

Análise de inteligibilidade de material didático para Educação de Jovens e Adultos (EJA): Compreensão leitora avaliada pelo Coh-Matrix-Port

Nély Silva da Motta Mesquita

Pontifícia Universidade Católica do Rio de Janeiro

Resumo

O objetivo desse artigo é apresentar os resultados da pesquisa sobre inteligibilidade textual da Coleção Cadernos Didáticos de EJA, destinada a jovens e adultos egressos do ensino regular. A avaliação dos parâmetros de inteligibilidade foi efetuada com o auxílio de uma ferramenta computacional o Coh-Matrix-Port. Nesse artigo iremos detalhar as métricas selecionadas da ferramenta e apresentar os resultados obtidos com as submissões dos textos ao Coh-Matrix-Port comparando-os com os índices Flesch referentes aos mesmos textos, como forma de orientar professores a selecionar o material adequado aos seus alunos.

Palavras Chave: leitura; métricas de legibilidade; educação de jovens e adultos; gêneros textuais.

Abstract

The aim of this article is to present the results of research about the textual readability of the EJA Didactic Notebook Collection for young and adult learners who left regular education. The evaluation of the parameters of readability was conducted with the help of a computational tool, the Coh-Matrix-Port. In this article we explain in detail the selected metrics of the tool and we present the results obtained when the texts were submitted to the Coh-Matrix-Port. We compare these results with the Flesch results for the same texts, as a way of instructing teachers on how to select adequate material for their students.

Keywords: reading; readability metrics, education of teens and adults, genres

INTRODUÇÃO

Nesse artigo apresentamos os resultados de pesquisa realizada para dissertação de mestrado na área da Psicolinguística, em que se investigou a inteligibilidade de material didático destinado à Educação de Jovens e Adultos (EJA) com o auxílio de ferramenta computacional, o Coh-Matrix-Port, desenvolvido pelo Projeto PorSimples, nome pelo qual ficou conhecido o projeto que envolve a Simplificação Textual do Português para Inclusão e Acessibilidade Digital (Aluísio et al., 2008).

O foco da pesquisa foi avaliar o custo para o processamento da leitura de textos contidos na Coleção Cadernos Didáticos de EJA, representado pelos recursos linguísticos que foram submetidos ao Coh-Metrix-Port e cujos resultados foram comparados aos índices Flesch obtidos, e que, apesar de estarem incluídos na ferramenta, é uma fórmula que avalia superficialmente a inteligibilidade do texto. Essa fórmula foi adaptada para o português em 1996 (Martins et al.) do *Flesch Reading Ease* enquanto o Coh-Metrix-Port foi adaptado ao português do Coh-Metrix (Universidade de Memphis) em 2008 pelos pesquisadores da USP/São Carlos (Aluísio et al.). Resolvemos comparar os resultados porque o Índice Flesch considera apenas o número de palavras em sentenças e o número de letras ou sílabas por palavras, enquanto que os outros índices incorporados pelo Coh-Metrix-Port passam de 40, como por exemplo, Incidência de palavras funcionais, Incidência de palavras de conteúdo, Ambiguidade de substantivos, Sobreposição de palavras de conteúdo, Referência anafórica, entre outros.

O objetivo desse artigo é apresentar esses resultados de modo a facilitar a seleção de material didático adequado ao nível de compreensão do aluno. Essa análise poderá ser feita pelo professor, desde que seja ancorada nesses parâmetros psicolinguísticos apresentados na pesquisa.

INTELIGIBILIDADE/LEGIBILIDADE/LEITURABILIDADE

O termo inteligibilidade está relacionado ao texto e às suas especificidades que podem ser determinantes para que a leitura seja facilitada, como o uso de palavras frequentes e de estruturas sintáticas menos complexas (Aluisio & Scarton, 2010).

A inteligibilidade textual é construída pelo leitor em sua interação com o texto e seu autor, buscando a compreensão leitora através do processamento de material linguístico: léxico, relações sintáticas marcadas ou não por conjunções, sinais de pontuação etc. A compreensão leitora será facilitada se as relações entre as partes do texto estiverem marcadas de forma adequada com conjunções, advérbios de sequência, elementos de correferência (como pronomes, elipses, entre outros).

