



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE COMUNICAÇÃO E EXPRESSÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

Bruna Alexandra Franzen

**Processamento anafórico em compreensão leitora de texto acadêmico-científico: “o mesmo” sob as lentes da Psicolinguística**

Florianópolis

2022

Bruna Alexandra Franzen

**Processamento anafórico em compreensão leitora de texto acadêmico-científico: “o mesmo” sob as lentes da Psicolinguística**

Tese submetida ao Programa de Pós-Graduação em Linguística da Universidade Federal de Santa Catarina para a obtenção do título de doutora em Linguística.

Orientadora: Prof.<sup>a</sup> Dra. Ana Cláudia de Souza

Florianópolis

2022

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Franzen, Bruna Alexandra

Processamento anafórico em compreensão leitora de texto acadêmico-científico : "o mesmo" sob as lentes da Psicolinguística / Bruna Alexandra Franzen ; orientadora, Ana Cláudia de Souza, 2022.

228 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão, Programa de Pós Graduação em Linguística, Florianópolis, 2022.

Inclui referências.

1. Linguística. 2. Leitura. 3. Anáfora Correferencial. 4. "O mesmo". 5. Compreensão. I. Souza, Ana Cláudia de. II. Universidade Federal de Santa Catarina. Programa de Pós Graduação em Linguística. III. Título.

Bruna Alexandra Franzen

**Processamento anafórico em leitura de texto acadêmico-científico: “o mesmo” sob as lentes da  
Psicolinguística**

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. José Ferrari Neto, Dr.  
Universidade Federal da Paraíba (UFPB)

Prof.<sup>a</sup> Lêda Maria Tomitch, Dra.  
Universidade Federal de Santa Catarina (UFSC)

Prof. Márcio Martins Leitão, Dr.  
Universidade Federal da Paraíba (UFPB)

Prof.<sup>a</sup> Otilia Lizete de Oliveira Martins Heinig, Dra.  
Universidade Regional de Blumenau (FURB)

Prof.<sup>a</sup> Thais de Souza Schlichting, Dra.  
Universidade Regional de Blumenau (FURB)

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de doutora em Linguística.

---

Prof. Dr. Valter Pereira Romano

Coordenador do Programa

---

Prof. Dra. Ana Cláudia de Souza

Orientadora

Florianópolis, 2022.

Aos meus pais, às minhas irmãs e às minhas sobrinhas, que são meu amparo em todos os momentos.

## AGRADECIMENTOS

Momento de agradecer, afinal essa caminhada não foi feita na solidão. Muito pelo contrário, ela foi construída na interação e na ajuda de muitas pessoas. Por isso, agradeço!

Aos meus pais, às minhas irmãs, aos meus cunhados e às minhas sobrinhas. Mesmo sem saber, são eles que me dão força para seguir nos momentos mais obscuros e desanimadores. Agradeço especialmente ao Nuna (em memória), meu cunhado que partiu sem me ver doutora, mas que, desde o momento da aprovação no doutorado, já me chamava com esse título. Eu dizia: calma, doutora só daqui quatro anos! Bom, ele não estará aqui para me chamar de doutora, então, que bom que ele já confiava que o título viria. Certamente, muito da força veio dessa certeza em me ver construir esse caminho. Obrigada, “cOnhado”! Agradeço, também, de forma muito especial, à minha sobrinha Júlia, que esteve ao meu lado sempre, trazendo alegria e diversão nos momentos difíceis, segurando a minha mão e me fazendo rir. Minha parceira em muitas histórias. Além disso, me auxiliou no momento de checagem das referências! Obrigada, Jú!

Agradeço à minha orientadora, Ana Cláudia de Souza, que, desde o início, me acolheu de forma carinhosa e atenciosa. Com muita ética e respeito, trilhou comigo esse caminho, buscando sempre o melhor de mim e acreditando em mim quando eu já não acreditava mais. Teve paciência nos momentos em que a ansiedade me dominava, me ensinou sobre leitura, sobre compreensão (em todos os aspectos), sobre Psicolinguística, mas, para muito além disso, me ensinou sobre amparo, amizade e sobre a importância de boas risadas e de momentos de leveza. Foi minha orientadora no doutorado e se tornou minha parceira de trilhas e amiga querida na vida. Obrigada por tudo, Ana! Sem você, nada disso teria sido possível.

Agradeço à Otilia Heinig, minha sempre orientadora, quem me inseriu no mundo da pesquisa e nunca mais largou minha mão. Acompanhou de perto toda a minha trajetória e está comigo em todos os momentos. Um abraço amigo, um café gostoso, um dia de sol, um joguinho numa sexta chuvosa na praia, ela sempre está presente. As palavras jamais serão suficientes para te agradecer. Obrigada, Ótis, obrigada por seres tão presente e tão parceira. Obrigada por me permitir fazer parte da tua vida e por estares na minha em todos os momentos!

Agradeço à minha amiga Thataruga, parceira de vida e de academia. Trilhou comigo o processo da pesquisa desde os primórdios da graduação. Foi minha caloura e minha veterana. Foi a culpada por eu ter decidido entrar no doutorado. Foi a pessoa que mais ouviu minhas reclamações e minhas dores. Riu e chorou comigo. Nos dias difíceis, compartilhou vídeos fofos de cachorrinhos e memes engraçados. Mesmo longe geograficamente, sempre esteve presente. Obrigada, miga!

De modo muito especial, agradeço, também, à minha amiga Luana, que foi parceira durante essa trajetória do doutorado, amiga de caronas, de dividir apartamento e, como se não bastasse, ainda ajudou a avaliar alguns dos testes aplicados e respostas obtidas no teste de compreensão em leitura. Obrigada, Lu! Obrigada pela tua companhia, por tua amizade e por teu carinho.

Agradeço aos professores que aceitaram avaliar meu trabalho e participar da banca de qualificação e de defesa desta pesquisa. Obrigada, professor José Ferrari Neto, professor Márcio Leitão, professora Lêda Tomitch, professora Otilia Heinig e professora Thais Schlichting. Os trabalhos de vocês foram suporte e inspiração para que esta pesquisa acontecesse e o olhar de vocês é fundamental para a conclusão dela. Obrigada por aceitarem estar comigo neste momento tão importante. Vocês foram escolhidos com muito carinho e cuidado.

Agradeço aos meus queridos colegas de orientação, que compartilharam comigo tanto do seu conhecimento e que me auxiliaram de diferentes formas em distintos momentos dessa caminhada. Obrigada, Dohane, Marga, Cris, Pâmela, Letícia e Jonas.

Agradeço, ainda, aos professores Leandro, Júlio e Hector, do curso de Engenharia de Controle e Automação da UFSC-Florianópolis, que abriram para mim as portas do curso, me

auxiliaram na seleção dos textos de área, avaliaram o teste de compreensão em leitura e me deram dicas valiosas sobre os alunos do curso.

Agradeço ao CTC e ao curso de Engenharia de Controle e Automação da UFSC, pois, desde o primeiro contato, foram muito solícitos e abrigaram a pesquisa com confiança e disposição.

Aos alunos do curso de Engenharia de Controle e Automação que participaram do estudo piloto e do estudo principal, meu muito obrigada! Sem vocês, esta pesquisa não seria possível. Obrigada pela paciência e pela participação.

Agradeço, também, à professora Hanna Kivistö de Souza, que me ajudou a montar o experimento de leitura automonitorada no DMDX, resolveu problemas para os quais eu não conseguia encontrar solução e fez tudo isso com paciência e cuidado em um momento que a ansiedade e o desespero já batiam fortes em mim. Muito obrigada, professora!

Agradeço, também, ao engenheiro de controle e automação Flávio Barbosa, que, na época de seu doutorado, me ajudou também com a avaliação do teste de compreensão em leitura e, depois, me ajudou na correção desses testes. Teu olhar e tua ajuda foram fundamentais para que a pesquisa acontecesse com o rigor necessário. Obrigada!

Agradeço à Cristiane Seimetz Rodrigues que, gentilmente, fez a avaliação das frases a serem utilizadas no experimento de leitura automonitorada e do julgamento de aceitabilidade, contribuindo de forma imprescindível para a melhoria do experimento.

Do mesmo modo, agradeço à engenheira Carla, que teve um papel fundamental na avaliação prévia do teste de compreensão leitora a ser aplicado aos participantes. Agradeço pelo olhar atento e pelas sugestões.

Agradeço, ainda, à professora Cristiane Dall Cortivo Lebler e à sua turma de estágio de 2019/II, que contribuíram, também, na avaliação das frases para o experimento de leitura automonitorada e do julgamento de aceitabilidade.

Agradeço, também, aos alunos da turma de estágio 2019/II da professora Ana Cláudia de Souza, os quais aceitaram participar como voluntários da pilotagem do experimento de leitura automonitorada. Agradeço pelo tempo cedido e pelas contribuições durante a testagem.

Agradeço, ainda, aos alunos do ensino médio do Instituto Federal Catarinense de Ibirama, os quais, no ano de 2019, sob orientação da professora Thais, realizaram o teste de aceitabilidade piloto.

Agradeço ao professor Rafael Matielo, que realizou a rodada estatística desta tese e teve a paciência de ensinar esta pesquisadora ainda amadora no campo da estatística. Obrigada pela paciência, pelas horas de explicações e pelos socorros dados nos momentos de desespero.

Por fim, agradeço à Capes, que, por meio da bolsa de pesquisa, contribuiu para que eu pudesse realizar este estudo de forma mais tranquila com um suporte financeiro. Muito obrigada!

O produto final é este texto, mas, para além dele, há um mundo de descobertas, de pessoas, de olhares. Há sempre alguém no meio do caminho para estender a mão, dar amparo, um abraço amigo, um café no meio do furacão, uma palavra, um carinho. Pessoas! Esse texto é feito de pessoas, só existe por conta de todas essas pessoas e de tantas outras que não foram citadas aqui, mas que também passaram por mim, me constituem e me fizeram chegar até aqui. Às vezes, o caminho parece solitário. O momento de escrita normalmente o é, mas ao parar para analisar, não há como se sentir só diante de tantas mãos estendidas nos mais diferentes momentos desses quase cinco anos. Gratidão por tê-los encontrado, gratidão por ter chegado aqui!

## RESUMO

Estudar a leitura sob as lentes da Psicolinguística envolve pensar nos diferentes processos que ocorrem cognitivamente para que a compreensão seja construída. Nesse sentido, diversos fatores que se relacionam ao texto, ao leitor, à situação e ao contexto de leitura assumem papel relevante. No que diz respeito à presente investigação, busca-se olhar para um elemento inerente ao texto, que atua na sua microestrutura (KINTSCH; RAWSON, 2013) e auxilia na integração textual (KINTSCH; RAWSON, 2013; JUST; CARPENTER, 2013; OAKHIL; CAIN; ELBRO, 2017): a anáfora correferencial. De modo mais específico, centra-se na realização anafórica construída com o sintagma “o mesmo”, que não tem na correferência uma de suas funções mais salientes e aceitáveis segundo as gramáticas escritas. Busca-se, assim, pensar acerca do papel dessa anáfora correferencial na compreensão em leitura, sobretudo por se tratar de uma construção tida como não adequada para alguns estudiosos da língua (BAGNO, 2012; SACCONI, 1994; ALMEIDA, 2009). Diante disso, este estudo tem por objetivo compreender de que forma e em que medida “o mesmo” usado como anáfora correferencial influencia no processamento e na compreensão em leitura de texto acadêmico-científico. Esse objetivo geral se desdobra em cinco objetivos específicos: 1. Conhecer o custo de tempo de resposta diante do sintagma “o mesmo” como anáfora correferencial; 2. Explorar o tempo e a acurácia diante de anáforas correferenciais de forma a relacioná-los a estratégias utilizadas e ao desempenho em leitura de texto acadêmico-científico; 3. Identificar a aceitabilidade do uso de “o mesmo” como elemento anafórico correferencial em construções sentenciais do PB; 4. Explorar o desempenho e o comportamento estratégico do leitor diante do texto e do sintagma “o mesmo”; 5. Caracterizar e compreender as estratégias utilizadas durante a leitura de texto acadêmico-científico, no que tange à anáfora, e a relação de tais estratégias com a resolução bem-sucedida dessa anáfora. A hipótese aventada prevê que “o mesmo” segue um padrão de processamento e de integração semelhante ao do pronome “ele”. Assim, esperava-se que os resultados em termos de processamento mostrassem que “o mesmo” pudesse exercer a função de anáfora correferencial, sem que a compreensão em leitura fosse dificultada por maiores custos para o processamento. Trata-se de um estudo experimental com exploração de questões específicas acerca do comportamento leitor e de estratégias empregadas diante das anáforas correferenciais realizadas com o sintagma “o mesmo”. O enfoque é explicativo, com abordagem do problema feita de forma mista, realizando os processos analíticos por meio de discussões qualitativas e quantitativas. Os participantes da pesquisa foram vinte e seis acadêmicos do curso de Engenharia de Controle e Automação da Universidade Federal de Santa Catarina. A coleta de dados ocorreu em dois momentos. No primeiro, aplicaram-se as técnicas de protocolo verbal (ERICSSON; SIMON, 1993) e de leitura automonitorada (MITCHELL, 2004; KAISER, 2013). No segundo momento, ocorreram o teste de aceitabilidade (SCHÜTZE; SPROUSE, 2013) e o teste de compreensão em leitura (BRASIL, 2016), com a gravação da tela do computador, captando os movimentos realizados pelos participantes por meio do mouse e do teclado. Para análise, as estratégias e o comportamento dos participantes foram categorizados e, após tabulados, foram cruzados com os dados relativos ao tempo de leitura e ao desempenho na compreensão em leitura. O teste de leitura automonitorada mostrou que o pronome e o sintagma são processados de modo semelhante, sem que haja diferenças significativas entre ambos os anafóricos. No que diz respeito à relação entre tempo de leitura e desempenho local, há uma tendência de que quanto maior a pontuação nas questões locais, menor o tempo de leitura das frases construídas com o sintagma e com o pronome. Tal resultado vai em direção ao esperado, já que quanto mais rápida e acurada é a leitura, entende-se que maior será o índice de acerto.



