foi avaliar as melhores opções de: a) desenho do exercício, especialmente a escolha de textos e formulação linguística dos

1 Trabalho em cooperação com os professores Aduino Locatelli Taufer e Daniela Fávero Netto e seus estudantes-bolsistas de pré-iniciação científica do Colégio de Aplicação da UFRGS no âmbito do Programa PICMEL – Edital FAPERGS 015/2013, Processo 0349-2551/14-0.

comandos da tarefa; b) estratégias de implementação, considerando-se atrair o interesse e promover o comprometimento dos estudantes envolvidos para com a aprendizagem em foco. Nele buscamos verificar a percepção dos estudantes sobre o quão complexo um texto poderia ser para eles mesmos e para outras pessoas, especialmente no que se refere ao vocabulário empregado.

Dada a temática do exercício, torna-se importante situar alguns conceitos-chave nele envolvidos. Afinal, o que é entendido como complexidade (de um texto) ou dificuldade (de compreender esse texto)? Quem se tem em mente como leitor? O que se entende por leitura? As perguntas são inúmeras e suas respostas demandariam um espaço que não temos aqui. Entretanto, cabe dizer, *grosso modo*, que, no âmbito dos Estudos da Linguagem e da Linguística Aplicada, boa parte das pesquisas sobre complexidade realça um ou outro elemento: o texto em si ou o seu leitor, suas especificidades, competências e demandas. Em geral, ou se trata do texto como uma estrutura linguística de tessitura, ou se enfatizam as etapas do processamento cognitivo e psicolinguístico da leitura de determinados sujeitos (crianças, jovens, adultos, falantes não nativos de uma língua, etc.).

Atualmente, avançamos rumo às concepções diferenciadoras entre a capacidade de ler e de entender um texto escrito e a capacidade de *letramento* (TFOUNI, 1995) de um indivíduo. A título de explicação, destacamos apenas o entendimento de que a “leitura configura-se como atividade essencialmente significativa, dirigida a um objetivo, dependendo de conhecimentos anteriores e encaminhada pelas expectativas do leitor” (PEREIRA, 2009, p. 22). Quando se toca nos temas complexidade textual e leitura, é preciso esclarecer vários pontos: o que é entendido como complexidade ou dificuldade e para quem, o que se entende por leitura, entre outros aspectos. Aqui entram em cena conceitos teóricos altamente complexos, alguns muito imbricados, tais como o de *compreensão leitora*, *proficiência* ou *competência de leitura*, *letramento*, *estratégias de compreensão de leitura*, *legibilidade*, *inteligibilidade*, *acessibilidade* e *alfabetismo*. A seguir, exploramos, muito sinteticamente, alguns desses conceitos.

Legibilidade, segundo Rudolf Flesch (1949), engloba elementos gráficos do texto (como a presença de figuras, tabelas, quadros, tamanho da letra, formato do texto) que auxiliam o leitor. Também inclui elementos linguísticos da superfície do texto, como extensão das frases, número de frases por parágrafo, número de sílabas médio das palavras, entre outros aspectos. A *acessibilidade* de um texto seria semelhante à sua *legibilidade*, mas com facilidades visuais para pessoas com deficiências, por exemplo. A *acessibilidade* implicaria uma condição do texto que o tornasse compreensível também por parte de leitores com algum tipo de limitação, como seu nível de escolaridade, por exemplo. *Inteligibilidade*, por outro lado, diz respeito a características que tornam um texto mais ou menos compreensível. Assim, um texto pode ser *legível*, de acordo com os critérios de Flesch (puramente superficiais), e *acessível* para pessoas com alguma deficiência ou limitação, mas, ainda assim, ser pouco *inteligível/compreensível*, pois nesse fator entram elementos da *natureza* do texto, suas características linguísticas (uso de formas gramaticais e itens lexicais frequentes, ordenação gramatical, uso de elipses, e assim por diante), e também aspectos como o *público* a que o texto se destina. Portanto, para estimar a *inteligibilidade* de um texto, precisaríamos ponderar, inicialmente, sobre qual é ou seria a *competência* de leitura do seu leitor-alvo. Tal competência está atrelada à sua experiência de leitura, seu nível de letramento e familiaridade com a natureza do

texto. Já a natureza do texto vai desde o assunto abordado até o tipo e gênero discursivo envolvido (reportagem de jornal, romance, instruções de montagem de um móvel, artigo científico, e assim por diante). Nesse âmbito, o nível de complexidade de um texto pode envolver elementos gráficos (legibilidade, acessibilidade), elementos linguísticos (inteligibilidade) e elementos extralinguísticos (natureza, assunto, tipo, gênero) em relação ao leitor. Naturalmente, há uma miríade de aspectos envolvidos entre essas noções aqui recortadas conforme vemos, por exemplo, em Kleiman (2003).

