

1

Introdução

Esta pesquisa insere-se no contexto da leitura acadêmica, mais especificamente na compreensão de leitura de textos argumentativos. É pela leitura que o sujeito aperfeiçoa-se, adquire novos conhecimentos, interage criticamente na sociedade e exerce sua cidadania. Se a leitura é fundamental na aquisição de conhecimentos, ela é indispensável na formação acadêmica dos indivíduos, já que está diretamente relacionada à produção científica. É inevitável pensar que os alunos, ao chegarem aos bancos universitários, já possuem a habilidade leitora que deles será exigida. Supõe-se que o sujeito neste nível de educação formal já esteja letrado em leitura. Isso quer dizer que, sozinho, deveria ser capaz de adquirir conhecimentos e que deveria ter as habilidades de reflexão e compreensão requeridas para interagir no meio universitário.

No entanto sabemos pela divulgação na mídia que os testes de leitura realizados por programas governamentais negam essa suposição. O teste do Enem, em nível brasileiro, e o teste de Pisa¹, em nível internacional, vêm demonstrando que os alunos brasileiros, na sua grande maioria, saem do Ensino Médio ainda despreparados para a demanda em leitura que a universidade ou a vida social lhe impõem. A universidade, tida como a última possibilidade de educação formal do sujeito, deveria estar atenta ao pouco rendimento dos seus alunos em leitura e pensar uma forma de intervenção.

Percebemos que existem bons e maus leitores, mas não sabemos descrever exatamente o que os distingue. Isso nos leva a refletir sobre o processo de leitura, a

¹Cf. relatórios divulgados na internet pelo site do INEP:
http://www.inep.gov.br/download/internacional/pisa/PISA2006-Resultados_internacionais_resumo.pdf
http://www.inep.gov.br/download/enem/Relatorio/ENEM_2007.pdf

difficuldade de compreensão textual e sobre as diferenças entre leitores proficientes e não proficientes. Com o objetivo de entender melhor essa distinção, investigamos o gênero resumo, como representante possível da macroestrutura de um texto, para perceber quais características linguísticas em nível textual podem nos fornecer subsídios para a diferenciação entre níveis de compreensão de leitura.

A opção pelo resumo escrito como indicador de nível de compreensão não é ingênua. Sabemos que a escritura de um texto é uma tarefa que apresenta dificuldades específicas e que isso pode mascarar a evidência de compreensão textual. No entanto, como não há como ter acesso direto às representações e operações mentais envolvidas no processo de compreensão, assumimos que o resumo escrito poderia ser utilizado como um parâmetro do processo de compreensão textual de alunos, abstraindo a interferência da dificuldade de produção textual nos resumos produzidos.

Escolhemos trabalhar com resumos porque acreditamos que estes exemplificam, em tese, a capacidade de compreensão de um texto, por possivelmente conterem a representação global do texto que foi sumarizado. Para isso tomamos o texto como uma unidade de análise que tem, na cadeia de um discurso – no nosso caso, o argumentativo escrito – o leitor como sujeito cognitivo ativo e interpretante de uma rede de relações linguísticas.

A leitura é uma habilidade humana complexa, que exige do leitor não só habilidades de ordem cognitiva mas também conhecimentos linguísticos para a construção de sentidos de um texto. As pesquisas em compreensão leitora podem ser orientadas tanto para o processamento da leitura em relação aos aspectos cognitivos inerentes ao leitor, quanto para as características textuais que deflagram a compreensão. A maioria dos leitores realiza automaticamente, em menor ou maior grau, o processo cognitivo de decodificação, sem mesmo perceber como ele acontece. Mas a noção de cognição vai além do processamento do código. A teoria cognitiva hoje se baseia em modelos que sustentam o estudo da leitura como uma ação de natureza interativa, em que o processo básico de decodificação das palavras aciona em paralelo os conhecimentos linguísticos e o conhecimento de mundo do leitor.

Dentre as características textuais que facilitam a compreensão e acionam os processos cognitivos superiores do conhecimento da linguagem, vamos nos referir à macro e microestruturas que, agrupadas em unidades sintático-semânticas, orientam o leitor na compreensão de determinado gênero textual (superestrutura). Ao pensarmos no texto como uma unidade de compreensão, nos perguntamos até que ponto esses elementos textuais, importantes na construção de sentidos, são mantidos quando os sujeitos sumarizam a informação de um texto fonte. Nosso trabalho teve como objetivo investigar quais as operações de sumarização privilegiadas pelos sujeitos ao resumir um texto do gênero artigo de opinião, em que predominam sequências estruturais argumentativas, e avaliar se a macroestrutura textual presente nos resumos pode ou não ser analisada como um parâmetro de compreensão textual.