No que tange à aceitabilidade, ambos os anafóricos foram considerados aceitáveis, ainda que a aceitabilidade do pronome tenha sido significativamente maior que a do sintagma. Além disso, a partir da análise do comportamento dos participantes durante a realização do teste de compreensão em leitura e dos resultados obtidos no teste, é possível inferir que o uso da anáfora realizada com o sintagma “o mesmo” não se mostra como um complicador no desempenho dos participantes. Tal resultado se revela verdadeiro também quando são analisadas as estratégias utilizadas pelos participantes. Ou seja, as anáforas realizadas com o sintagma “o mesmo” não parecem ter afetado de forma negativa a compreensão geral do texto. Portanto, de modo geral, os resultados sugerem que, em termos de processamento, “o mesmo” pode exercer a função de anáfora correferencial sem que haja problemas para a construção da base textual e sem que compreensão em leitura seja dificultada por maior demanda de processamento.

Palavras-chave: Leitura. Anáfora Correferencial. “O mesmo”. Compreensão. Psicolinguística Aplicada.

## ABSTRACT

Studying reading through the lenses of Psycholinguistics involves thinking about the different processes that occur cognitively so that understanding can be built. In this sense, several factors that relate to the text, to the reader, to the situation and to the reading context play a relevant role. In respect to the present investigation, we intend to look at an element inherent to the text, which acts on its microstructure (KINTSCH; RAWSON, 2013) and assists in textual integration (KINTSCH; RAWSON, 2013; JUST; CARPENTER, 2013; OAKHIL; CAIN; ELBRO, 2017): the coreferential anaphora. More specifically, it focuses on the anaphoric realization constructed with “o mesmo”, which does not have in coreference one of its most salient and acceptable functions in Brazilian written grammars. Thus, we propose to investigate the role of this co-referential anaphora in reading comprehension, especially because it is a construction considered inappropriate by some language scholars (BAGNO, 2012; SACCONI, 1994; ALMEIDA, 2009). Therefore, this study aims to understand how and to what extent "o mesmo" used as a coreferential anaphor influences the processing and reading comprehension of academic-scientific text. This general goal is divided into five specific objectives: 1. To know the cost of response time in the face of the syntagma “o mesmo” as a coreferential anaphora; 2. Explore the time and accuracy of coreferential anaphoras in order to relate them to the strategies used and the performance in reading academic-scientific texts; 3. Identify the acceptability of the use of “o mesmo” as a coreferential anaphoric element in BP (Brazilian Portuguese) sentence constructions; 4. Examine the performance and strategic behavior of the reader in front of the text and the syntagma “o mesmo”; 5. Characterize and understand the strategies used when reading academic-scientific text, with regard to anaphora, and the relationship of such strategies with the successful resolution of this anaphora. The hypothesis raised for this study predicts that “o mesmo” follows a pattern of processing and integration similar to that of the pronoun “ele”. Thus, it was expected that results in terms of processing would demonstrate that “o mesmo” could perform the function of co-referential anaphora, without reading comprehension being hindered by higher processing costs. It is an experimental study with exploration of specific questions about reader behavior and strategies used in the face of coreferential anaphora carried out with “o mesmo”. It focuses on an explanatory perspective, with a mixed approach to the problem, carrying out the analytical processes through qualitative and quantitative discussions. The research participants were twenty-six undergraduate students from the Control and Automation Engineering course at the Federal University of Santa Catarina. Data collection took place in two moments. In the first, the techniques of verbal protocol (ERICSSON; SIMON, 1993) and self-paced reading (MITCHELL, 2004; KAISER, 2013) were applied. In the second moment, the acceptability test (SCHÜTZE; SPROUSE, 2013) and the reading comprehension test (BRASIL, 2016) took place, with the recording of the computer screen, capturing the movements performed by the participants through the mouse and keyboard. For analysis, the strategies and behavior of the participants were categorized and, after tabulation, were crossed with data related to reading time and performance in reading comprehension. The self-paced reading test showed that the pronoun “ele” and “o mesmo” are processed in a similar way, without significant differences between both anaphoras. Regarding the relationship between reading time and local performance, there is a tendency that the higher the score on local questions, the shorter the reading time of sentences constructed with the syntagma and with the pronoun. This result is in the direction of what was expected, since the faster and more accurate the reading, it is understood that the higher the accuracy rate will be. Regarding acceptability, both anaphoras were considered acceptable, although the acceptability of the pronoun was significantly higher than that of “o mesmo”. In addition, from the analysis of the participants' behavior during the reading comprehension test and the results obtained in the test, it is possible to infer that the use of the anaphora performed with “o mesmo” does not prove to be a complicating factor in the participants' performance. This result is also true when

the strategies used by the participants are analyzed. That is, the anaphors made with syntagma “o mesmo” do not seem to have negatively affected the general understanding of the text. Therefore, in general, the results suggest that, in terms of processing, “o mesmo” can perform the function of co-referential anaphora without any problem in the construction of the textbase and without the reading comprehension being hampered by greater processing demand.

**Keywords:** Reading. Co-referential anaphora. "O mesmo". Comprehension. Applied Psycholinguistics.

## LISTA DE FIGURAS

Figura 1: O quebra-cabeça da leitura.....	60
Figura 2: Coleta de dados .....	98

## LISTA DE GRÁFICOS

Gráfico 1: teste dos Postos Sinalizados de Wilcoxon – tempo total em milissegundos.....	149
Gráfico 2: teste dos Postos Sinalizados de Wilcoxon – tempo segmento pós-anafórico .....	152
Gráfico 3: Desempenho Geral Categorizado.....	165
Gráfico 4: Desempenho Local Categorizado.....	165

## LISTA DE QUADROS

Quadro 1: Caracterização dos Participantes da Pesquisa .....	92
Quadro 2: Descrição e percentual de estudantes nos sete níveis de proficiência em leitura – PISA 2015.....	116
Quadro 3: identificação do nível e do aspecto de cada questão .....	119
Quadro 4: Parâmetros e Critérios seguidos no Teste de Compreensão em Leitura .....	120
Quadro 5: Estratégias utilizadas durante o Protocolo Verbal.....	138
Quadro 6: Resumo das variáveis e dos testes aplicados.....	142
Quadro 7: Elaboração diante do sintagma “o mesmo”.....	183
Quadro 8: Paráfrase diante do sintagma “o mesmo”.....	184
Quadro 9: Avaliação do texto diante da anáfora "o mesmo" .....	194
Quadro 10: Monitoramento da Compreensão e Releitura diante da anáfora "ele".....	195

## LISTA DE TABELAS

Tabela 1: Acurácia “O mesmo” .....	146
Tabela 2: Acurácia “ele” .....	146
Tabela 3: Estatística do teste de acurácia .....	147
Tabela 4: Estatística Descritiva – Leitura Automonitorada – Tempo Total em milissegundos .....	147
Tabela 5: Normalidade – Teste de Leitura Automonitorada – Tempo Total em milissegundos .....	148
Tabela 6: Postos do tempo total de leitura em milissegundos .....	148
Tabela 7: Estatística do tempo total de leitura em milissegundos .....	149
Tabela 8: Estatística Descritiva – Leitura Automonitorada – Tempo em milissegundos Segmento Pós-anafórico .....	150
Tabela 9: Normalidade – Teste de Leitura Automonitorada – Tempo Segmento Pós-anafórico .....	150
Tabela 10: Postos do tempo de leitura do segmento pós-anafórico em milissegundos.....	151
Tabela 11: Estatística do tempo de leitura do segmento pós-anafórico em milissegundos....	151
Tabela 12: Estatística Descritiva – Leitura Automonitorada – Média em milissegundos.....	154
Tabela 13: Normalidade – Teste de Leitura Automonitorada – Média.....	154
Tabela 14: Estatística Descritiva – Teste de Compreensão Leitora – Local .....	155
Tabela 15: Normalidade – Teste de Leitura Automonitorada – Média.....	155
Tabela 16: Correlações - Pearson .....	156
Tabela 17: Correlações - Spearman.....	156
Tabela 18: Estatística Descritiva – Teste de Aceitabilidade.....	158
Tabela 19: Teste de Aceitabilidade – Frequência “O mesmo” .....	158
Tabela 20: Teste de Aceitabilidade – Frequência “Ele” .....	159
Tabela 21: Teste de Aceitabilidade - Postos.....	159
Tabela 22: Estatística do teste de aceitabilidade - Testes dos Postos Sinalizados de Wilcoxon .....	160
Tabela 23: Estatística Descritiva – Teste de Compreensão Leitora .....	162
Tabela 24: Desempenho Individual no Teste de Compreensão Leitora .....	162
Tabela 25: Desempenho Individual no Teste de Compreensão Leitora - Local .....	163
Tabela 26: Estatística Descritiva – Teste de Compreensão Leitora Média .....	166
Tabela 27: Normalidade – Teste de Compreensão Leitora – Média .....	166

Tabela 28: Estatística Descritiva – Monitoramento Teste de Compreensão Leitora .....	166
Tabela 29: Normalidade – Monitoramento Teste de Compreensão Leitora .....	167
Tabela 30: Estatística Descritiva – Monitoramento Teste de Compreensão Leitora Local ...	168
Tabela 31: Normalidade – Monitoramento Teste de Compreensão Leitora Local .....	168
Tabela 32: Correlações de Pearson do Mapeamento do Comportamento Geral x Mapeamento do Comportamento “O Mesmo” .....	169
Tabela 33: Correlações de Postos de Spearman do Mapeamento do Comportamento Geral x Mapeamento do Comportamento “O Mesmo” .....	169
Tabela 34: Correlações de Pearson do Mapeamento do Comportamento Geral x Desempenho Compreensão Leitora Geral.....	170
Tabela 35: Correlações de Postos de Spearman do Mapeamento do Comportamento “O Mesmo” x Desempenho Compreensão Leitora Local.....	171
Tabela 36: Estatística Descritiva – Protocolo Verbal.....	175
Tabela 37: Normalidade – Protocolo Verbal.....	176
Tabela 38: Correlações de Postos de Spearman dos Protocolos Verbais e Compreensão Leitora Local.....	177
Tabela 39: Correlações de Pearson dos Protocolos Verbais e Compreensão Leitora Geral ..	178
Tabela 40: Correlações de Postos de Spearman dos Protocolos Verbais e Compreensão Leitora Geral .....	179
Tabela 41: Estatística Descritiva – Protocolo Verbal – Comparação “ele” e “o mesmo” .....	180
Tabela 42: Teste de Wilcoxon - Protocolo Verbal – Comparação “ele” e “o mesmo” .....	182
Tabela 43: Comparação de médias do tempo em milissegundos – Leitura Automonitorada	187