Tentando quantificar a determinação da competência do leitor, há medidas e testes, como o teste de *letramento* adotado pelo Indicador de Alfabetismo Funcional (INAF) (RIBEIRO, 2010). Nele, o leitor é classificado em um de quatro níveis de letramento: analfabetismo (incapacidade de realizar tarefas simples de leitura), alfabetismo em nível rudimentar (capacidade de localizar informações em textos curtos e familiares), alfabetismo em nível básico (capacidade de ler e compreender textos de média extensão e localizar informações fazendo algumas inferências) e alfabetismo em nível pleno (capacidade de ler textos longos e relacionar suas partes, comparar informações e fazer inferências e sínteses). De acordo com resultados recentes do INAF, o analfabetismo funcional no Brasil diminuiu de 39 para 27% entre 2001 e 2010, e o índice de alfabetismo funcional aumentou de 60 para 73% no mesmo período. Um dado alarmante foi o de que 59% dos indivíduos que têm da 5ª à 8ª série do Ensino Fundamental e 57% dos indivíduos que cursaram o Ensino Médio foram considerados alfabetizados apenas em nível básico. O índice de letramento do Ensino Superior, conforme Ribeiro (2010), é composto por 62% de indivíduos em nível pleno e 34% em nível básico. Esses dados seriam também relevantes para o que poderia ser a complexidade de um texto para diferentes perfis de leitores.

Por opção metodológica, destacamos, no nosso experimento inicial, apenas o *Índice Flesch* (IF), uma medida superficial de legibilidade ou de complexidade textual, a qual é de natureza estatística. Essa medida é capaz de estimar o quanto um determinado texto seria mais ou menos complexo (ou legível) para um dado tipo de leitor. A origem do IF remonta ao ano de 1948 e aos trabalhos de Flesch (FLESCHE, 1949) sobre o *Plain English* (inglês simples). Seu índice correlaciona o número total de palavras de um texto, o número de sílabas por palavra e o número de palavras por sentença e sua adequação para determinados públicos de leitores conforme sua escolaridade. No Brasil, pesquisadores (MARTINS et al., 1996) do Instituto de Ciências Matemáticas e da Computação da Universidade de São Paulo (ICMC-USP) adaptaram o IF do inglês para o português brasileiro e adequaram o nível de escolaridade ao cenário de nosso sistema de ensino (ver Tabela 1).

A fórmula do índice Flesch é a seguinte: $\text{Score} = 206.835 - (1.015 \times \text{TMS}) - (84.6 \times \text{MSP})$, onde:

Score = posição numa escala de 0 (difícil) a 100 (fácil), com até 30 = muito difícil e acima de 70 = adequado para todos os públicos.

TMS = tamanho médio das sentenças (o número de palavras dividido pelo número de sentenças).

MSP = número médio de sílabas por palavras (o número de sílabas dividido pelo número de palavras).

Tabela 1: Escores do Índice Flesch, níveis de complexidade por faixa de escolaridade.