Nessa pesquisa vamos nos deter nas operações de sumarização realizadas por um grupo de alunos universitários que resumiu um texto do gênero artigo de opinião e na análise de alguns elementos linguísticos textuais que concorrem para a construção da textualidade e, portanto, para a interpretação dos sentidos do texto. Acreditamos que a compreensão leitora advém da construção de uma representação coerente, ativada não somente pelas habilidades cognitivas mas também pelas pistas sintático-semânticas do texto e pelo conhecimento de mundo do leitor.

Os estudos realizados durante esta pesquisa na área da Psicolinguística mostram que, se um sujeito não consegue identificar certas pistas sintático-semânticas de um texto, ele, provavelmente, não terá acesso ao sentido do texto. Faz-se necessário entender quais são essas pistas para que indiquem quais habilidades cognitivas e linguísticas pontuais e globais são necessárias para a construção da representação de um texto.

Queremos entender de que forma os fatores acima citados contribuem para a elaboração de um resumo que reflita a macroestrutura de um texto, com o objetivo de elucidar aspectos que colaboram para a compreensão em leitura. Compreende-se por macroestrutura a definição de van Dijk (2004), segundo o qual elas são uma espécie de estrutura profunda semântica abstrata do texto que forma sua coerência global. A

macroestrutura “explica o que é mais relevante, importante, ou proeminente na informação semântica do discurso como um todo” (p. 51). Conforme o autor, o resumo seria a expressão mais fidedigna de uma macroestrutura textual. Nesse sentido tem-se que a macroestrutura é formada a partir da estrutura global textual que é mais facilmente lembrada e pode revelar a compreensão, armazenamento (em nível abstrato mental) e reprodução de um texto. Tomamos nesta pesquisa o conceito de macroestrutura numa adequação com resumos. Supõe-se que, ao descobrirmos quais elementos dessa macroestrutura estão presentes em um resumo, estabeleceríamos sua linha de coesão e coerência textual e a compreensão, ou não, de um texto. Com a análise desses resumos, pretendeu-se ter acesso aos elementos da macroestrutura textual que, conforme vimos, pode indicar a compreensão de um texto.

Outro objetivo desta pesquisa foi contribuir para a avaliação de métodos de sumarização automática de textos. Gostaríamos de destacar quais informações contidas nos resumos humanos podem ser mais relevantes para um sumarizador e em que medida elas são recuperadas automaticamente. Para isso, compararam-se os resumos elaborados pelos alunos com os resumos feitos por um sumarizador automático implementado para o português do Brasil, o GistSumm².

Sumarizadores automáticos são programas computacionais que produzem um sumário ou resumo de textos. Em geral, a composição dos resumos segue critérios lexicais, partindo das palavras-chave recorrentes em um texto; sintáticos, como, por exemplo, eliminando sintagmas em função adjuntiva; ou posicionais, por exemplo, pontuando a posição de uma sentença no texto como indicadora de sua relevância, para, a partir daí, selecionar as sentenças que comporão o resumo.

Para cumprir os objetivos propostos nesta pesquisa, vamos abordar a leitura, primeiramente, como um processo psicolinguístico em que o leitor lança mão de habilidades cognitivas³ disponíveis em seu arcabouço mental e do seu conhecimento

² Cf. definições e conceitos mais detalhados no Capítulo 04: “A sumarização automática”.

³ Estas habilidades cognitivas estarão mais bem definidas no Capítulo 02: “A compreensão leitora-Revisão da literatura.

enciclopédico acumulado. Posteriormente, ao analisarmos os elementos textuais⁴ sintáticos de coesão e coerência, bem como os elementos proposicionais de ordem semântica presentes na macroestrutura dos resumos, temos a intenção de reconhecer os diferentes níveis de compreensão leitora.