## **LISTA DE ABREVIATURAS E SIGLAS**

IFC - Instituto Federal Catarinense

LA - Leitura Automonitorada

OCDE - Organização para Cooperação e Desenvolvimento Econômico

PB - Português Brasileiro

PDF - Portable Document Format

PISA - Programa Internacional de Avaliação de Estudantes

PV – Protocolo Verbal

SPSS - Statistical Package for Social Science

TA - Teste de Aceitabilidade

TCL - Teste de Compreensão em Leitura

TCLE - Termo de Consentimento Livre e Esclarecido

TLSP - Tempo de Leitura do segmento pós-crítico no experimento de leitura automonitorada

TLT - Tempo de Leitura Total no experimento de leitura automonitorada

UFSC - Universidade Federal de Santa Catarina

## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>20</b>
1.1 JUSTIFICATIVA E DELIMITAÇÃO DO ESTUDO .....	21
1.2 OBJETIVOS, PERGUNTAS E HIPÓTESES.....	26
1.3 ORGANIZAÇÃO DA TESE .....	28
<b>2 FUNDAMENTAÇÃO TEÓRICA.....</b>	<b>30</b>
2.1 A LEITURA .....	30
<b>2.1.1 O texto.....</b>	<b>31</b>
<b>2.1.2 O leitor .....</b>	<b>33</b>
2.1.2.1 <i>As estratégias de leitura .....</i>	35
<b>2.1.3 A situação da leitura.....</b>	<b>37</b>
<b>2.1.4 Modelos cognitivos de compreensão em leitura.....</b>	<b>38</b>
<b>2.1.5 Alguns aspectos sobre memória .....</b>	<b>49</b>
2.2 A NOÇÃO DE REFERÊNCIA E OS PROCESSOS DE REFERENCIAÇÃO NA LINGUÍSTICA.....	52
2.3 A ANÁFORA CORREFERENCIAL EM LEITURA.....	57
<b>2.3.1 As investigações sobre o processamento anafórico em leitura .....</b>	<b>61</b>
2.4 A FORMA ANAFÓRICA ESTUDADA NESTA TESE: “O MESMO” .....	74
<b>2.4.1 Os estudos sobre o sintagma “o mesmo” .....</b>	<b>78</b>
<b>3 MÉTODO .....</b>	<b>84</b>
3.1 OBJETIVOS, PERGUNTAS E HIPÓTESES.....	85
3.2 PARTICIPANTES.....	90
<b>3.2.1 Contrapartida aos participantes: oficinas de escrita acadêmica.....</b>	<b>94</b>
3.3 INSTRUMENTOS E PROCEDIMENTOS DE COLETA DE DADOS .....	97
<b>3.3.1 Protocolos Verbais de Leitura .....</b>	<b>98</b>
3.3.1.1 <i>Procedimento de coleta de dados .....</i>	101
<b>3.3.2 Teste de Leitura Automonitorada.....</b>	<b>102</b>
3.3.2.1 <i>Procedimentos de coleta de dados .....</i>	104
<b>3.3.3 Teste de Julgamento de Aceitabilidade .....</b>	<b>109</b>
3.3.3.1 <i>Procedimento de Coleta de dados .....</i>	111
<b>3.3.4 Teste de Compreensão em leitura .....</b>	<b>113</b>
3.3.4.1 <i>Procedimento de coleta de dados .....</i>	122
3.4 TEXTO UTILIZADO NOS PROCEDIMENTOS: PROTOCOLO VERBAL E TESTE DE COMPREENSÃO EM LEITURA .....	123

3.5 ESTUDO PILOTO .....	129
<b>3.5.1 Principais resultados da pilotagem .....</b>	<b>131</b>
3.6 PROCEDIMENTOS GERAIS DA COLETA DE DADOS.....	132
3.7 CRITÉRIOS E PROCEDIMENTOS DE ANÁLISE DOS DADOS .....	134
<b>4 APRESENTAÇÃO, ANÁLISE E DISCUSSÃO DOS RESULTADOS .....</b>	<b>145</b>
4.1 DESCRIÇÃO E ANÁLISE DOS RESULTADOS .....	145
<b>4.1.1 O uso de “o mesmo” como anáfora correferencial implica maior custo no tempo de resposta? .....</b>	<b>145</b>
<b>4.1.2 Qual a relação entre desempenho local e tempo de resposta na leitura de “o mesmo” e de “ele” como anáforas correferenciais? Há diferenças na relação entre tempo de resposta de “o mesmo” e desempenho local e tempo de resposta de “ele” e desempenho local? .....</b>	<b>153</b>
<b>4.1.3 Qual a aceitabilidade do uso de “o mesmo” como elemento anafórico correferencial em construções sentenciais do PB? .....</b>	<b>157</b>
<b>4.1.4 Qual a relação entre o comportamento geral, o comportamento diante da anáfora correferencial construída com o sintagma “o mesmo” e o desempenho diante das retomadas anafóricas e do texto? .....</b>	<b>161</b>
<b>4.1.5 Que estratégias de leitura são utilizadas pelo leitor e qual a relação de tais estratégias com o desempenho diante das retomadas anafóricas e do texto? Ainda, é possível, no que diz respeito às estratégias empregadas pelos participantes, observar semelhanças frente às diferentes anáforas lidas? .....</b>	<b>172</b>
4.2 DISCUSSÃO DOS RESULTADOS .....	185
<b>4.2.1 <i>Outliers</i> e o N amostral: o que revelam sobre os resultados obtidos? .....</b>	<b>193</b>
<b>5 CONSIDERAÇÕES FINAIS.....</b>	<b>196</b>
<b>REFERÊNCIAS.....</b>	<b>201</b>
<b>APÊNDICE A – Instruções para a realização da técnica de Protocolo Verbal.....</b>	<b>213</b>
<b>APÊNDICE B – Instruções para a realização da leitura automonitorada .....</b>	<b>214</b>
<b>APÊNDICE C – Instruções para a tarefa de Julgamento de Aceitabilidade.....</b>	<b>215</b>
<b>APÊNDICE D – Estímulos utilizados no julgamento de aceitabilidade e na leitura automonitorada.....</b>	<b>216</b>
<b>APÊNDICE E – Instruções para o teste de Compreensão em Leitura.....</b>	<b>218</b>
<b>APÊNDICE F – Teste de Compreensão em Leitura gabaritado.....</b>	<b>219</b>
<b>APÊNDICE G – Texto utilizado nas tarefas de coleta de dados .....</b>	<b>227</b>

## 1 INTRODUÇÃO

Este é um trabalho cujo foco se volta para a leitura, pensada em uma perspectiva psicolinguística, com o intuito de se construir uma reflexão relevante dentro das especificidades que aqui serão apresentadas. As investigações sobre leitura trazem contribuições acerca de diferentes pontos, buscando alcançar as nuances necessárias para que ela aconteça. Neste estudo, olharemos para um aspecto relativo ao texto, considerando sua relação com o leitor: a anáfora correferencial realizada com o sintagma “o mesmo”. Estudamos, portanto, um aspecto concernente à leitura, mas considerando a amplitude que a engloba e os pontos que compõem os meandros do texto e que refletem na compreensão de um indivíduo em certo contexto e partindo de determinados objetivos.

Ocorre, então, que a leitura é o grande fio condutor deste estudo, mas ela sozinha promove reflexões inacabáveis. Por conta disso, é importante um afunilamento, que delimita o tema e, ao mesmo tempo, dá profundidade a ele. Nas linhas que seguem, as bases da leitura em uma perspectiva psicolinguística, juntamente com os eixos centrais deste estudo: compreensão leitora e processamento da anáfora, serão trazidas para a reflexão, com a finalidade de construir o caminho que orienta esta pesquisa.

Iniciamos pensando nas bases teóricas que já se apresentam no momento de abrir esta introdução: a Psicolinguística e a leitura. A Psicolinguística busca estudar o comportamento humano mediado pela linguagem. Mais especificamente, investiga os mecanismos cognitivos envolvidos na produção e na compreensão da linguagem (SCOVEL, 1998). Tais mecanismos podem ser estudados em uma perspectiva que parte da aquisição da linguagem, focando no entendimento sobre como uma língua é adquirida, especialmente a aquisição da língua materna pelas crianças; ou a partir de uma perspectiva que parte do processamento da linguagem, cujo foco se volta para a compreensão ou para a produção de uma língua oral, escrita ou visuoespacial. Nesse sentido, este estudo está inserido no rol de pesquisas que estudam o processamento na recepção escrita, com foco na compreensão em leitura, tratando-se, portanto, de um processo de segunda ordem, o qual é aprendido por meio de instrução formal (PINTO, 2019).

A perspectiva de leitura com a qual trabalhamos será mais bem delimitada em nosso referencial teórico. Por ora, cabe adiantar que ela parte de uma abordagem cognitiva e é entendida, para além da decodificação e dos processos iniciais relacionados à sua aprendizagem, como uma competência que envolve processos complexos e multicomponenciais com o objetivo de atingir alguma sorte de compreensão (SOUZA;

SEIMETZ-RODRIGUES; WEIRICH, 2019; OAKHIL; CAIN; ELBRO, 2017; DANEMANN, 1996; GOODMAN, 1991). Para esta investigação, focamos na leitura de texto acadêmico científico, contendo anáforas correferenciais realizadas com o elemento “o mesmo” e com o pronome “ele”. Na seção que segue, delimitamos e justificamos nossas escolhas para este estudo.

## 1.1 JUSTIFICATIVA E DELIMITAÇÃO DO ESTUDO

O foco da investigação está na compreensão de texto acadêmico-científico, o que nos levou a delimitar a área com a qual trabalhamos, pois os textos acadêmicos, ainda que tenham suas similaridades, como a linguagem própria da academia, a organização textual e estrutural, são diferentes nas mais diversas áreas acadêmicas e científicas; isso porque tratam de temáticas específicas, com terminologias próprias de cada área (SCHLICHTING, 2021) e com demandas específicas, já que os textos podem apresentar gráficos, cálculos, quadros, que devem ser integrados ao texto escrito – e a depender da área esses recursos podem ser mais ou menos utilizados.

Por isso, este estudo foca na leitura a partir de participantes de uma área específica, a saber: Engenharia de Controle e Automação, acadêmicos que estão em formação inicial de nível superior. O trabalho com a Engenharia se deu porque é uma área que tem se mostrado interessada em estudos que investiguem a leitura e a escrita (FRANZEN, 2012; FISCHER; HEINIG, 2014; SCHLICHTING, 2016; SCHLICHTING, 2021). Como essa área tende a se voltar majoritariamente para questões técnicas e matemáticas, é um campo em potencial para refletir sobre aspectos que envolvem a leitura a fim de se quebrar o senso comum que diz que engenheiro não lê. Desse modo, além de estudar a leitura em si, buscamos, também, estabelecer um encontro e um diálogo, viabilizado pela leitura, entre o campo de conhecimento da Linguística e da Engenharia.

A leitura tem sido alvo de investigação nos mais diversos campos, e as pesquisas abordam a temática a partir de diferentes perspectivas. É um tema caro, que requer um olhar atento. Essa importância pode ser constatada quando são verificados os resultados de diferentes testes com foco em leitura. Um deles, realizado internacionalmente e cujos dados são comparados entre os diversos países participantes, é o PISA (Programa Internacional de

Avaliação de Estudantes). A despeito de todas as críticas<sup>1</sup> muitas vezes lançadas ao Programa (BONNET, 2006; BART; DAUNAY, 2018; DAUNAY; BART, 2019), duas questões devem ser ressaltadas: 1. A avaliação realizada é cuidadosamente elaborada, com critérios claros e sustentação teórica consistente. É possível que ela não dê conta de alcançar todas as especificidades educacionais e populacionais de cada país participante; entretanto, propicia um panorama importante do conhecimento dos alunos especialmente em três áreas: Leitura, Ciências e Matemática. 2. Os índices de leitura dos estudantes brasileiros estão muito abaixo do esperado. Os primeiros testes realizados pelo PISA (BRASIL, 2001) já mostravam uma carência no que tange à competência leitora de quem está em fase final da educação básica, normalmente ingressante no Ensino Médio.

Esses dados não têm se mostrado muito diferentes nos últimos anos, e o Brasil continua não apresentando avanços significativos em leitura. De modo mais específico, no teste realizado em 2018, o Brasil manteve sua posição no que diz respeito à leitura e segue atrás de mais de 50 países. Os resultados têm apontado para o fato de que os estudantes brasileiros<sup>2</sup> apresentam dificuldades em múltiplos aspectos, dentre os quais compreender texto mais longos, comparar diferentes ideias, fazer inferências a partir de distintas fontes etc. O que os resultados obtidos sucessivamente nos testes realizados revelam é a existência de uma defasagem em níveis que requerem maior elaboração para a compreensão em leitura. Nesse sentido, ressaltamos o papel essencial de investigações que têm por foco a leitura, a fim de que se possa construir subsídios que auxiliem a modificar esse panorama e a contribuir com o processo democrático desse país, “criando condições, em aspectos com quais se trabalha detidamente, para a redução da desigualdade social no que diz respeito ao acesso à cultura letrada.” (FRANZEN; SOUZA, 2020, p.26).