A legibilidade está relacionada aos aspectos gráficos do texto, e aí entram fatores como ilustrações, diagramas, fotografias, formato e cor das letras, que contribuem na construção do sentido.

A leitura de um texto é sua capacidade de ser compreendido pelo leitor, não só por sua acessibilidade em termos de material linguístico, como também pela facilidade que o leitor encontra de interagir com a informação disponibilizada por esse texto (McNamara et al., 2009).

Esses conceitos são importantes, de forma a estabelecer os parâmetros para avaliação do grau de dificuldade de leitura que um texto pode apresentar. Em nossa pesquisa estivemos direcionados na inteligibilidade de um material didático para EJA e, para aferi-la, comparamos os índices obtidos com a submissão dos textos ao Coh-Matrix-Port.

MATERIAL ANALISADO

O material que foi analisado é uma coleção composta de 13 cadernos de textos de gêneros variados, cada qual contendo uma média de 25 textos a respeito do tema trabalho. Cada texto tem uma orientação dos Cadernos do Professor para que possa ser abordado pelas disciplinas diferentes: Português, Matemática, Ciências etc., e de acordo com o segmento de Ensino Fundamental que o aluno está cursando: 1º ou 2º. Esses textos foram submetidos à ferramenta computacional Coh-Matrix-Port para a obtenção de índices de inteligibilidade textual.

O público-alvo da Coleção Didática são alunos jovens ou adultos egressos do ensino regular. A concepção pedagógica do material está relacionada ao trabalho com a leitura em contextos da vida prática do aluno jovem ou adulto, que está inserido em um contexto letrado, mas não domina a leitura e a escrita de forma adequada.

A Coleção foi disponibilizada pela Secretaria de Educação Continuada, Alfabetização e Diversidade, criada pelo MEC para implementar políticas públicas voltadas para pessoas com 15 anos ou mais que não completaram o Ensino Fundamental. Foi elaborada por uma equipe de professores da Unitrabalho, uma Rede Nacional de Universidades que apoia a busca de melhores condições de vida e de trabalho.

Além do Caderno do Aluno e do Caderno do Professor com atividades sugeridas para cada texto do Caderno do Aluno nas diferentes disciplinas, como já foi dito, há também um Caderno Metodológico com a concepção pedagógica do material e com perguntas frequentes feitas por professores que acompanharam as oficinas pedagógicas,

realizadas nas cinco regiões do Brasil.

Como já foi assinalado, o tema principal é trabalho e cada Caderno tem um subtema relacionado ao assunto: Caderno 1: Cultura e Trabalho; CD 2: Diversidades e Trabalho; CD 3: Emprego e Trabalho; CD 4: Economia Solidária e Trabalho; CD 5: Globalização e Trabalho; CD 6: Juventude e Trabalho; CD 7: Meio Ambiente e Trabalho; CD 8: Mulher e Trabalho; CD 9: Qualidade de Vida, Consumo e Trabalho; CD 10: Segurança e Saúde no Trabalho; CD 11: Trabalho no Campo; CD 12: Tecnologia e Trabalho; CD 13: Tempo Livre e Trabalho.

METODOLOGIA E RESULTADOS DE ANÁLISE

Uma tabela foi elaborada com os resultados dos índices obtidos pelo Coh-Matrix-Port, para que fosse feita uma comparação entre os 13 Cadernos Didáticos, em termos de dificuldade de processamento de leitura.

As métricas do Coh-Matrix-Port que foram selecionadas (a ferramenta possui mais de 40) para fins de avaliação e comparação dos textos quanto à sua inteligibilidade foram: *Número de palavras por sentenças; Número de verbos por número de sentenças; Sentenças por parágrafos; Sílabas por palavras de conteúdo; Incidência de negações; Frequências; Palavras antes de verbos principais; Incidência de conectivos; Aditivos positivos e negativos; Temporais positivos e negativo; Causais positivos e negativos; Lógicos positivos e negativos; Sobreposição de argumentos adjacentes; Referência anafórica adjacente.*