ESCORE	NÍVEL DE COMPLEXIDADE	INDICAÇÃO
100-75	Muito fácil	Adequado para 1ª a 4ª séries do ensino fundamental
75-50	Fácil	Adequado para 5ª a 8ª séries do ensino fundamental
50-25	Difícil	Adequado para 9ª a 11ª séries do ensino médio
25-0	Muito difícil	Adequado para ensino superior

2 Seleção de textos para o exercício

A seleção inicial dos textos para o nosso exercício de compreensão de leitura em EAD procurou abarcar diferentes perfis discursivos – textos jornalísticos, técnicos e literários. Partimos de materiais já utilizados (FINATTO et al., 2015), reunidos em um *corpus*, tendo como apoio técnicas e conceitos derivados da Linguística de Corpus, tal como proposta no Brasil por Berber Sardinha (2004), e aplicamos recursos de léxico-estatística ao conjunto de textos reunido (seguindo indicações de Biderman, 1978, 1996). Conforme Finatto et al. (2014):

a léxico-estatística é uma disciplina da Linguística que trata de aspectos quantitativos do uso da linguagem e do sistema linguístico, com apoio de procedimentos estatísticos. Esses procedimentos tornam-se importantes naquelas situações em que se trata de compreender a língua em funcionamento, ou seja, em textos (FINATTO et al., 2014, p. 54).

Privilegiamos textos, em tese, mais familiares e/ou interessantes para os nossos jovens leitores, estudantes do Ensino Médio de uma escola pública, mas também colocamos alguns textos (em tese) potencialmente difíceis. Feito um reconhecimento inicial sobre o perfil dos alunos da escola e um exame inicial dos textos, definimos um conjunto de 10 unidades de textos relativamente curtos para utilizar no nosso exercício. Conforme já mencionado, preferimos textos curtos ou segmentos de texto que mantivessem reconhecíveis as características mais salientes de um todo, categoria textual ou tipo discursivo. Também levamos em conta a necessidade de ter um conjunto variado de textos, que empregassem diferentes tipos de vocabulário.

Aplicamos aos textos selecionados recursos de léxico-estatística (seguindo indicações de BIDERMAN [1978, 1996]), conforme Finatto et al. (2014). Para estimar, preliminarmente, o provável nível de complexidade dos textos escolhidos, verificamos, em cada texto, apenas o seu IF, com apoio do novo sistema Coh-Matrix-Port² (CUNHA, 2015), em acesso livre pelo NILC/ICMC-USP³. Esse *software* é uma adaptação para o português brasileiro das métricas do sistema Coh-Matrix desenvolvidas para o inglês pela Universidade de Memphis, dos Estados Unidos. Esse sistema automático oferece várias

2 Disponível em: <<http://143.107.183.175:22680>>, na versão denominada *Dementia*.

3 Núcleo de pesquisa brasileiro da área de Linguística Computacional/PLN. Mais informações em: <<http://www.nilc.icmc.usp.br/nilc/index.php>>.

contagens para um dado texto nele inserido (48 métricas no total, como, por exemplo, incidência de pronomes pessoais, sintagmas nominais, palavras por sentença, negações, parágrafos, entre outras).

A Tabela 2 sintetiza os dados de cada um dos 10 textos apenas quanto ao IF. Observe-se que adaptamos a abrangência dos escores: a faixa dos considerados *fáceis* e *muito fáceis* está em uma só categoria (“fácil”), o mesmo valendo para os textos considerados *difíceis* e *muito difíceis* (reunidos na categoria “difícil”).

Tabela 2: Dados dos textos selecionados, escores e índice Flesch pelo sistema Coh-Metrix-Port.

DESCRIÇÃO	TEXTO	ESCORE	NÍVEL DE COMPLEXIDADE
Trecho do conto Majestic Hotel, de Sergio Faraco. In: <i>Dançar tango em Porto Alegre</i> . Porto Alegre: LP&M, 2006, p. 83.	1	Índice Flesch = 75	Fácil
Resumo de artigo científico sobre corrosão de refratários utilizados na siderurgia. Parte III: Caracterização de refratários comerciais. In: Revista Cerâmica 59 (2013) p. 39.	2	Índice Flesch = 0	Difícil
Trecho do romance <i>Quincas Borba</i> , de Machado de Assis, em domínio público.	3	Índice Flesch = 71	Fácil
Trecho de um manual de instruções de um ventilador.	4	Índice Flesch = 49	Difícil
Artigo da lei ordinária 5.172 de 25/10/1966, que trata do Sistema Tributário Nacional e estipula as normas gerais de direito tributário (Código Tributário Nacional), p. 318.	5	Índice Flesch = 45,5	Difícil
Notícia de jornal, <i>Correio do Povo</i> , 28/08/2014.	6	Índice Flesch = 52	Fácil
Trecho da bula do medicamento maleato de enalapril.	7	Índice Flesch = 42	Difícil
Artigo 132 do Código Penal Brasileiro, p. 81-82.	8	Índice Flesch = 67	Fácil
Trecho do livro	9	Índice Flesch = 60	Fácil

didático <i>Fundamentos da Química</i> , de Ricardo Feltre (São Paulo: Ed. Moderna, 1996, p. 240).			
Trecho da crônica de Juremir Machado da Silva, <i>Produto à venda</i> (<i>Correio do Povo</i> , 28/08/2014).	10	Índice Flesch =62,5	Fácil