Diante desses aspectos, o estudo da leitura se mostra relevante para todo e qualquer público. Afinal, não é suficiente apenas, ano após ano, ressaltar os baixos resultados dos brasileiros em diversas avaliações. É necessário conhecer e compreender os processos envolvidos na leitura de diferentes textos, feita com distintos objetivos e em variadas áreas para, então, poder refletir sobre os possíveis caminhos a se seguir.

---

<sup>1</sup> Uma das principais críticas está no fato de a prova não considerar especificidades culturais, regionais e socioeconômicas de quem responde ao teste. Além disso, há o fato de a prova se preocupar em demasia com a preparação do estudante para o mercado de trabalho e não para atuar em uma sociedade de forma crítica e democrática. Outra crítica, essa mais direcionada à avaliação da leitura em si, está no tratamento dado ao texto literário e a forma de abordá-lo na prova (DAUNAY; BART, 2019).

<sup>2</sup> Optamos, com base em Schwidnt (2018; 2020), por empregar o gênero masculino, compreendendo este como forma não marcada da língua portuguesa. Essa compreensão está pautada nas discussões de gênero gramatical, não contemplando as reflexões de gênero social, que entendemos fundamentais, mas extrapolam os limites desta investigação.

Ao tratarmos de leitura, precisamos, ainda, considerar a compreensão, pois, não basta ler; é preciso compreender. Este estudo entende que a compreensão é construída à medida que os processos suscitados pelas informações contidas no texto interagem em seus distintos níveis e, cooperativamente, são integrados ao conhecimento do leitor, que, então, forma uma representação mental do texto (KINTSCH; RAWSON, 2013; FOX; ALEXANDER, 2009; KINTSCH, 1998). Isso significa que a compreensão em leitura será resultante de uma combinação de fatores, que envolvem leitor, texto e situação de leitura.

Quando se trata de compreensão em leitura, existem diversas nuances de estudo. É possível, por exemplo, refletir sobre a influência da estrutura sintática no processamento das informações, pensar sobre o acesso lexical do texto lido, analisar o papel da metáfora na compreensão etc. Enfim, são diferentes olhares a serem lançados para desenvolver subsídios que deem suporte aos processos envolvidos em uma leitura bem-sucedida com alguma sorte de compreensão. Desde os níveis mais iniciais, imbricados no processo de decodificação de palavras, até os níveis superiores, como a realização de inferências e o monitoramento da compreensão, há muitas questões a serem investigadas. Fato é que, para que haja compreensão do texto lido, não basta ao leitor compreender palavras ou frases isoladas; é necessário que as partes do texto sejam integradas e organizadas, a fim de que se construa um todo coerente, que leve o leitor a atingir o objetivo traçado para aquela leitura.

Dito isso, o presente trabalho estuda a leitura focando em um nível que envolve a construção coesiva de um texto e que é um dos responsáveis pela integração das diferentes partes que o constituem, proporcionando cadência e continuidade à leitura, participando da formação da estrutura local, chamada de microestrutura textual<sup>3</sup>, e tendo papel fundamental na construção da base textual<sup>4</sup> e do modelo situacional<sup>5</sup> (KINTSCH, 1998; KINTSCH; RAWSON, 2013), pois opera na integração do todo textual: as anáforas correferenciais<sup>6</sup> (todos esses conceitos serão mais bem explicados no capítulo da fundamentação teórica). Assim, no que

---

<sup>3</sup> Um dos níveis de processamento envolvido na compreensão de um texto é o da análise semântica, que determina o significado do texto. Na análise semântica, os significados das palavras são combinados, formando proposições, que se interrelacionam em uma microestrutura – a menor estrutura de um texto. A formação de uma microestrutura coerente é o primeiro passo para a compreensão textual (KINTSCH; RAWSON, 2013; KINTSCH, 1998; KITSCH, VAN DIJK, 1978).

<sup>4</sup> O significado do texto é dado pela base textual, que é formada a partir do que foi construído na microestrutura e na macroestrutura – a macroestrutura, por sua vez, se forma a partir da integração de porções maiores do texto (KINTSCH; RAWSON, 2013; KINTSCH, 1998; KITSCH, VAN DIJK, 1978).

<sup>5</sup> Um modelo situacional é formado a partir da integração entre as informações trazidas pelo texto e o conhecimento prévio relevante do leitor. A compreensão profunda depende, também, da formação desse modelo situacional (KINTSCH; RAWSON, 2013; KINTSCH, 1998; KITSCH, VAN DIJK, 1978).

<sup>6</sup> Neste trabalho, sempre que nos referirmos ao objeto investigado, usamos o termo “anáfora correferencial”, que será mais bem detalhado no capítulo destinado à fundamentação teórica.

tange à leitura, nosso olhar se volta para o processamento das construções anafóricas, que dão encadeamento e progressão ao texto e que, normalmente, são processadas automaticamente, exercendo papel fundamental na compreensão. Nesta tese, o foco está na anáfora que recupera um antecedente determinado e explícito no texto, buscando, mais especificamente, pela influência da anáfora realizada com “o mesmo” na compreensão em leitura. Na compreensão, as relações anafóricas atuam de modo a dar progressão ao texto, garantindo que as conexões sejam feitas e mantendo determinados referentes ativos na memória do leitor (KINTSCH; RAWSON, 2013; JUST; CARPENTER, 2013; KOCH; ELIAS, 2015; OAKHILL; CAIN; ELBRO, 2017).

Diversos estudos têm sido realizados com o intuito de investigar as diferentes formas anafóricas, as distintas relações semânticas e as diferentes construções sintáticas que operam no processamento da anáfora (LEITÃO, 2005; QUEIROZ, 2009; MAIA, 2013; TEIXEIRA, 2013; NÓBREGA-LIMA, 2014, dentre outros). Esta, além disso, relaciona-se diretamente a outros aspectos da cognição, como a memória, sobretudo a memória de trabalho<sup>7</sup>, pois o papel da anáfora é recuperar uma informação já introduzida previamente e, para que essa recuperação ocorra de forma exitosa, o indivíduo precisa ter tal referente ativo em sua memória, com a finalidade de processar corretamente a informação. A anáfora é uma das estratégias utilizadas para manter ativo um determinado “objeto” já apresentado. Para isso, podem ser usados diferentes recursos linguísticos, como pronomes plenos ou nulos, a repetição do próprio termo já citado, hipônimos, hiperônimos ou sinônimos. Isso quando se trata de anáforas diretas, foco deste estudo. Dentre essas formas, cada uma com suas especificidades, chama a atenção o uso de um elemento específico como anafórico. Elemento esse que tem aparecido com frequência no tipo de texto com o qual se está trabalhando nesta investigação (texto acadêmico-científico) e que assume a função do pronome, mas que não se caracteriza como tal. É o elemento “o mesmo”<sup>8</sup>, em construções como:

O currículo X foi implantado progressivamente, a partir do 1º semestre de 2006. **O mesmo** substituiu o currículo Y, com a troca de disciplinas fase a fase.

<sup>7</sup> Aqui, consideramos a Memória de Trabalho, porque compreendemos que esta é a responsável por manter uma informação ativa ao mesmo tempo em que ela é processada. Ou seja, não basta apenas armazenar a informação em uma memória de curta duração, mas é necessário que tal informação seja processada e então relacionada a informações armazenadas na memória de longa duração.

<sup>8</sup> Neste trabalho, sempre que utilizamos a expressão “o mesmo” incluímos todas as variações de gênero e número: a mesma, os mesmos, as mesmas.



Nesse contexto, “o mesmo” exerce a função de anáfora correferencial, retomando seu antecedente “O currículo X”. O anafórico utilizado poderia ser substituído pelo pronome “ele”:

O currículo X foi implantado progressivamente, a partir do 1º semestre de 2006. **Ele** substituiu o currículo Y, com a troca de disciplinas fase a fase.

O pronome de terceira pessoa, como será discutido de modo mais aprofundado no capítulo de fundamentação teórica, é um elemento característico para a realização de uma anáfora correferencial, mas “o mesmo” também tem sido utilizado. Encontramos esse uso em diversos contextos escritos, sobretudo em textos acadêmicos. A título de ilustração, trazemos exemplos extraídos de textos de pesquisadoras da Linguística: Sicuro-Corrêa (2006, p. 23) e Kleiman (2016, p.17), respectivamente:

Quando dados da aquisição da linguagem são explicitamente considerados na construção de um modelo de língua, estes são obtidos em um estado consideravelmente avançado do processo de aquisição, quando a criança já produz enunciados reconhecidamente linguísticos e o tratamento **dos mesmos** não é mediado por uma teoria do processamento linguístico que permita relativizar conclusões acerca do estado de conhecimento da língua por parte da criança ou embasar explicações que remetam a questões de desempenho.

Encontra-se, no final de cada capítulo, um pequeno resumo **dos mesmos** [...]

Isso mostra que o uso desse elemento é feito por autores renomados e que possuem o domínio de gêneros acadêmicos. O que ocorre é que, gramaticalmente, “o mesmo” não tem essa função; por isso, o seu uso como anáfora correferencial acaba gerando questionamentos e reflexões por alguns estudiosos da língua (BAGNO, 2012; SACCONI, 1994; ALMEIDA, 2009). No entanto, ele existe e acontece com bastante frequência, inclusive em textos acadêmicos formais que passaram por avaliação cega de pares, como apresentado acima.

Diante disso, questiona-se sobre os custos dessa construção anafórica para a leitura. Isso porque a leitura dessas anáforas responsáveis pela coesão local em um texto tende a ocorrer de modo automático em leitores hábeis. Esses mesmos leitores hábeis usam o raciocínio analítico quando encontram algum problema no texto (KINTSCH; RAWSON, 2013). Esta investigação, portanto, busca, por meio de diferentes técnicas (protocolos verbais, leitura automonitorada, teste de compreensão em leitura com gravação de tela e julgamento de aceitabilidade), verificar se o processamento de “o mesmo” segue padrão de processamento semelhante ao do pronome “ele” e, ainda, lança um olhar para as estratégias de leitura e para o desempenho em teste de

compreensão para saber se o fluxo da leitura e o seu produto tende a ser interrompido ou prejudicado quando essa forma anafórica é encontrada ou se ela é tida como natural por leitores hábeis da língua portuguesa, cumprindo o seu papel na integração textual. A seguir, esses aspectos a serem alcançados serão mais bem detalhados por meio dos objetivos, das perguntas e das hipóteses que orientam esta pesquisa.

## 1.2 OBJETIVOS, PERGUNTAS E HIPÓTESES

Diante de todo o apresentado, o grande **objetivo** do presente estudo de tese está em compreender de que forma e em que medida “o mesmo” usado como anáfora correferencial influencia no processamento e na compreensão em leitura de texto acadêmico-científico. A partir desse objetivo geral, são formulados os cinco objetivos específicos:

1. Conhecer o custo de tempo de resposta diante do sintagma “o mesmo” como anáfora correferencial.
2. Explorar o tempo e a acurácia diante de anáforas correferenciais de forma a relacioná-las a estratégias utilizadas e ao desempenho em leitura de texto acadêmico-científico.
3. Identificar a aceitabilidade do uso de “o mesmo” como elemento anafórico correferencial em construções sentenciais do PB.
4. Explorar o desempenho e o comportamento estratégico do leitor diante do texto e do sintagma “o mesmo”.
5. Caracterizar e compreender as estratégias utilizadas durante a leitura de texto acadêmico-científico, no que tange à anáfora, e a relação de tais estratégias com a resolução bem-sucedida dessa anáfora.

Buscamos, portanto, conhecer o tempo de resposta diante da anáfora correferencial utilizada, o comportamento diante do texto, o desempenho na compreensão em leitura e as estratégias empregadas durante a leitura de texto que contém essas anáforas. Isto é, refletimos sobre o processo e, ainda, sobre o produto da leitura. Aqui, a compreensão e o processamento são considerados a partir do desempenho, do comportamento, do tempo de resposta e das estratégias.