Além desses índices do Coh-Matrix foi incorporado também o *Índice Flesch* (que foi incluído pelos pesquisadores da USP/São Carlos na ferramenta, e foi adaptado para o português por Martins et al., 1996). Apesar de ser uma fórmula que avalia superficialmente a inteligibilidade do texto, visto que está baseada em aspectos mais materiais, como *Número de palavras em sentenças* (entenda-se, aqui, períodos porque conta-se o que está contido entre dois pontos), e *Número de letras ou sílabas por palavras*, trata-se, como apontam Scarton, Almeida e Aluísio (2009), de um índice que indica, grosso modo, a complexidade da tarefa de leitura. O índice é obtido com as métricas *Flesch Reading Ease* e *Flesch-Kincaid Grade Level*. A primeira “é a única métrica de inteligibilidade já adaptada para o português (Martins et al., 1996) e incorpora

o conceito de séries escolares da segunda”. A fórmula é: $ILF=164.835-[1.015x(\text{Número de palavras por sentença}) - [84.6x(\text{Número de sílabas do texto}/\text{Número de palavras do texto})]$, levando em consideração tamanho de sentenças e de palavras, e relacionando esse índice com níveis de escolaridade.

Textos cujo índice varia entre 100-75 são analisados como **muito fáceis**, adequados às primeiras séries do Ensino Fundamental; textos cujo índice varia de 75-50 são tidos como **fáceis**, adequados às séries finais do Ensino Fundamental. Os que tem Índice Flesch entre 50-25 são analisados como **difíceis** e seriam voltados para Ensino Médio. Por fim, os que ficam entre 25-0 são classificados como **muito difíceis** e direcionados para o ensino universitário.

A seguir, apresentaremos as métricas que foram selecionadas do Coh-Metrix-Port, com a justificativa da seleção para a nossa pesquisa. O primeiro índice, **Números de Verbos por Número de Sentenças** foi por nós calculado com base em dois índices – **Número de Sentenças** e **Incidência de Verbos**. A fórmula usada para calcular a incidência de verbos, definida na ferramenta Coh-Metrix-Port, - (número de verbos/(número de palavras/1000) – nos permitiu chegar ao número de verbos de cada texto. Em seguida, dividimos o número de sentenças pelo número de verbos. A ideia era identificar se o que a ferramenta define como sentença (que na verdade é período, já que é computado o que está entre pontos) corresponderia a períodos simples ou compostos. Se o índice for 2,2, por exemplo, pode significar que a maior parte dos textos daquele Caderno Didático possui períodos compostos de duas orações (cada oração correspondendo a um verbo). Trata-se de um cálculo aproximado, visto que, até onde pudemos identificar, a ferramenta não parece distinguir verbos de locuções verbais. De todo modo, considerando-se que o problema se observa na análise de todos os textos, podemos usar esse parâmetro para comparar os textos. Entendemos que, em princípio, períodos compostos sejam, em termos de processamento, mais complexos do que períodos simples. Logicamente, sabemos que outros fatores, como a natureza sintática do vínculo oracional (coordenação ou subordinação) é também fator determinante dessa complexidade.

O índice **Sentenças por Parágrafos** pode indicar se os parágrafos são longos ou curtos, o que potencialmente pode ser relacionado ao grau de complexidade informacional do parágrafo.

Sílabas por Palavra de Conteúdo foi também mantido, pois pode ser relevante para casos onde há dificuldade com a decodificação do sinal gráfico, como no caso dos analfabetos funcionais. Leitores iniciantes, no ato da leitura, adotam predominantemente um processamento ascendente (*bottom-up*), bastante preso ao próprio *input* visual e, para chegar ao sentido das palavras, precisam do apoio da linearidade das cadeias de letras (cf. Kleiman, 2001, p.36).

Acrescente-se ainda que as palavras de conteúdo desempenham um papel fundamental na leitura para a compreensão. Resultados de trabalho clássico na área de movimento ocular indicam que os leitores fixam 80% do movimento dos olhos nesse tipo de palavra, o mesmo não sendo observado para palavras funcionais (Carpenter & Just, 1983, *apud* Perfetti, 1999).

O índice **Palavras Antes de Verbos Principais** foi selecionado porque, segundo a documentação do Coh-Metrix, é um bom índice para avaliar a carga da memória de trabalho relacionada à manutenção de informação relativa ao sujeito, quando do processamento do verbo. Resultados experimentais apontam para a relevância de informação gramatical e semântica (relativa a papéis temáticos) codificada no verbo (cf. Frazier & Rayner, 1982; Britt, 1994; MacDonald, 1994; *apud* Perfetti, 1999). O elemento inicial da sentença para que possa ser corretamente analisado e interpretado depende, portanto, da informação codificada no verbo. É necessário manter esse elemento inicial, para que seu papel temático possa ser identificado adequadamente. Nesse sentido, quanto maior for a sequência de palavras antes de verbos principais, maior o custo em termos do processamento da leitura.