3 Elaboração do exercício de leitura

O nosso exercício de compreensão de leitura, hospedado num *site* institucional⁴, procurou relacionar o perfil do estudante, seus hábitos de leitura e a sua percepção sobre o quão difícil a compreensão de um determinado texto poderia ser para o próprio respondente e para diferentes pessoas. Reduzimos as opções de resposta sobre a complexidade dos textos para não gerar sobrecarga de informação, pois a tarefa inicialmente modelada envolvia, além de ler 10 textos e responder um conjunto de questões sobre cada um deles, completar um questionário preliminar relativamente longo.

Como parte preliminar do exercício (parte 1), apresentamos 10 questões objetivas sobre o perfil socioeconômico do leitor, abrangendo faixa etária, sexo, escolaridade, renda familiar mensal e hábitos/preferências de leitura. Utilizamos itens comparáveis aos dados oficiais do Índice de Analfabetismo Funcional e de outros órgãos governamentais (p. ex., do IBGE).

Após essas perguntas preliminares, na parte 2 seguem os textos selecionados e as respectivas questões sobre eles. Para cada texto (ou trecho, quando fosse o caso), foram colocadas três perguntas sobre: (1) a dificuldade do texto percebida pelo leitor (com as opções de resposta sendo “fácil”, “nem fácil nem difícil” – ou seja, neutralidade – e “difícil”, seguindo o agrupamento dos escores mencionado anteriormente; (2) uma provável adequação do texto lido para pessoas com ensino superior, médio ou fundamental; (3) o gênero a que pertenceria o texto lido (parte 3).

Esse questionário, frisando-se isso para o respondente, não apresentava alternativas “certas” ou “erradas”, apenas indagava o que o leitor, subjetivamente, achava. No entanto, para viabilizar a colocação de um gabarito para nossa ponderação e poder dar algum retorno para os respondentes, a ideia foi usar o Índice Flesch dos textos selecionados como um parâmetro de sua avaliação de complexidade.

Assim, apresentar o índice Flesch aos respondentes, resgatando sua natureza e origem, foi considerada uma etapa de aprendizagem interessante para os estudantes após a aplicação do exercício. Desse modo, a atividade serviria também como pano de fundo, para que os alunos tivessem um contato inicial com técnicas de contagem de palavras, observação de vocabulário em textos e com a Linguística de Corpus e a Estatística Lexical. Esse índice Flesch, ressaltamos, não é normativo, tampouco absoluto, uma vez que serve apenas como um ponto de referência inicial, entre tantos possíveis,

4 Disponível em: <<http://www.ufrgs.br/textecc/traducao/teorias/pesquisa.php>>.

para a elaboração de um “gabarito” de complexidade, bem como para o tema da complexidade textual.

Esse exercício (veja amostra, Figura 1) funcionou como um retorno “autoavaliativo”: visualiza-se o IF de cada texto, na tela, após o envio da tarefa completa pelo *site*, apresentando o IF aos respondentes, resgatando sua natureza e origem, constituindo-se uma válida forma de aprendizagem.