O grande **problema** desta investigação se volta, portanto, para: de que forma e em que medida o sintagma “o mesmo”, como anáfora correferencial, atua na integração local, na formação da base textual e na compreensão em leitura de texto acadêmico-científico? A **hipótese geral** formulada para o estudo prevê que: “o mesmo” segue um padrão de

processamento e de integração semelhante ao do pronome “ele”. Assim, espera-se que os resultados em termos de processamento mostrem que “o mesmo” pode exercer a função de anáfora correferencial, sem que a compreensão em leitura seja dificultada por maiores custos para o processamento.

De modo mais específico, têm-se as seguintes perguntas e hipóteses, que são apresentadas neste momento e retomadas de modo mais detalhado no capítulo 3 do método:

Pergunta 1: O uso de “o mesmo” como anáfora correferencial implica maior custo no tempo de resposta?

Hipótese a: Os custos de tempo de resposta não são afetados pelo uso de “o mesmo” como anáfora correferencial. Ou seja, o custo de tempo de resposta de “o mesmo” é semelhante ao custo de tempo de resposta do pronome ele.

Pergunta 2: Considerando que o desempenho local do participante é semelhante ao desempenho geral, qual a relação entre desempenho local e tempo de resposta na leitura de “o mesmo” e de “ele” como anáforas correferenciais? Há diferenças na relação entre tempo de resposta de “o mesmo” e desempenho local e tempo de resposta de “ele” e desempenho local?

Hipótese b: O tempo de resposta na leitura de “o mesmo” afeta positivamente o desempenho local, assim como ocorre com o tempo de resposta na leitura de “ele”. Isto é: quanto menor o tempo de resposta na leitura desses dois anafóricos, melhor o desempenho em leitura local do texto.

Pergunta 3: Qual a aceitabilidade do uso de “o mesmo” como elemento anafórico correferencial em construções sentenciais do PB?

Hipótese c: A aceitabilidade do uso de “o mesmo” não é diferente da aceitabilidade do uso de “ele” como anafóricos.

Pergunta 4 (exploratória): Qual a relação entre o comportamento geral, o comportamento diante da anáfora correferencial construída com o sintagma “o mesmo” e o desempenho diante das retomadas anafóricas e do texto?

Pergunta 5 (exploratória): Que estratégias de leitura são utilizadas pelo leitor e qual a relação de tais estratégias com o desempenho diante das retomadas anafóricas e do

texto? Ainda, é possível, no que diz respeito às estratégias empregadas pelos participantes, observar semelhanças frente às diferentes anáforas lidas?

Para atingir os objetivos, responder às perguntas e checar as hipóteses, lançamos mão de técnicas comportamentais, que têm por foco captar a compreensão como processo e como produto. Os dados foram coletados a partir de quatro instrumentos, a saber: Protocolo Verbal (PV), Leitura Automonitorada (LA), Teste de Aceitabilidade (TA) e Teste de Compreensão em Leitura (TCL) aplicado em computador com gravação da tela, captando os movimentos de mouse e de teclado realizados pelos participantes. A combinação de diferentes técnicas tem papel significativo na pesquisa científica, a fim de que se possa contornar quaisquer limitações que uma determinada técnica possa apresentar (SOUZA; FRANZEN; SCHLICHTING, 2019; TOMITCH, 2008).

No capítulo que apresenta e explica o método, os objetivos, as perguntas e as hipóteses serão retomados e cruzados com os instrumentos que auxiliaram a construir respostas às indagações do estudo. Assim, estudamos a compreensão em leitura, focando na atuação de anáforas correferenciais, como elementos coesivos, em textos acadêmico-científicos da área específica dos participantes da pesquisa, qual seja, Engenharia de Controle e Automação. Com os delineamentos feitos até aqui, o leitor deste trabalho já conhece o lugar de onde falamos e, de modo amplo, qual o nosso objeto de estudo. Para encerrarmos esta primeira parte, trazemos o mapa deste texto, apresentando a sua organização.

### 1.3 ORGANIZAÇÃO DA TESE

Após esta introdução, são apresentados os fundamentos que dão base para esta pesquisa. Inicia-se com a compreensão de leitura que orienta esta tese, buscando luz nas diferentes teorias desenvolvidas acerca dos modelos de leitura construídos ao longo de anos de pesquisa na área. Em seguida, é realizada uma discussão sobre referenciação e seus processos a partir das perspectivas da linguística. Com isso, chega-se à anáfora correferencial e seu papel na compreensão em leitura. Para aprofundar essas reflexões, são trazidos os resultados de estudos que têm investigado o processamento anafórico. Desse modo, é possível relacionar teorias de compreensão em leitura com resultados de estudos que se centram fundamentalmente no processamento da anáfora e nas variáveis que podem atuar em tais processos. Fazer essa relação é importante para que, na sequência, possamos discutir sobre as variáveis com as quais essa pesquisa está trabalhando e, assim, entender sua atuação na compreensão em leitura.

Após a apresentação da base teórica, no capítulo 3, os objetivos, as perguntas e as hipóteses da pesquisa são detalhados para, então, discorrer sobre os fundamentos metodológicos e o método que orientou a experimentação realizada, apresentando os participantes da pesquisa e especificando as quatro técnicas utilizadas na coleta de dados, a saber: protocolos verbais, leitura automonitorada, teste de aceitabilidade e teste de compreensão em leitura com gravação de tela. Após os instrumentos terem sido apresentados, discute-se o texto utilizado nos testes aplicados e se abre uma seção para tratar do estudo piloto realizado. Na sequência, são descritos os procedimentos gerais de coleta de dados. Para finalizar, são trazidos os critérios e procedimentos que guiam a análise quantitativa e qualitativa dos dados obtidos a fim de explicitar como o capítulo de descrição, análise e discussão foi construído.

Desenvolvidos esses pontos, os resultados estatísticos são apresentados. Para tanto, optou-se por criar seções de acordo com cada pergunta de pesquisa que buscamos responder. Nesse sentido, os resultados são apresentados em cinco subseções, nas quais detalhamos questões específicas de cada teste e, também, dos entrecruzamentos feitos para que fosse possível responder às perguntas que guiam esta investigação. Feito isso, trazemos as discussões dos resultados obtidos. Nessa seção, buscamos responder à grande questão investigativa, relacionando os resultados obtidos nas respostas a cada pergunta específica e refletindo sobre nuances acerca deles.

O texto se encerra, então, com as considerações acerca de todo o processo investigativo construído e dos resultados a que chegamos acerca da compreensão em leitura de texto acadêmico-científico que contém anáforas correferenciais realizadas com o sintagma “o mesmo”.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, são trazidos os fundamentos que sustentam as discussões e reflexões realizadas nesta tese. Para compreender o ponto sobre o qual a pesquisa se debruça, é preciso conhecer o que tem sido postulado sobre leitura na área em que este estudo se insere, a Psicolinguística. Afinal, é ela que dá base a esta investigação. Por isso, em primeiro lugar, apresentamos a concepção de leitura que guia o trabalho, discorremos sobre as três abordagens de processamento de leitura mais proeminentes e apresentamos alguns aspectos relevantes sobre a memória. Tudo isso com o objetivo de discutir sobre a construção<sup>9</sup> da compreensão em leitura. Na sequência, ponderamos sobre o que a linguística textual postula acerca da referenciação e de seus processos, dando ênfase às anáforas, elemento central neste estudo. Com isso elaborado, refletimos sobre o papel das anáforas na compreensão em leitura, sobre os estudos acerca do processamento anafórico e, então, chegamos à variável-alvo da presente pesquisa, o sintagma “o mesmo”. Desse modo, os fundamentos estão organizados em torno de dois grandes eixos, a saber: a compreensão em leitura e a anáfora.

### 2.1 A LEITURA

Ao estudar a leitura, há diferentes aspectos que podem ser considerados, tais como aspectos cognitivos, sociais e culturais. Como já citado, este trabalho se volta para os aspectos cognitivos envolvidos nela. Diante disso, a leitura, nesta tese, é entendida e estudada como uma competência que envolve processos cognitivos complexos e que, de acordo com Danemann (1996, p.513, tradução nossa), “consiste na execução coordenada de processos oculomotores, perceptuais e de compreensão.”<sup>10</sup> Essa definição não esgota aquilo que entendemos por leitura, mas é o ponto de partida para conhecer o lugar de onde falamos. Nesse lugar, entendemos que a leitura envolve compreensão, ou seja, a concretização da leitura acontece quando a compreensão é construída. Essa construção se dá de modo bastante complexo por meio de estratégias e conhecimentos que refletem em processos cognitivos que ocorrem antes, durante e depois da leitura. Nessa complexidade, estão englobados fatores tais quais:

---

<sup>9</sup> Utilizamos o termo “construção” no sentido de processo, ou seja, a construção da compreensão se dá a partir da interação de diferentes e diversos processos complexos que envolvem aspectos linguísticos, mas, também, outros aspectos cognitivos (memória, atenção, percepção, motivação etc). Dessa forma, a compreensão depende da interrelação de processos que ocorrem a partir dos diferentes níveis linguísticos presentes em um texto e não se trata de um ponto a que se chega, mas uma representação que se desenvolve.

<sup>10</sup> Reading is a complex cognitive skill, consisting of the coordinate execution of a collection of oculomotor, perceptual, and comprehension processes”.

1. Os objetivos que são previamente estabelecidos para a leitura: processos que ocorrem antes da leitura e que dizem respeito ao leitor, é o que incita a leitura.
2. Aquilo que é fornecido pelo texto: estímulo para os processos que ocorrem durante a leitura.
3. A representação que se constrói a partir dos conhecimentos prévios relevantes: representação esta que envolve a intersecção texto-leitor e que é construída no processo, mas que pode continuar a ser (re)formulada após a leitura, enquanto o leitor “reconsidera e reconstrói o que foi compreendido” (GOODMAN, 1991 p.33).
4. O contexto ou a situação na qual a leitura acontece.

Esses são alguns dos aspectos, dentre tantos outros (SOLÉ, 1998), envolvidos na compreensão, que comporta produto e processos e, no que tange aos processos, é multicomponencial. Cabe, assim, esclarecer que a leitura tem por grande finalidade a compreensão daquilo que é lido (OAKHILL; CAIN; ELBRO, 2017). Compreender o que é lido, por sua vez, é uma competência a ser desenvolvida. Tal afirmação implica o entendimento de diferentes questões, a saber: O que entendemos por leitura? O que está envolvido na compreensão em leitura? Por que apenas saber ler palavras e frases não resulta no entendimento daquilo que é lido? Como se chega à compreensão do texto? Busca-se discutir essas questões ao longo desta seção a fim de que se possa construir uma representação mental acerca da temática central deste estudo.

Assim, a leitura é, conforme Souza, Seimetz-Rodrigues e Weirich (2019, p.166), “uma competência desenvolvida a partir da relação que o leitor (que sempre se encontra em processo de formação e de novas aprendizagens) estabelece com o texto escrito, com vistas a alcançar determinado objetivo, relativo a alguma compreensão, em situação e contexto específicos”. Então, para entender como a leitura acontece, é preciso refletir sobre as peças que, ao atuarem em conjunto, formam esse complexo quebra-cabeças: o texto, o leitor e a situação de leitura. Para explicar cada uma dessas peças, faremos uma breve separação para que se possa realizar uma análise mais minuciosa, conhecer suas especificidades e entender como elas se encaixam, porque, no fim das contas, a leitura só acontece com a combinação de todas elas.

### **2.1.1 O texto**

O processo da leitura só é possível com a presença de um material escrito. O texto é o estímulo, constituído de diferentes níveis linguísticos que deverão ser processados pelo leitor

para que a compreensão seja construída. O texto, pensado de modo isolado, é compreendido como uma sequência linguística organizada de modo coerente (KOCH; TRAVAGLIA 2015), isto é, a coerência estabelecida entre os elementos linguístico é que gera a textualidade e faz de tal sequência um texto. Essa sequência linguística se organiza a partir de níveis linguísticos. A compreensão, então, é construída “com base nos elementos linguísticos presentes na superfície textual” (KOCH; ELIAS, 2015, p.57).

De acordo com Kleiman (2009, p. 45), o texto é compreendido “como uma unidade semântica onde os vários elementos de significação são materializados através de categorias lexicais, sintáticas, semânticas, estruturais”. Assim, o texto, a partir desses diferentes níveis articulados de modo coerente e coeso, fornece ao leitor os subsídios para que processos cognitivos ocorram a fim de se atingir alguma sorte de compreensão. O significado do texto é, portanto, construído e representado cognitivamente. Cada um dos níveis linguísticos reflete em processos cognitivos específicos e articulados, por isso, cada um deles exerce um papel na construção da compreensão.