O índice **Incidência de Negações** (consideradas as negações não, nem, nenhum, nenhuma, nada, nunca e jamais) também foi utilizado, pois a incidência maior de negações parece dificultar a compreensão. Resultados de experimentos envolvendo identificação de figuras a partir de sentenças afirmativas e negativas indicaram que estas são mais difíceis em termos de processamento do que aquelas (Carpenter & Just, 1975; Clark & Chase, 1972; Trabasso et al., 1971 *apud* Tettamanti et al., 2008). Conforme os autores esclarecem nesse artigo, “a compreensão de sentenças negativas demanda maior tempo de processamento, e estas apresentam maior percentagem de erros, conforme resultados de testes que envolvem pareamento de frases e figuras”. Além disso, estudos sobre acessibilidade de informação mencionada no escopo da negação, envolvendo

tarefas de reconhecimento de palavras ou tarefas de *priming* mostraram que os tempos gastos para responder eram significativamente mais longos para os itens com negações (MacDonald and Just, 1989; Kaup, 2001; Kaup and Zwaan, 2003 *apud* Tettamanti et al., 2008).

Na pesquisa realizada pela Profa. Sandra Maria Aluísio e Carolina Evaristo Scarton, do Projeto PorSimples, sobre análise de inteligibilidade em textos reescritos para crianças, o número de negações é menor em textos reescritos (Aluísio & Scarton, 2009, p. 53), o que também pode ser tomado como indicativo de que as negações constituem um custo maior para a leitura.

O índice **Frequências** indica a média de todas as frequências das palavras de conteúdo. O valor da frequência das palavras é retirado da lista de frequências do *corpus* Banco do Português (Sardinha, 2003). A presença de palavras pouco frequentes em um texto pode, evidentemente, também contribuir para sua complexidade. Pesquisas na área de reconhecimento de palavras apontam para o reconhecimento e acesso mais rápido e mais eficiente de palavras frequentes na língua (Moreno & Van Orden, 2000). Como afirma Perfetti (2005), para a eficácia da compreensão, faz-se necessário que a identificação de palavras seja implementada com sucesso, de modo que o significado das mesmas seja recuperado e integrado à representação mental do texto em construção.

Incidência de Conectivos foi outro índice considerado. Ele inclui Aditivos positivos e negativos, Temporais aditivos e negativos, Causais aditivos e positivos e Lógicos positivos e negativos. Alguns exemplos podem ajudar a esclarecer esta relação:

1.a. *Estamos caminhando para nosso objetivo assim como vocês também estão.*
(conectivo aditivo positivo)

1.b. *Estamos caminhando para nosso objetivo ainda que vocês não estejam.*
(conectivo aditivo negativo)

2.a. *A população japonesa dá exemplo ao mundo porque permanece disciplinada.*
(conectivo causal positivo)

2.b. *A população japonesa está sofrendo grandes privações apesar da ajuda mundial.* (conectivo causal negativo)

Em princípio, pode-se considerar que um texto com presença de conectivos facilita o processamento, porque as relações semânticas entre as sentenças são indicadas por esses elementos, o que não ocorre quando o vínculo de ligação entre orações não é

expresso. Neste caso, o leitor precisa, muitas vezes, realizar processos inferenciais para determinar de que tipo são as relações.

Perfetti (2005) afirma que leitores pouco proficientes concentram-se mais na leitura de palavras isoladas do que em buscar coerência no texto. Citando um experimento de Cain e Oakhill (1999), ele conclui que os leitores habilidosos são melhores nas atividades de inferência de conexões do texto. Daí a importância da presença de conectivos para leitores pouco proficientes, que precisam desses elementos para estabelecer as relações semânticas interfrásicas.