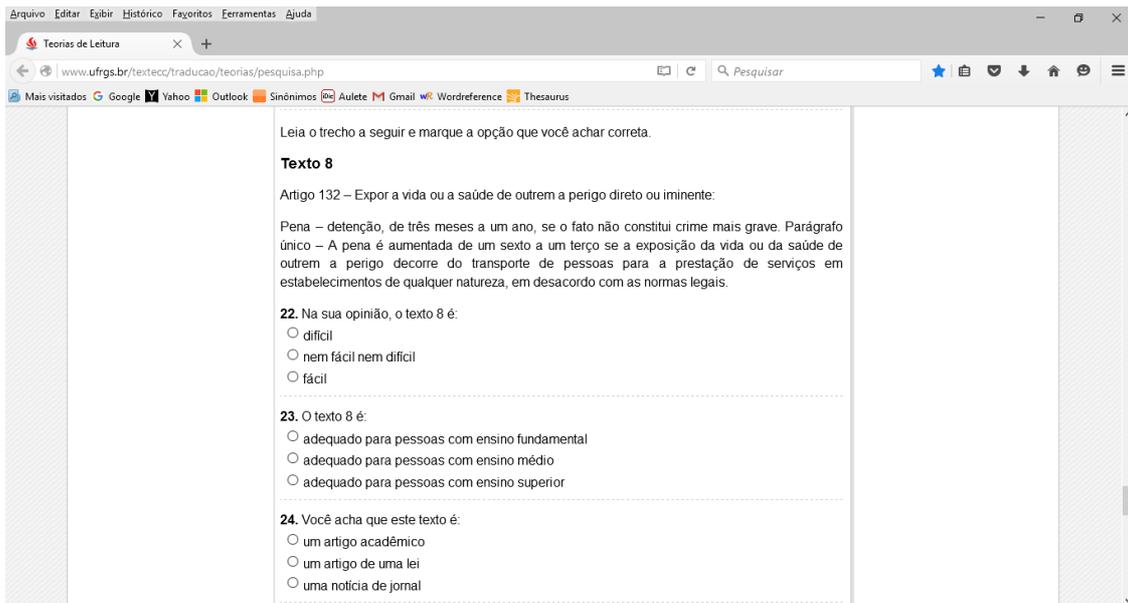


Figura 1: Exercício de Leitura, versão preliminar de 2015.

Fonte: <www.ufrgs.br/textecc/traducao/teorias/pesquisa.php> (nova versão em março 2016).

4 Realização do exercício, resultados e discussão

Antes de realizar o exercício com os estudantes, primeiro fizemos uma reunião com os professores responsáveis pelas disciplinas de Língua Portuguesa e Literatura, nossos parceiros. Apresentamos a eles a natureza da pesquisa e o exercício-piloto. Conversamos sobre dúvidas relacionadas às variáveis da primeira parte da pesquisa (parte 1, perguntas não numeradas) e às demais opções relacionadas aos tipos de textos escolhidos (partes 2 e 3 do exercício, perguntas 1 a 30). Apresentamos o tema da complexidade textual, o IF, as técnicas de Linguística de Corpus, o pré-processamento dos textos e as ferramentas envolvidas.

Em outro encontro, apresentamos, ao vivo em sala de aula, com os professores, o exercício para um grupo de apenas seis alunos do último ano do Ensino Médio. Esses alunos eram bolsistas de pesquisa de língua portuguesa dos professores em uma pesquisa de Língua Portuguesa do programa PICMEL-FAPERGS. Por isso, já tinham tido contato com conceitos básicos sobre Linguística de Corpus e Estatística Lexical. Foram apresentados a natureza da pesquisa e o *site* – em construção – para seu acesso. Tanto o questionário prévio quanto os conceitos que o embasam (gênero discursivo, complexidade textual, estatística de vocabulário e índice Flesch) foram apresentados aos

alunos dessa pequena turma com a supervisão de seus professores.

A participação de cada um na atividade, a ser feita no *site*, seria anônima e não obrigatória. A tarefa seria realizada após a aula, via Internet. Cada um dos seis alunos para quem a atividade foi apresentada poderia ainda convidar outros colegas, amigos ou mesmo parentes seus para participarem do exercício, solicitando-se que eles mesmos explicassem a seus convidados do que se tratava. Assim, o pequeno grupo de seis alunos, conforme imaginamos, funcionaria como disseminador do trabalho e incentivador dos colegas e até de membros de suas famílias para a atividade. Entretanto, as respostas efetivamente válidas obtidas de seus convidados foram apenas 13, conforme adiante detalhado.