A representação do texto inicia com uma ideia (SCLIAR-CABRAL, 2018), que será materializada a partir da codificação dos fonemas em grafemas, que, unidos, formam os itens lexicais. Ocorre que, para que um texto seja formado, não se pode simplesmente dispor itens lexicais lado a lado, é preciso que esses itens estejam organizados respeitando a sintaxe da língua. Todavia, para formar um texto, também não basta dispor, lado a lado, frases que respeitem a sintaxe dessa língua. Essas frases precisam estar conectadas de modo articulado, desenvolvendo um todo semanticamente coerente. Isso significa que um texto é formado por níveis linguísticos distintos (fonológicos, morfológicos, lexicais, sintáticos, semânticos, discursivos, pragmáticos etc), mas que são interdependentes. Então, no texto, existe uma estrutura linguística que diz respeito às relações locais, que envolvem aspectos morfológicos, lexicais, frasais e as relações interfrasais locais, e uma estrutura linguística formulada por meio da coerência estabelecida pelas conexões entre os parágrafos e o todo em um texto. A partir dessas informações contidas no texto, o leitor constrói, no momento da leitura, a microestrutura e a macroestrutura de tal texto (KINTSCH, 1998; KINTSCH, 2013; KINTSCH; RAWSON, 2013). Além disso, há que se considerar, ainda, que um texto é escrito por um autor, em um espaço-tempo específicos, com um objetivo específico. Isso também refletirá no todo do texto e nos sentidos expressos pelos recursos linguísticos utilizados. Esses aspectos que formam um texto são relevantes para entender os processos que ocorrem no momento da leitura de um texto, isso porque esses diferentes níveis linguísticos refletem em diferentes níveis de processamento.



Para este estudo, interessam dois níveis distintos de processamento: um nível microestrutural, que diz respeito à relação interfrasal, atuando na coesão local, que é a anáfora correferencial, e um nível macroestrutural<sup>11</sup>, que se relaciona à coerência global, pois olhamos, também, para a compreensão do texto (não somente para o processamento da anáfora). Tais níveis provêm do estímulo (o texto) e refletem cognitivamente (no leitor) durante a leitura. Isso porque os níveis de processamento estão diretamente relacionados aos níveis que caracterizam o estímulo que vai ser processado. Então, além de pensar acerca dos aspectos cognitivos envolvidos na leitura, precisamos, também, analisar os aspectos relativos ao texto. Para tanto, são mobilizadas teorias provindas da Linguística Textual (KOCH; ELIAS, 2018; KOCH; ELIAS, 2015; KOCH; TRAVAGLIA, 2015; KLEIMAN, 2009; MARCUSCHI, 2001) em diálogo com as bases fornecidas pela Psicolinguística. Isso porque a Linguística Textual provê subsídios para entender os níveis linguísticos envolvidos na construção do texto, nos seus elementos coesivos e na tecitura formada a partir de tudo que o envolve. Assim, quando discutirmos aspectos relativos à referenciação e à anáfora propriamente dita, a Linguística Textual é chamada novamente para a cena.

### 2.1.2 O leitor

O processo de leitura é concretizado e construído, como vimos, a partir de um texto. No entanto, um texto sozinho é inerte. Assim, o leitor é peça central no quebra-cabeça da compreensão, porque é ele quem orquestra toda a leitura (SOUZA, 2012). O texto oferece o *input* que, ao ser compreendido, é transformado em *intake*<sup>12</sup> e manipulado pelo leitor a partir de seus conhecimentos prévios, seus objetivos, suas experiências linguísticas e de mundo. Durante a leitura, o leitor mobiliza diferentes conhecimentos adquiridos ao longo de sua vida (conhecimentos linguísticos, textuais, conhecimento de mundo, dentre outros) (KLEIMAN, 2009) para que possa processar a informação fornecida pelo texto e, a partir disso, construir uma base semântica significativa do todo. Isso ocorre por meio dos processos cognitivos que são ativados de acordo com os níveis linguísticos acessados a partir do texto.

---

<sup>11</sup> A macroestrutura de um texto é formada a partir da integração de porções textuais maiores, identificando os pontos centrais do texto (KITSCH, VAN DIJK, 1978; KINTSCH, 1998; KINTSCH; RAWSON, 2013). O nível macroestrutural será mais bem explicado na seção 2.1.4.

<sup>12</sup> O *intake* é o processo que ocorre quando o *input* é compreendido, ou seja, o leitor não atua diretamente sobre o estímulo, mas sobre os processos que ocorrem a partir daquele estímulo. É, portanto, um estágio entre o *input* e a compreensão final, estágio em que o leitor se apropria e, assim, processa o que está lendo. (SILVA, 1999; KUMARAVADIVELU, 1994).

A leitura e a compreensão serão direcionadas pelo leitor e orientadas por seus objetivos. Assim, a compreensão do texto lido está diretamente relacionada às intenções do leitor. É o leitor quem vai construir os sentidos e estabelecer a coerência para o material lido (KOCH; ELIAS, 2015). Tudo isso, considerando aquilo que está sendo apresentado no texto a partir dos recursos linguísticos utilizados e das conexões efetuadas.

Assim, para que a leitura aconteça, é necessário que o leitor disponha de uma série de conhecimentos declarativos, condicionais e procedimentais, que envolve um conjunto de habilidades e estratégias (GAGNÉ; YEKOVICH; YEKOVICH, 1993; PARIS; LIPSON; WIXSON, 1994; SOUZA; SEIMETZ-RODRIGUES; WEIRICH, 2019) usadas para que seja possível construir os sentidos necessários à compreensão do texto lido. Dos conhecimentos declarativos, fazem parte os linguísticos, os textuais, os do conteúdo tratado e os conhecimentos de mundo. Os conhecimentos condicionais estão relacionados aos conhecimentos do leitor sobre o contexto e as intenções (HEINIG, 2019; PARIS; LIPSON; WIXSON, 1994) para a leitura em um determinado espaço social, ou seja, o conhecimento condicional do leitor faz com que ele adeque seus conhecimentos declarativos e procedimentais (RUDELLE; UNRAU, 2013; HEINIG, 2019; PARIS; LIPSON; WIXSON, 1994). Já as habilidades estão ligadas a processos cognitivos que tendem a ser mais automáticos, como os aspectos relativos à decodificação, ao reconhecimento de palavras, à computação sintática, aos mecanismos de articulação e coesão textual (MANOLI; PAPADOPOULOU, 2012; SOUZA; SEIMETZ-RODRIGUES; WEIRICH, 2019). Em um *continuum* com os processos automáticos, estão os processos controlados, dos quais se faz uso de modo deliberado a partir da metacognição. As estratégias, que serão abordadas na seção seguinte, são ações feitas deliberadamente com a finalidade de atingir ao objetivo que se tem com a leitura efetuada.

Consideramos, então, como vimos na seção anterior, que o texto é formado por diferentes níveis linguísticos que refletem nos níveis de processamento. Assim, a leitura acontece a partir de diferentes processos, que podem ser compreendidos em níveis: inferiores e superiores. Em primeiro lugar, o leitor precisa ter fluência - quando a leitura de palavras e frases ocorre com velocidade e de forma acurada. Ou seja, são considerados proficientes aqueles leitores que possuem processos automatizados (LABERGE; SAMUELS, 1974) e conseguem, portanto, identificar as palavras de modo rápido e preciso (ADAMS, 2013), liberando, assim, recursos para que o processamento geral da informação ocorra sem entraves, o que faz o código não ser um obstáculo. Nesse sentido, “leitores fluentes são capazes de processar texto de forma

automática e simultaneamente compreender e refletir sobre as características sintáticas e semânticas do texto lido com prosódia<sup>13</sup>” (RASINSKI et al, 2011, p. 286-287, tradução nossa).

No entanto, embora a automaticidade no reconhecimento das palavras seja necessária, ela não é suficiente para que um leitor seja proficiente (DANEMANN, 1996). Ao se considerarem leitores que já possuem essa automaticidade no nível do reconhecimento das palavras e frases, o foco se volta para os níveis superiores envolvidos na leitura. Nessa perspectiva, os leitores precisam ser estratégicos, devem ser capazes de adequar a sua leitura de acordo com os objetivos a serem alcançados (SOLÉ, 1998), construindo a base textual e, quando o objetivo é a compreensão, formando um modelo situacional coerente do lido. Para tanto, é necessário que se integrem de modo adequado as diferentes partes de um texto, que se conectem as diferentes informações apresentadas, que se consiga relacionar diferentes anáforas com seus antecedentes, que se realizem diferentes tipos de inferências (conectivas, elaborativas e lógicas), relacionando as informações fornecidas pelo texto ao conhecimento prévio em torno do tópico (DANEMANN, 1996). O leitor proficiente é, então, ativo, estratégico, flexível e sabe se autorregular. Esse leitor pode até encontrar entraves durante a leitura, mas, quando consegue resolvê-los, o faz de modo rápido e adequado sem prejudicar a compreensão e, quando não consegue resolvê-los, sabe que não conseguiu e, em geral, sabe o porquê. Kleiman (2016) denomina esse leitor de experiente (chamado, neste trabalho, de proficiente) e cita duas características básicas dele, que resumem muito do que já apresentamos:

primeiro, ele lê porque tem algum objetivo em mente, isto é, sua leitura é realizada sabendo para que está lendo, e, segundo, ele compreende o que lê, o que seus olhos percebem seletivamente é interpretado, recorrendo a diversos procedimentos para tornar o texto inteligível quando não consegue compreender (KLEIMAN, 2016, p.76).

Dessa forma, a construção da compreensão se dará levando em conta o objetivo traçado pelo leitor, e é a partir desse objetivo que ele monitora sua leitura e seleciona as estratégias que o auxiliarão nesse intento. Ainda, quando a autora cita “procedimentos”, está se referindo às estratégias a serem utilizadas durante a leitura. Sobre elas, falaremos na próxima seção.

### *2.1.2.1 As estratégias de leitura*

Como ressaltado, o leitor é uma das peças centrais para que a leitura ocorra. Durante o ato de ler, há uma interação do leitor sobre o texto. Isso envolve um leitor ativo (SOLÉ, 1998),

---

<sup>13</sup> “[...] fluent readers are able to simultaneously automatically process text and simultaneously understand and reflect the syntactic and semantic features of the text by reading with prosody.”

que se debruça sobre o que está lendo de modo a construir a compreensão e atingir o objetivo estabelecido para aquela leitura. Discutimos, ainda, sobre o fato de o leitor proficiente não ser somente aquele que lê de maneira fluente, mas, também, aquele que sabe usar de estratégias específicas para atingir ao seu objetivo. É válido, então, refletirmos sobre o que são estratégias a fim de que possamos entender os movimentos possíveis do leitor e, ainda, no momento de análise dos dados desta pesquisa, compreender as estratégias usadas pelos participantes diante do texto que leram.

Quando tratamos de estratégias, saímos do nível da cognição e entramos no nível da metacognição. Para apresentar a concepção de estratégia abordada neste estudo, é preciso, primeiro, compreender que há duas formas desenvolvidas pela literatura para pensar essas estratégias que podem ser utilizadas pelos leitores. Uma dessas formas entende estratégia em um sentido hipônimo, isto é, de modo mais direcionado, que busca responder a demandas mais pontuais e que podem ser mobilizadas antes, durante e depois da leitura.

Já no sentido hiperônimo, as estratégias são mobilizadas de modo mais amplo durante a leitura, e o leitor as mobiliza com um objetivo específico. Nessa perspectiva, seguida por esta pesquisa, as estratégias são ações orientadas, tiram o leitor do modo automático e, portanto, requerem engajamento e reflexão consciente acerca do que está sendo lido. Um leitor proficiente usa de diferentes estratégias durante todo o processo de leitura. Nesse sentido, pode direcionar as estratégias utilizadas para questões mais pontuais da construção de significados ou, ainda, pode mobilizá-las para atingir o objetivo traçado. As estratégias podem ser definidas como “tentativas deliberadas e direcionadas para o objetivo de controlar e modificar os esforços do leitor para decodificar o texto, entender as palavras e construir significados<sup>14</sup>” (AFFLERBACH; PEARSON; PARIS, 2008, p.368, tradução nossa).

Os leitores estratégicos conseguem adaptar e flexibilizar a sua leitura, utilizando os elementos constituintes do texto para fazer as ligações necessárias. Assim, as estratégias demandarão atenção por parte do leitor, envolvendo planejamento e recursos cognitivos – tudo isso com a finalidade de construir significados (AFFLERBACH; CHO, 2011). As estratégias, portanto, são processos direcionados e aplicados de modo intencional, representam o conhecimento procedimental (VAN DIJK; KINTSCH, 1983). Um leitor estratégico trabalha no sentido de atingir seu objetivo com a leitura, lançando mão das ações necessárias para isso.