Degand e Sanders (2002) afirmam, com base no trabalho de Gaddy, van den Broek & Sung (2001) que há realmente um suporte empírico à opinião de que os conectivos e outros marcadores de coerência desempenham uma função facilitadora durante o processo de leitura. Reportam, ainda, que os conectivos propiciariam um processamento mais rápido de segmentos de texto imediatamente subsequentes (cf. Bestgen & Vonk 1995; Britton, Glynn, Mayer & Penland 1982; Deaton & Gernsbacher 2000; Haberlandt 1982; Sanders 1992; Sanders & Noordman 2000).

Para fins da pesquisa, consideramos que a comparação dos conectivos de tipo negativo poderia ser relevante, visto que a operação envolvida no processamento de conectivos negativos – aqueles que param a ampliação de eventos – pode implicar quebra de expectativa e maior dificuldade de compreensão.

Outros índices considerados, além dos já mencionados, foram os seguintes: **Sobreposição de Argumentos Adjacentes** e **Referência Anafórica Adjacente**.

No caso **da Sobreposição de Argumentos Adjacentes**, diferentemente dos demais parâmetros, números mais elevados podem ser tomados como facilitadores da leitura, pois, em princípio, quando há sobreposição de argumentos, a informação é retomada, facilitando a manutenção da referência e o processo de progressão temática. O exemplo dado no manual do Coh-Metrix-Port é o seguinte:

“(1) Dentro do lago, existem peixes, como a traíra e o dourado, além da palometa, um tipo de piranha. (2) Ela é uma espécie carnívora que se alimenta de peixes. (3) No verão, elas ficam mais próximas das margens da barragem, atraídas pela movimentação das pessoas e por restos de comida que alguns turistas deixam na água quando lavam os pratos.”

De acordo com o Manual, “as sentenças (1) e (2) compartilham um substantivo

(peixes) e, portanto, este par incrementa 1 no valor de correferência. Como também há dois pares de sentenças adjacentes ((1) com (2) e (2) com (3)), o valor final da métrica é $\frac{1}{2}=0,5$.”

A **Referência Anafórica Adjacente** diz respeito à proporção de referências anafóricas entre sentenças adjacentes. O exemplo dado no manual do Coh-Metrix é o seguinte:

*“Dentro do lago, existem peixes, como **a traíra** e o dourado, além **da palometa**, um tipo de **piranha**. **Ela** é uma espécie carnívora que se alimenta de peixes.”*

Segundo os autores, os “candidatos” à retomada anafórica pelo pronome “ela” são *traíra*, *palometa* e *piranha*. Assim, como há três “candidatos” e uma sentença adjacente, o valor final da métrica seria $3/1 = 3$. Logo, quanto mais alto for o valor final dessa métrica mais difícil será, em princípio, o processamento da anáfora, pois mais possibilidades existirão.

Conforme esclarece Perfetti (2001), citando Frederiksen (1981), a integração de informação referencial ao longo do texto tem maior custo de processamento para leitores menos habilidosos. Foi verificado que, quando há dois possíveis referentes para um pronome, esses leitores levam mais tempo para atribuir o pronome ao referente pretendido, ao menos quando o referente não corresponde ao primeiro nome da frase precedente.

Os resultados obtidos após a comparação dos índices selecionados do Coh-Metrix-Port indicam que os Cadernos Didáticos são adequados ao Ensino Fundamental. Elaboramos uma tabela com os valores obtidos pela submissão dos textos à ferramenta para obtenção dos índices de inteligibilidade dos Cadernos e, a seguir, uma tabela com a hierarquia de complexidade dos CDs. Essas tabelas são apresentadas a seguir:

Tabela 4: Resultados dos índices do Coh-Metrix selecionados para fins de comparação dos CDs quanto à legibilidade