Quanto ao funcionamento do questionário de leitura no *site*, o respondente precisaria completar todas as perguntas antes de poder avançar para verificar seu desempenho. Assim, se, por algum motivo durante a realização da pesquisa, o respondente pula uma pergunta, surge uma mensagem de que há perguntas a serem respondidas. Quando o respondente finaliza a pesquisa, recebe a indicação de quantos “acertos” fez (com a porcentagem, tendo-se o IF do texto como parâmetro para a dificuldade do texto) e as respostas de todas as perguntas. Nesse momento, o respondente é apresentado ao IF de cada texto.

Antes de avançarmos para a amostragem de respostas aos textos, vejamos os dados socioeconômicos dos seis alunos disseminadores da tarefa na Tabela 3 a seguir. Infelizmente, apesar de seus esforços ou da nossa estratégia para multiplicar a adesão de mais colegas e até de seus familiares, as respostas total e efetivamente válidas e as mais acertadas ou homogêneas, entre inúmeras recebidas, muitas aleatoriamente respondidas, foram apenas as desses mesmos seis alunos.

Tabela 3: Dados socioeconômicos e de hábitos de leitura dos seis alunos disseminadores participantes do exercício.

Faixa etária Menos de 17 anos: 5 Entre 17 e 21 anos: 3	Sexo Feminino: 5 Masculino: 1
Trajetória de estudos No ensino público: 3 Parte no ensino público e parte no ensino privado: 3	
Renda familiar 3 a 5 salários mínimos: 3 5 a 7 salários mínimos: 2 mais de 10 salários mínimos: 1	
Hábito de leitura Sim: 2 Não: 4	O que mais gosta de ler Livros: 3 Gibis e animes: 2
O que menos gosta de ler Revistas: 1 Livros: 1 Gibis: 2 Jornais: 2	O que mais lê, independentemente de gostar Livros: 1 Jornais: 3 Revistas: 2
Quantos livros lê por ano Mais de 4: 1 Entre 3 e 4: 3 Entre 2 e 3: 2	

No grupo dos 13 respondentes “convidados”, a maioria é jovem e está cursando Ensino Médio ou Superior, possui renda familiar de mais de 3 salários mínimos e manifestou ter o hábito de ler, embora deva ser ressaltado que as quatro respostas negativas nesse item vêm de alunos do Ensino Médio. Nas preferências de leitura, os livros são os mais lidos, tanto por gosto quanto independentemente de gosto, mas esse dado se altera quando os alunos do Ensino Médio são considerados: eles gostam mais de ler livros, mas leem mais jornal, independentemente de gostar; o que menos gostam de ler são os gibis e os jornais. A frequência de leitura desse grupo ficou em mais de 3 livros por ano. Em termos de desempenho, o número de acertos desse grupo de 13 pessoas teve uma média de 17 dos 30 possíveis. No entanto, para os propósitos deste artigo, vale destacar os resultados dos seis alunos disseminadores, cuja média ficou em 16 acertos. No pequeno grupo, chama a atenção que um aluno – menos de 17 anos – líder no *ranking* de acertos – 20 – respondeu não ter o hábito de ler, mas possuía a maior renda familiar entre todos – mais de 10 salários mínimos – tendo experiência de ensino tanto na rede privada quanto na pública.