O leitor, quando se torna proficiente, possui a habilidade de retirar significado daquilo que lê. Para Goodman (1973, p.23), o processo de compreensão inicia-se com a decodificação do sinal gráfico e termina com o significado construído pelo leitor. O sucesso para a compreensão de um texto requer que o leitor identifique palavras, entenda as estruturas sintáticas e extraia o significado de sentenças individuais, relacionando-as ao todo do texto, a seu gênero, propósito, relação com outros textos e assim por diante. Se um texto é bem compreendido, está-se gerando uma representação na memória como uma estrutura coerente e não em forma de informações desconexas. Estas representações são importantes para qualquer tarefa que exija o acesso à informação retirada do texto e retida na memória. Os elementos retidos no modelo de compreensão ou macroestrutura são oriundos, de acordo com van Dijk (2004), da interpretação das proposições do texto durante a leitura. A macroestrutura organiza, através de regras recursivas de generalização, apagamento e construção, as informações que vão sendo processadas, transformando o texto em um sumário menos complexo e mais acessível.

A pesquisa encontrou fundamentação e aporte teórico em leituras no campo da Ciência Cognitiva, Psicolinguística, Linguística Textual, Análise do Discurso e Linguística Computacional. Na presente investigação, estudou-se o resumo como um possível exemplo de compreensão leitora de um texto de natureza argumentativa. A pesquisa se utilizou de dois corpora de investigação: 14 resumos sublinhados com as idéias principais do texto e 10 resumos elaborados livremente, respectivamente, por 14 e 10 estudantes universitários. Os resumos livres foram ainda avaliados por dois

⁴ Cf. definições e conceitos mais detalhados no Capítulo 02, seção 2.2: Caracterização do texto: macroestrutura, microestrutura e superestrutura.

profissionais da área de Letras que tiveram a tarefa de escolher três deles: um resumo que considerassem ótimo, um que achassem satisfatório e outro, insatisfatório.

Neste estudo exploratório verificamos as operações de sumarização privilegiadas na produção de resumos de um texto argumentativo. Procuramos identificar, nos resumos dos sujeitos, a manutenção das idéias principais de um texto fonte e verificar quais elementos constituintes em nível sintático e semântico foram mantidos, alterados ou apagados na macroestrutura, discutindo as implicações dessas escolhas para a composição de um exemplo de compreensão geral do texto fonte. O estudo se concentrou também na análise da construção da cadeia referencial nos resumos, a partir da identificação de expressões referenciais do texto fonte e da sua manutenção nos textos resumidos. Investigamos ainda como aparecem nos resumos as relações de ligação entre proposições por meio de conectivos e que operações foram realizadas nas proposições do texto fonte ao elaborarem-se os resumos, para evidenciar se elas incidiam mais em termos predadores ou termos argumentais.

Apresentamos, a seguir, a forma como organizamos esta dissertação.

No capítulo 2, buscamos entender como ocorre o processamento e a compreensão na leitura e apresentamos as estruturas constituintes de um texto. Caracterizamos o texto, a partir da Linguística Textual, como uma unidade de sentido que apresenta coerência e coesão. Abordamos, à luz da Psicolinguística, os conceitos de macroestrutura textual, como uma estrutura de informações semânticas resultante de um texto, e os conceitos de microestrutura textual e de superestrutura textual. Indicamos os elementos textuais de coesão, como os conectivos e as relações semânticas que estes sinalizam, e a cadeia referencial, como pistas para a percepção da macroestrutura. Evidenciamos, neste estudo, as proposições que nos orientam para a rede de inter-relações textuais no processamento da compreensão.

Considerações sobre gênero e sua influência na compreensão leitora serão apresentadas no capítulo 3. Destacamos nesse capítulo as propriedades do texto argumentativo e discutimos alguns conceitos sobre o resumo. Nessa categorização

procuramos definir os dois gêneros envolvidos na pesquisa, o texto argumentativo e o resumo, já que a percepção da estrutura retórica do gênero a ser sumarizado é importante na compreensão e sumarização de um texto.

No capítulo 4 nos referimos a métodos e programas computacionais sumarizadores e à importância dos estudos linguísticos para o aprimoramento da geração de sumários automáticos.

No capítulo 5 descrevemos a metodologia, os instrumentos e os critérios utilizados para a investigação. Analisamos os dados quantitativos e qualitativos referentes ao estudo exploratório.

Na conclusão, procuramos refletir sobre a pesquisa efetuada, apresentamos os resultados obtidos em nossa pesquisa, apontamos suas limitações, bem como suas contribuições e nossas perspectivas para trabalhos futuros.

Apresentamos ao final as referências bibliográficas utilizadas como fundamentação para esta pesquisa e, ao final, os anexos, incluídos para facilitar a leitura crítica e a compreensão das análises efetuadas por esta pesquisa.