Legibilidade/CD	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	CD9	CD10	CD11	CD12	CD13
Índice Flesch	57,81	58,51	56,00	43,93	49,78	50,51	40,89	54,04	42,75	45,31	51,72	44,89	56,24
Número de palavras por sentenças	11,37	11,21	12,43	11,79	14,14	14,60	17,32	13,11	12,98	15,16	14,75	14,17	14,24
*Número de verbos por número de sentenças	1,66	1,74	1,93	1,83	1,91	2,17	2,29	2,01	1,80	2,33	2,15	2,01	2,34
Sentenças por parágrafos	1,37	1,61	2,21	1,88	2,02	1,99	2,53	1,98	1,69	1,94	1,96	1,89	2,08
Sílabas por palavras de conteúdo	2,78	2,76	2,80	3,05	2,94	2,86	3,01	2,83	3,04	2,93	2,83	2,98	2,72
Incidência de Negações	3,27	2,87	4,49	2,18	1,72	2,79	1,34	2,36	1,22	1,75	1,99	2,54	2,98
Frequências	225328	257834	223584	204273	195364	248271	186646	222916	220180	213385	219365	205720	268033
Palavras antes de verbos principais	5,14	4,31	5,00	5,19	5,47	4,94	5,78	5,00	5,71	4,23	4,80	4,72	4,49
Incidência de conectivos	73,97	76,06	77,70	79,96	67,04	70,55	67,42	74,61	74,58	67,71	71,75	76,59	79,47
Sobreposição de argumentos adjacentes	0,41	0,37	0,43	0,39	0,41	0,47	0,48	0,42	0,49	0,43	0,30	0,42	0,48
Referência anafórica adjacente	0,17	0,11	0,15	0,09	0,13	0,36	0,09	0,16	0,08	0,22	0,12	0,17	0,24

Tabela 5 - Hierarquia de legibilidade dos CDs

Legibilidade/ CD	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	CD9	CD10	CD11	CD12	CD13
Índice Flesch	12	13	10	3	6	7	1	9	2	5	8	4	11
Número de palavras por sentenças	12	13	10	11	7	4	1	8	9	2	3	6	5
*Número de verbos por número de sentenças	13	12	8	10	9	4	3	7	11	2	5	6	1
Sentenças por parágrafos	13	12	2	10	4	5	1	6	11	8	7	9	3
Sílabas por palavras de conteúdo	11	12	10	1	5	7	3	8	2	6	9	4	13
Incidência de Negações	2	4	1	8	11	5	12	7	13	10	9	6	3
Frequências	10	12	9	3	2	11	1	8	7	5	6	4	13
Palavras antes de verbos principais	5	12	7	4	3	8	1	6	2	13	9	10	11
Incidência de conectivos	6	9	11	13	1	4	2	8	7	3	5	10	12
Sobreposição de argumentos adiacentes	5	2	8	3	4	10	12	7	13	9	1	6	11
Referência anafórica adiacente	10	4	7	3	6	13	2	8	1	11	5	9	12
Média	8,7	9,2	7,3	6,6	5,2	7,1	3,8	7,3	7,6	6,9	5,9	7	8,4
Ordenação	11	12	8	4	2	7	1	8	9	5	3	6	10

Para fins de compreensão dessa tabela, observa-se que a numeração de 1 a 13 corresponde a uma ordem de complexidade, sendo 1 o número atribuído ao Caderno mais complexo e 13, ao mais simples. Depois de identificada, para cada índice, a ordem de complexidade dos cadernos, foi realizada uma média das posições para cada Caderno, o que está registrado na penúltima linha. Na última linha da tabela, foi verificada, com base na média de cada caderno, a posição global no ranking por grau de complexidade. Cumpre esclarecer que, para o cálculo da média, foi excluído o índice Flesch. A ideia era tentar isolar esse parâmetro dos demais para que, em um segundo momento, pudéssemos avaliar em que medida o índice Flesch, embora sendo uma medida superficial, refletiria, tomado isoladamente, a complexidade dos textos.

Tabela 7: Comparação dos CDs, em termos de ranking de legibilidade, com base no Índice Flesch e no conjunto de parâmetros selecionados

Legibilidade /CD	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	CD9	CD10	CD11	CD12	CD13
Índice Flesch	12	13	10	3	6	7	1	9	2	5	8	4	11
Ordenação	11	12	8	4	2	7	1	8	9	5	3	6	10