Na parte dos textos das respectivas questões, estávamos interessados em saber: **(1)** o quanto cada um dos textos poderia ser **mais ou menos fácil para o respondente** compreender; **(2) e (3)** por qual(is) tipo(s) de LEITOR (por escolaridade) o respondente achava que o texto seria **melhor/pior compreendido**; **(3)** a que categoria de documento ou a que **tipo de texto** o trecho lido pertenceria. Dos 10 textos, apenas 2 apresentaram equívocos nas respostas quanto ao gênero: o texto 9 (excerto de um livro didático que recebeu uma resposta no exercício identificando-o como um resumo e foi considerado “nem fácil nem difícil” pela maioria – 4 respostas – e como “fácil” – pelas 2 restantes); e o texto 10 (50% dos respondentes o marcaram acertadamente como o trecho de uma crônica, mas os outros 50% o identificaram como um resumo, sendo que 4 respostas o consideraram “nem fácil nem difícil”, e as 2 respostas restantes como “fácil”). As diferenças nas respostas ficaram basicamente na percepção de dificuldade/facilidade dos textos. Por exemplo, o texto 4 (manual de instruções), considerado difícil pelo IF, foi percebido pelos alunos como “fácil” (4 das 6 respostas), o mesmo ocorrendo com o texto 7 (4 respostas considerando-o “fácil” e 2 como “nem fácil nem difícil”). O texto 5 (trecho de uma lei) recebeu unanimidade quanto à dificuldade, enquanto o texto 2 (um resumo de artigo científico) ficou com 5 respostas como “difícil” e apenas 1 como “nem fácil nem difícil”. O texto 8, que também era uma lei, obteve 4 respostas para “nem fácil nem difícil” e 2 para “fácil”. Quanto aos textos literários, vê-se que eles tiveram uma percepção de neutralidade: o texto 1 (conto de Sergio Faraco) obteve 3 respostas para “fácil” e 3 para “nem fácil nem difícil”, e o texto 3 (trecho de Machado de Assis) ganhou 5 respostas para “nem fácil nem difícil” e 1 para “fácil”. O texto 6 (notícia da área econômica) também merece destaque nesse aspecto de neutralidade: 5 respostas para “nem fácil nem difícil” e apenas 1 para “fácil” (talvez a área da notícia – economia – tenha contribuído para essa percepção mais neutra).

Fora esses resultados, o que instigou o pequeno grupo dos alunos disseminadores após a realização do exercício foi a ideia de seguir estudando Linguística de Corpus e o tema da complexidade textual com o apoio da Matemática e de ferramentas computacionais da Internet. Pelo que apuramos, todos consideraram que foram capazes de motivar seus colegas e conhecidos para a atividade. Também relataram que o desenho da atividade e a escolha dos textos parecia-lhes compatível com seus interesses e

expectativas de aprendizagem. Também os professores desse pequeno grupo manifestaram surpresa pelo pouco comprometimento dos demais convidados a quem a atividade foi apresentada. Na opinião desses professores, o trabalho autônomo, sem identificação, fora do horário de escola, totalmente feito em um *site*, e sem supervisão ou “endosso” de um professor, teria facilitado tamanho insucesso. Para nós, numa visão pessimista, não conseguimos motivar o grande grupo de alunos nem para os textos, tampouco para o tema da complexidade.

5 Considerações finais

A proposta da atividade de compreensão de leitura e de percepção de complexidade de textos foi realmente produtiva apenas para um pequeno grupo de alunos, um grupo com condições especiais de ensino-aprendizagem, proporcionadas por seu engajamento em um projeto de pré-iniciação científica. Infelizmente, a disseminação para colegas, familiares e conhecidos não foi bem-sucedida. Embora tenhamos tido um alto número de tarefas preenchidas (cerca de 200), a quase totalidade pareceu ter sido realizada aleatoriamente, já desde a parte preliminar, ou repetidamente enviada por uma mesma pessoa. Isso apontou que a atividade, aplicada em grande escala, na modalidade “cada um por si”, precisaria contar com uma melhor preparação prévia, melhores insumos pré-tarefa e melhor desenho de recursos técnicos. Essa preparação, conforme imaginamos, deverá ser presencial e relativamente extensa sobre os conceitos envolvidos no tema da complexidade textual. Por isso, a versão atualizada da atividade, disponível em versão melhorada, para nova testagem, oferece agora: a) um material prévio sobre o tema da complexidade textual, conceitos relacionados e modos de acesso a ferramentas para processamento de textos – especialmente o Coh-Metrix-Port – e trabalhos relacionados; b) identificação de hora e data de envio das respostas para consulta do administrador/supervisor da tarefa; c) dados **prévios** sobre o IF de cada texto e indicações de materiais para estudo de noções relacionadas a ele.