Nesse sentido, é possível diferenciar estratégias de habilidades, pois estas, como já apresentamos, são compreendidas como processos automáticos, que, de acordo com

---

<sup>14</sup> Reading strategies are deliberate, goal-directed attempts to control and modify the reader's efforts to decode text, understand words, and construct meanings of text.

Afflerbach, Pearson, Paris, (2008, p.368, tradução nossa), “resultam na decodificação e compreensão com velocidade, eficiência e fluência e, geralmente, ocorre[m] sem consciência dos componentes ou controles envolvidos<sup>15</sup>”. Embora sejam processos diferentes, habilidades e estratégias são os dois lados de uma mesma moeda (MANOLI; PAPADOPOULOU, 2012) – porque estratégias são as próprias habilidades trazidas para a metacognição durante o processo de leitura. Por esse motivo, um leitor proficiente é habilidoso e sabe, também, se valer das estratégias necessárias e requeridas diante da tarefa específica. Como as estratégias são ações deliberadas do leitor, é possível fazer inferências sobre os processos cognitivos que ocorrem durante a leitura a partir das estratégias utilizadas para a construção da compreensão de um texto (AFFLERBACH; CHO; 2011). Há que se considerar, contudo, que as estratégias também são catalogadas a partir de inferências feitas pelo pesquisador por meio do comportamento do leitor diante do texto, ou seja, por meio da forma como ele atua e se comporta diante do texto (KLEIMAN, 2016). Muito do que se sabe sobre as estratégias foi descoberto a partir de estudos realizados por meio de protocolos verbais (AFFLERBACH; JOHNSTON, 1984; ERICSSON; SIMON, 1993; PRESSLEY; AFFLERBACH, 1995; TOMITCH, 1995; SOUZA, 2004; FINGER-KRATOCHVIL, 2010; PROCAILO, 2017), que fornecem dados relevantes sobre o comportamento metacognitivo do leitor diante de um texto.

Nesse ponto, é relevante esclarecermos que, nos objetivos desta pesquisa, trazemos dois termos para as estratégias catalogadas e analisadas: estratégias e comportamento estratégico. Isso porque, no que tange aos dados obtidos por meio do mapeamento de tela durante a realização do teste de compreensão leitora, temos indícios diretos do comportamento estratégico do leitor, isto é, as estratégias realizadas por meio dos movimentos em tela com as idas e vindas ao texto. Já nos dados obtidos com os protocolos verbais, inferimos as estratégias a partir do que foi relatado oralmente, ou seja, a partir dos pensamentos verbalizados pelos participantes, pudemos depreender as ações que estavam sendo feitas para que eles construíssem alguma sorte de compreensão.

### **2.1.3 A situação da leitura**

Tendo conhecido o papel do leitor no processo de leitura, há, ainda, que se considerar que a leitura ocorre em uma situação. Leitor e texto se encontram em uma situação específica,

---

<sup>15</sup> Reading skills are automatic actions that result in decoding and comprehension with speed, efficiency, and fluency and usually occur without awareness of the components or control involved.

que determina os passos que o leitor dá na (inter)ação sobre o texto. Assim, “o sentido de um texto não existe *a priori*” (KOCH; ELIAS, 2015, p.57). Não porque quem lê pode compreender o que bem entender, mas porque existem aspectos que vão além do texto em si e do próprio leitor.

De modo objetivo, a situação envolve, de acordo com Souza (2011, p.9), os fatores relacionados ao “entorno de cada evento de contato com textos escritos visando à obtenção de informação e/ou aprendizagem e produção de sentidos”. Então, as condições em que o material é lido também influenciam nos processos de compreensão. Isso porque o leitor não está isolado de seu entorno e os processos cognitivos envolvem, também, aspectos atencionais e emocionais, os quais são afetados pela situação em que a leitura acontece.

Quando estamos falando em situação de leitura, consideramos as condições de espaço, tempo, de estado físico e emocional do leitor. Esses fatores afetam a concentração e influenciam no processamento dos diferentes níveis linguísticos e na formação de um modelo mental coerente. Assim, a situação envolve fatores externos que atuam sobre o leitor, que é quem deve reger a leitura. Ocorre que, não havendo condições de fazer essa regência, a leitura pode ficar prejudicada.

Como foi apresentado, leitor, texto e situação estão envolvidos no processo de leitura. Não existe leitura sem esses aspectos; não é possível retirar qualquer deles da equação. Sem algum deles, o quebra-cabeça não se completa, e a compreensão falha. Como já explicitado, aqui, fizemos a separação apenas para que fosse possível compreender suas especificidades, mas todos os três são fatores que atuam na leitura. A intersecção entre eles ficará ainda mais perceptível na seção seguinte, quando trazemos os modelos cognitivos de compreensão que, ao longo dos anos, buscaram estudar como a compreensão em leitura se dá cognitivamente. São esses modelos cognitivos que dão base ao modo como olhamos e analisamos esse quebra-cabeça com a finalidade de se construir a compreensão em leitura.

#### **2.1.4 Modelos cognitivos de compreensão em leitura**

Para entender melhor os níveis de processamento, os componentes envolvidos na compreensão e, por sua vez, a perspectiva de leitura que orienta esta tese, é importante refletir sobre os diferentes modelos que buscam explicar os processos cognitivos envolvidos no ato de ler, isto é, os modelos procuram elucidar o que acontece na mente do leitor durante os processos de construção da compreensão. Ressaltamos a “perspectiva de leitura que orienta a tese”, porque, embora o modelo de Kintsch guie a maior parte do trabalho e das análises, buscamos

trazer discussões pautadas na perspectiva que entende o processo de leitura como interativo, ou seja, a compreensão se dá a partir de processos que interagem entre si e tais processos são resultados do que o texto fornece e do que o leitor faz com esse estímulo fornecido em uma situação específica. Para chegar a essa perspectiva, discorreremos de modo geral acerca dos três modelos principais que têm orientado as discussões em torno da compreensão em leitura ao longo do tempo, a saber: modelos de processamento ascendente (*bottom-up*), modelos de processamento descendente (*top-down*) e os modelos interativos. Cada um deles traz contribuições importantes para as formulações em torno dos processos envolvidos na compreensão em leitura.

Nos modelos ascendentes, a compreensão é tida como unidirecional, em que o texto fornece todas as informações necessárias ao leitor, ou seja, para esses modelos, o texto está em evidência. Nessa perspectiva, todos os *inputs* fornecidos pelo texto são processados linearmente (GOUGH, 1976). Entende-se, assim, que o fluxo da informação se dá totalmente do texto para o leitor, e cada nível (grafêmico, fonêmico, lexical, sintático) é dependente do anterior para que seja processado, ou seja, o texto é processado em etapas. Nesse modelo, estão contidas explicações que consideram os níveis inferiores envolvidos na compreensão, chegando até o nível de integração sintática. No entanto, para a compreensão ocorrer, não basta a integração sintática, embora ela seja necessária.

A leitura, nesse viés, é, basicamente, um processo de decodificação linear, em que os passos executados pelo leitor são determinados pelas informações escritas no texto (LEFFA, 1999). Pode-se considerar, conforme destaca Souza (2004), que a perspectiva postulada por esses modelos ascendentes é bastante relevante para caracterizar a fase de aprendizagem da leitura, pois, quando um indivíduo está aprendendo a ler, é possível dizer que o processamento ocorre de modo mais serial e linear. Nessa etapa, o esforço primeiro do leitor está em decodificar a informação apresentada na palavra e então na frase e no texto. No entanto, tal processo demanda tempo e recursos cognitivos, limitando a atuação de níveis mais altos de análise e, por conseguinte, a construção de sentidos. O que se espera de um leitor proficiente é que esses níveis mais baixos, descritos por Gough (1976), ocorram de modo automático, para liberar recursos de memória, que serão usados nos processos de níveis superiores. O processamento da anáfora está incluído nessa integração dos níveis inferiores, deve, portanto, estar automatizado para que o leitor possa focar seus recursos em níveis mais altos de processamento.

Esse modelo dá conta, portanto, de explicar o processo da leitura até um determinado ponto. No entanto, para que a compreensão em leitura ocorra, é necessária a atuação do leitor

para além da decodificação linear do texto. Precisamos, então, caminhar mais um pouco para chegar ao ponto daquilo que entendemos como compreensão em leitura nesta tese. Para tanto, cabe tratar de um outro modo sob o qual a leitura foi estudada: a perspectiva dos modelos descendentes.

É com vistas a abordar e estudar o papel do leitor no processo de compreensão em leitura que surgem os modelos descendentes (*top-down*). Goodman (1976) traz para a cena da compreensão a atuação do leitor, postulando que a leitura é mais do que um processo preciso e linear. Ou seja, não basta ao leitor conhecer, identificar e integrar o código para que a leitura se concretize. Nessa proposta, a leitura é um jogo de adivinhações, e, para ser bem-sucedida, depende da habilidade que o leitor tem para selecionar as pistas relevantes deixadas pelo texto e utilizá-las de modo a construir sua própria compreensão. Nessa direção, as informações gráficas, semânticas e sintáticas servem apenas para confirmar o palpite do leitor. Em outras palavras, a leitura é feita a partir dos conhecimentos declarativos do leitor e envolve precisão nas suposições formuladas. Quando a suposição não é confirmada, correções precisam ser efetuadas. Nessa perspectiva, quanto mais veloz for a leitura e quanto mais habilidades estiverem desenvolvidas, menos as pistas gráficas são utilizadas. Vale ressaltar, ainda, que, para os modelos descendentes, ocorre uma adequação do texto aos conhecimentos do leitor. De modo resumido, pode-se dizer que o leitor seleciona algumas informações para, então, construir os seus significados a partir do texto.

Mais uma vez, a leitura é compreendida de modo unidirecional, assim como nos modelos ascendentes, com a diferença que o foco sai do texto e se centra no leitor. Certamente, há muito a ser aproveitado desses modelos descendentes, pois a compreensão em leitura não ocorre sem o leitor. No entanto, entendemos que não há como descrever a leitura em um processo unidirecional (nem de um lado e nem de outro). Isso porque para compreender o que se lê é necessário que haja texto, com todos os seus elementos e níveis, e é necessário que haja o leitor, com todo o seu conhecimento e com todos os aspectos cognitivos envolvidos no processamento da informação, além do contexto e da situação em que a leitura acontece.

De modo a conseguir conciliar as diferenças existentes nesses modelos que entendem a leitura de modo unidirecional, surgem os chamados modelos interativos, que argumentam pelo processamento em paralelo nos diferentes níveis envolvidos na leitura. É nessa perspectiva que situamos esta pesquisa. Assim, as informações apresentadas pelo texto e os conhecimentos trazidos pelo leitor, juntos, promovem a compreensão em leitura. De acordo com Rumelhart (2013, p. 732, tradução nossa), “[...] o processo de leitura é o produto da aplicação conjunta



simultânea de todas as fontes de conhecimento<sup>16</sup>.”. Isso significa dizer que ocorre uma combinação de processos de níveis superiores e inferiores, fazendo com que informações ascendentes e descendentes atuem cooperativamente, com a finalidade de interpretar o *input* recebido da maneira mais apropriada (RUMELHART; McCLELLAND, 1981). Nessa perspectiva, a percepção de palavras, por exemplo, pode ser influenciada pelo conhecimento ortográfico e de itens lexicais, pelos efeitos da sintaxe e pelos conhecimentos semânticos. Ainda, a percepção da sintaxe de dada palavra depende do contexto em que ela está inserida, o contexto influencia, também, na compreensão de todo um segmento lido (SAMUELS; KAMIL, 1984). Ou seja, os diferentes níveis interagem para que o processamento ocorra e, assim, um nível pode compensar possíveis dificuldades encontradas em outro nível. Neste momento do texto, podemos retomar o fato de nossa investigação trabalhar com dois níveis complexos de processamento, a saber: um nível micro, em que ocorrem a correferenciação e os processos de integração, que atuam na formação da base textual; e um nível macro, em que ocorre a formação de uma base textual. A partir de todos esses processos, desde o início da leitura, o leitor constrói a representação mental do texto lido (KINTSCH; VAN DIJK, 1978; KINTSCH, 1998; KINTSCH; RAWSON, 2013). Há diferentes níveis e processos que acontecem entre o nível em que a correferenciação se situa e o nível da formação da base textual. Entendemos que há muitos aspectos envolvidos nesse entremeio. Nesta pesquisa, abordaremos alguns deles para tentar colocar um tijolo na construção dessa ponte entre os níveis micros e os níveis macros envolvidos na compreensão.