Podemos observar que os cadernos 6, 7, e 10 apresentam coincidência de resultados entre o Índice Flesch e a comparação obtida com os outros índices selecionados e classificados por hierarquia de dificuldade. Já os cadernos 1, 2, 4, 8 e 13 apresentam diferença de apenas uma colocação no ranking de legibilidade. Os cadernos 3 e 12 apresentam uma diferença de duas colocações. Esses resultados indicam que, nessa comparação dos parâmetros selecionados do Coh-Metrix-Port com o Índice Flesch, os cadernos, em sua maioria, apresentam uma concordância de valores de legibilidade. No entanto, no caso dos CD5, CD9 e CD11, a diferença é mais expressiva, com destaque para o CD 9, que, dependendo do parâmetro (Flesch VS. Outros índices), figuraria em posições opostas em termos de complexidade. Isso revela que, embora o Índice Flesch possa ser tomado inicialmente como uma medida de legibilidade, não pode ser visto isoladamente na avaliação dos textos. É fundamental incorporar outros parâmetros nesse processo. Num trabalho futuro, seria importante considerar a possibilidade de atribuir pesos diferenciados aos parâmetros, a partir de uma reflexão sobre o que, em termos de processamento, poderia representar mais custo.

CONCLUSÃO

Dos professores exige-se o discernimento de decisões de escolha do material que será mais adequado aos seus alunos, porém, em nossa opinião, seria oportuno a provisão de mecanismos de aferição de dificuldades de leitura dos textos que deixassem os professores em uma situação mais segura para a tomada de decisões. Nesse particular, esperamos estar contribuindo com nossa pesquisa, ainda que outros desdobramentos sejam necessários, de forma a sugerir metodologias de aproximação dessas métricas do Coh-Metrix-Port ao trabalho do professor.

REFERÊNCIAS

ALUÍSIO, Sandra; SCARTON, Carolina Evaristo; ALMEIDA, Daniel Machado. Análise de Inteligibilidade de textos via ferramentas de Processamento de Língua Natural: adaptando as métricas do Coh-Metrix para o Português. In: **The 7th Brazilian Symposium in Information and Human Language Technology** (STIL, 2009) (São Carlos, Brasil, 8-11 setembro de 2009).

COSCARELLI, C. V. Entendendo a leitura. **Revista de Estudos da Linguagem**. Belo Horizonte: UFMG. v. 10, nº 1, p. 7-27, jan./jun. 2002.

DEGAND, L.; SANDERS, T. The impact of relational markers on expository text comprehension in L1 and L2. **Reading and Writing: An Interdisciplinary Journal**, v. 15, p. 739–757, 2002.

KLEIMAN, Angela. Abordagens da leitura. **Scripta**. Belo Horizonte, v.7, n.14, p.13-22, 2004.

_____. **Texto e Leitor: Aspectos cognitivos da leitura**. Campinas: Pontes, 2008.

LIBERATO, Yara; FULGENCIO, Lúcia. **É possível facilitar a leitura: um guia para escrever claro**. São Paulo: Editora Contexto, 2007.

MORENO, M. A.; VAN ORDEN, G. C. Cognitive psychology of word recognition. In M. Lax (Ed.). **International Encyclopedia of the Social Sciences**. Elsevier Science LTD., vol. 24, pp. 16555-16561, 2000.

PERFETTI, C.A. Comprehending written language: a blueprint of the reader. BROWN, Colin M.; HAGOORT, Peter (Eds.) In: **The neurocognition of language**. Oxford: Oxford University Press, 1999.

PERFETTI, C.A. Reading Skills. In: **International Encyclopedia of the Social & Behavioral Sciences**. Oxford: Pergamon, pp.12800-12805, 2001.

PERFETTI, C.A. The acquisition of reading comprehension skill. SNOWLING, Margaret J.; HULME, Charles (Eds.). In: **The science of reading: a handbook**. Oxford: Blackwell, 2005.

SCARTON, Carolina Evaristo; ALMEIDA, Daniel Machado. **Manual das Métricas do Coh-Metrix-Port**. São Paulo: Projeto PorSimples (USP), 2009.

TETTAMANTI, Marco et al. Negation in the brain: Modulating action representations. **NeuroImage**, n. 43, p. 358–367, 2008.

A AUTORA

Nély Silva da Motta Mesquita, mestre em Estudos da Linguagem pela PUC-Rio. Esse artigo é um desdobramento dos resultados obtidos na dissertação, cujo título é Legibilidade Textual nos Cadernos Didáticos de EJA: questões de processamento da leitura; apresentada à banca examinadora em abril de 2011 e em sua versão final em setembro de 2011. A área de interesse é a psicolinguística e a pesquisa sobre o processamento e a aquisição da linguagem.

E -mail: nelysmm@gmail.com