Ainda assim, um ponto muito positivo do experimento, na percepção de quem o planejou e conduziu, foi o processo de seleção dos textos e o uso de recursos estatísticos e computacionais para sua análise prévia. Além disso, a percepção das características do vocabulário dos textos com apoio de princípios matemáticos e de ferramentas informatizadas foi um ponto de motivação para os alunos disseminadores e para seus professores, os quais não tinham familiaridade com enfoques estatísticos de vocabulário. Assim, acreditamos que um trabalho posterior à atividade que desenhamos, envolvendo o processamento de cada texto completo e/ou do trecho em foco, feito com os alunos em uma nova atividade de EAD e presencial, poderia ser um ganho importante para conquistar o engajamento dos alunos nessa experiência de aprendizagem que envolve um tema tão importante. Afinal, o tema da complexidade de um texto – como uma medida sempre relativa – e o uso de ferramentas na Internet pareceu ser o grande motivador dos envolvidos nesse primeiro experimento. Para uma próxima etapa de utilização do recurso, já melhorado, a ser conduzida em 2017 com estudantes de graduação de Letras, pretendemos investir nessa via.

Agradecimentos

Ao programa PICMEL da FAPERGS, ao CNPq, à CAPES, ao Programa STIC-AmSud da CAPES (Projeto RITA), ao Programa CAPES-COFECUB (Projeto Cameleon), ao PPG-LETRAS-UFRGS e, em especial, à SEAD-UFRGS, pela concessão de bolsa de pesquisa em nível de pós-graduação sobre a melhoria da EAD na UFRGS.

Referências

BERBER SARDINHA, T. *Linguística de corpus*. São Paulo: Manole, 2004.

BIDERMAN, M. T. *Teoria Linguística: Linguística Quantitativa e Computacional*. Rio de Janeiro: Livros Técnicos e Científicos, 1978.

BIDERMAN, M. T. Léxico e Vocabulário Fundamental. *Alfa*, São Paulo, v. 40, p. 27-46, 1996.

CUNHA, A. L. V. da. *Coh-Matrix-Dementia: análise automática de distúrbios de linguagem nas demências utilizando Processamento de Línguas Naturais*, 2015. 151 p. Dissertação (Mestrado em Ciências da Computação e Matemática Computacional). Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo, 2015.

FINATTO, M. J. B.; EVERS, A.; PASQUALINI, B.; KUHN, Tanara Z.; PEREIRA, A. M. Vocabulário controlado e redação de definições em dicionários de português para estrangeiros: ensaios para uma léxico-estatística textual. *Revista Trama* (UNIOESTE Online), v. 10, p. 53-68, 2014.

FINATTO, M. J. B.; STEFANI, M.; PASQUALINI, B. F.; CIULLA, A.; EVERS, A.; SORTICA, M. *Leitura: um guia sobre teoria(s) e prática(s)*. Porto Alegre: UFRGS, 2015.

FLESCHE, R. *The art of readable writing*. New York: Harper & Row, 1949.

FULGÊNCIO, L.; LIBERATO, Y. *Como facilitar a leitura*. São Paulo: Contexto, 1998.

KLEIMAN, A. Avaliando a compreensão: letramento e discursividade nos testes de leitura. In: RIBEIRO, V. M. (Org.). *Letramento no Brasil*. São Paulo: Global Editora, 2003, p. 209-225.

MARTINS, T. B. F., GHIRALDELO, C. M., NUNES, M. G., OLIVEIRA JÚNIOR, O. N. *Readability formulas applied to textbooks in Brazilian-Portuguese*. Technical Report, São Carlos: ICMC/USP, 1996.

MCNAMARA, D. *Reading Comprehension: Theories, Interventions, Technologies*. New Jersey: Lawrence Erlbaum Publishers, 2007.

PASQUALINI, B. *Leitura, tradução e medidas de complexidade textual em contos da literatura para leitores com nível de letramento básico*. 2012. 155f. Dissertação (Mestrado

em Letras). Programa de Pós-graduação em Letras, Universidade Federal do Rio Grande do Sul, 2012.

PEREIRA, V. W. Predição leitora: procedimentos e desempenhos em ambiente virtual e ambiente não virtual. *Letras de Hoje*, Porto Alegre, v. 3, n. 44, p. 22-27, jul./set. 2009.

RIBEIRO, Vera M. Matriz de referência para a medição do alfabetismo nos domínios do letramento e do numeramento. *Estudos em Avaliação Educacional*, 21, n. 45, jan./abr. 2010, p. 147-168.

TFOUNI, L. V. *Letramento e alfabetização*. São Paulo: Cortez, 1995.

Recebido em 11 de julho de 2016.
Aprovado em 08 de setembro de 2016.