Entendemos a leitura, portanto, em uma perspectiva cuja centralidade não está exclusivamente no texto ou no leitor (SOLÉ, 1998). O que ocorre é uma cooperação entre os recursos dispostos no texto e as expectativas que tais recursos geram no leitor a partir do seu conhecimento prévio, dos objetivos e da situação em que a leitura ocorre. Isto é, o texto é o estímulo a partir do qual inicia o processamento em um determinado nível (decodificação, reconhecimento das palavras...) e se propaga a níveis mais elevados. Essa propagação ocorre a partir de processos *top-down*. Nessa perspectiva, o processo da leitura inicia, de acordo com Rumelhart e McClelland (1981), com um rol de hipóteses geradas pelo conhecimento de uma série de elementos: letras, sentenças, aspectos contextuais e aspectos não-linguísticos. Conforme a leitura progride, as hipóteses consistentes com o *input* inicial vão sendo confirmadas, e as inconsistentes são enfraquecidas. As hipóteses que se confirmam propiciam previsões ainda mais específicas acerca da sequência de estímulos e o processamento é

---

<sup>16</sup> “[...] the reading process is the product of the simultaneous joint application of all the knowledge sources.”

facilitado. Nessa forma de compreender a leitura, entram em jogo diferentes níveis de abstração e o processamento não segue uma ordem linear; não há uma sequência preferencial, porque é simultaneamente conduzido pelo leitor a partir das informações disponíveis no texto. Segundo Souza (2012, p.48), “[é] o leitor quem orchestra a leitura. É o leitor quem produz sentido. É o leitor quem transforma. Mas é o texto que o dirige e orienta.”. Em resumo, a perspectiva que orienta este trabalho entende que, para a leitura acontecer, há um texto, que é o estímulo, há um leitor (que é o articulador) e há um contexto (que é a situação). O texto traz informações linguísticas, que serão processadas pelo leitor. A correferenciação, aspecto central desta tese, atua no nível do texto, é uma das peças que dirige e orienta o leitor.

Esses modelos até então apresentados tratam do processamento, buscando estabelecer generalizações acerca da atividade cognitiva que opera na sequência temporal da leitura, mas, para poder refletir sobre os processos de leitura de modo mais acurado, é importante conhecer os componentes envolvidos nessa habilidade. Dessa forma, é possível descrever os elementos que atuam nos processos da leitura e chegar, na sequência, ao processamento anafórico. Dentro da perspectiva de leitura que adotamos neste trabalho, trazemos a explicação dos componentes envolvidos na leitura a partir de duas perspectivas teóricas. Uma desenvolvida por Gagné, Yekovich, Yekovich (1993), que explica os diferentes componentes que atuam no processamento da leitura. Essa perspectiva é mobilizada para que possamos esclarecer os diferentes níveis envolvidos no processamento em leitura de modo resumido. No entanto, a perspectiva que exploramos nesta tese e que a embasa é aquela desenvolvida por Kintsch e colegas ao longo dos anos (2013; 1998; 1988; 1983; 1978), que se atém de modo mais demorado aos processos de nível superior e sobre a qual discorreremos na sequência.

Gagné, Yekovich e Yekovich (1993) apresentam um modelo que descreve de modo bastante esclarecedor os componentes envolvidos nos diferentes níveis de processamento. A partir do que os autores discorrem, é possível ter indícios de como um leitor hábil processa a leitura em sua mente. Nessa perspectiva, para que a compreensão em leitura se concretize, são necessários os conhecimentos declarativos e os conhecimentos procedimentais. De modo mais específico, os conhecimentos declarativos são aqueles possíveis de serem verbalizados, dizem respeito ao conhecer “sobre”, ou seja, conhecimentos sobre a língua (as letras, os fonemas, os morfemas etc) e sobre o mundo (conhecimentos prévios sobre o tópico lido, conhecimentos sobre o funcionamento da sociedade, conhecimentos experienciais etc). Já os conhecimentos procedimentais são caracterizados por um conjunto de habilidades necessário para uma leitura fluente; envolvem o “como” fazer. São habilidades, na maioria das vezes, internalizadas e que se tornam automáticas em alguma medida. Esses conhecimentos procedimentais ou habilidades

são formados pelos componentes processuais, divididos em quatro níveis de compreensão: decodificação, compreensão literal, compreensão inferencial e monitoramento da compreensão. Embora esses processos sejam denominados de estágios, na leitura proficiente, eles ocorrem de modo paralelo, convergindo com um modelo de leitura interativo, como apresentado anteriormente. Assim, ainda que a leitura fluente não ocorra de modo sequenciado, pois, como já abordado, os processos ocorrem em paralelo, com estágios que interagem hierarquicamente (cf. RUMELHART, 2013), os processos são aqui apresentados em uma sequência que parte daqueles de níveis inferiores para os de níveis superiores, conforme propõem Gagné, Yekovich e Yekovich (1993).

A decodificação é considerada um processo de nível inferior, o mais específico da leitura. Isso significa dizer que ela dá base para que os demais processos ocorram. Esse processo acontece de modo ascendente e está relacionado à leitura de palavras. Há dois subprocessos pelos quais a decodificação pode ocorrer: emparelhamento e recodificação. De acordo com Tomitch (2012), no emparelhamento, o leitor acessa de modo direto o significado da palavra em sua memória. Isso quer dizer que a palavra já está salva no léxico mental do leitor e, no momento em que é visualizada, seu significado é diretamente ativado por meio do acesso lexical. Já no subprocesso de recodificação ocorre a tradução da palavra a uma sequência de sons (que podem ser apenas mentais). Em outras palavras, o estímulo grafêmico recebido por meio das letras é transformado em seu valor fonêmico. Após essa transformação, de acordo com Gagné, Yekovich e Yekovich (1993), o significado da palavra é ativado. Normalmente, esse processo ocorre quando o leitor encontra palavras que não lhe são familiares.

A recodificação é bastante comum, também, na fase da alfabetização, quando se está aprendendo a decodificar e ainda não se acessa o significado das palavras de modo direto. A tendência, contudo, é que, com o desenvolvimento da fluência leitora e da aquisição de vocabulário, esse processo seja gradualmente substituído pelo emparelhamento. Cabe, ainda, ressaltar que, embora a decodificação seja considerada um processo de nível inferior, é essencial para que a leitura ocorra. No entanto, decodificar não é sinônimo de ler. Ademais, é necessário que esse processo esteja automatizado para uma leitura fluente.

Outro processo que ocorre, também em níveis inferiores, é a compreensão literal, cuja função é obter o significado literal do que está escrito. Esse componente está dividido em dois subprocessos: o acesso lexical e o *parsing*. O primeiro acessa o significado das palavras que são identificadas. Os autores explicam que “o processo de decodificação ativa a percepção da palavra na memória declarativa e as operações de acesso lexical selecionam a interpretação

adequada da palavra a partir do conhecimento ativado<sup>17</sup>.” (GAGNÉ; YEKOVICH; YEKOVICH, 1993, p. 272). Com essa explicação, percebe-se que, embora seja feita a separação entre os conhecimentos declarativos e procedimentais, eles interagem a todo momento. Esses significados individuais das palavras precisam, por sua vez, ser combinados para que ocorra a formação de uma unidade maior de significação, essa é a função dos processos de *parsing*. Para tanto, são usadas as regras sintáticas e gramaticais da língua, formando unidades proposicionais.

A compreensão em leitura requer, contudo, mais que compreender palavras e frases, ou seja, não basta identificar as palavras, seus significados e as unidades menores de significação, é preciso construir um todo coerente e integrado por meio de uma rede complexa (KINTSCH, 1998; KINTSCH; RAWSON, 2013). Por conta disso, entra em cena o componente da compreensão inferencial, que envolve processos de níveis superiores e que se mostra, também, relevante para este trabalho.

A compreensão inferencial é responsável pelo entendimento mais aprofundado de unidades de textos maiores e mais complexas. É o caso, por exemplo, dos textos acadêmico-científicos. Ela é constituída por três subprocessos, a saber: integração, sumarização e elaboração, segundo Gagné, Yekovich e Yekovich (1993). A integração é o processo que estabelece as relações entre as diferentes partes do texto (sentenças, parágrafos), a partir dessas relações, constrói-se uma nova unidade de significado, resultando em uma “[...] representação declarativa mais coerente<sup>18</sup>” (GAGNÉ; YEKOVICH; YEKOVICH, 1993, p. 275, tradução nossa). É a partir dos processos de integração que o leitor estabelece as relações entre as sentenças e constrói a coerência do texto. Embora Gagné, Yekovich e Yekovich (1993) relacionem os processos de integração à compreensão inferencial, a integração ocorre, também, a partir das conexões locais, como no caso da anáfora correferencial. Acionaremos isso ao tratar do modelo proposto por Kintsch (2013; 1998).

Outro subprocesso que atua nesse nível é o da sumarização, que se trata de um esboço mental a partir do qual o leitor organiza, de modo hierárquico, as ideias contidas no texto em uma macroestrutura coerente (KINTSCH, 1998; KINTSCH; RAWSON, 2013). É por meio desse processo que são feitas as inferências globais que conduzem o leitor à compreensão da ideia principal do texto (TOMITICH, 2012). A elaboração, por sua vez, ocorre quando, a partir da mobilização dos conhecimentos prévios relevantes – que acontece durante todo o processo

---

<sup>17</sup> the decoding process activates the word percept in declarative memory, and the lexical access operations select the right interpretation of the word from all of the knowledge that has been activated.

<sup>18</sup> [...] a more coherent declarative representation.

da leitura -, o leitor atua sobre o texto, elaborando acerca dele e transcendendo as informações que estão explícitas na base textual (OAKHILL; CAIN; ELBRO, 2017). É no nível da compreensão inferencial que a informação é armazenada e que o conhecimento se torna significativo ao leitor.

Por último, dentro da organização desenvolvida por Gagné, Yekovich e Yekovich (1993), estão os processos envolvidos no monitoramento da compreensão, que são: estabelecimento de objetivos, seleção de estratégias, checagem do objetivo e remediação. Tais processos envolvem a metacognição, pois são usados de modo desautomatizado por leitores fluentes, com o intuito de controlar seus próprios processos cognitivos e garantir a compreensão em leitura. Como abordamos na seção anterior, o leitor estabelece um objetivo para leitura e, então, são selecionadas as melhores estratégias para atingi-lo. Para garantir que o objetivo seja atingido, é necessário fazer constantes checagens, que podem interromper o fluxo da leitura e, então, a remediação pode ser aplicada, caso se perceba que o objetivo estipulado não está sendo alcançado. Recorre-se, também, à remediação nos momentos em que se percebe que algo no texto ou na situação interrompe o fluxo da compreensão; lança-se mão, portanto, de recursos de modo que se possa resolver o impasse. Esses impasses podem ser gerados nos diferentes níveis envolvidos na compreensão, inclusive durante o estabelecimento da correferência anafórica, a partir de uma ambiguidade ou mesmo do uso de algum elemento anafórico que aumente o tempo de processamento, dificultando ou interrompendo o fluxo da leitura. O esperado é que, caso isso ocorra, o leitor se valha de estratégias para resolver o problema.

Pois bem, todos os níveis de processamento constantes no modelo componencial proposto por Gagné, Yekovich e Yekovich (1993) foram apresentados. Com o exposto até o momento, é possível entender a complexidade envolvida na compreensão em leitura. São diversos e distintos fatores ativados para que seja possível a construção de uma representação mental coerente do texto. A esse modelo, outros, como o alvitado por Kintsch (1998), podem ser relacionados com o propósito de auxiliar na compreensão de determinados processos. O modelo proposto por Kintsch (1998) é bastante relevante, porque foca estritamente em processos de nível superior, mais especificamente, naqueles envolvidos na integração, fornecendo, então, subsídios para entender como as pessoas compreendem textos. Isso não significa que o autor não reconheça os processos de nível inferior. O que ocorre é que tais processos são automáticos na leitura proficiente. Por conta disso, não os coloca como foco de atenção na descrição do modelo. Esse modelo nos auxilia, sobretudo, na definição daquilo que entendemos por compreensão e, também, a situar de modo mais consistente o nível envolvido no processamento da correferência anafórica.

Esse modelo é de base semântica e foca na representação do significado (KINTSCH, 1998), a interação ocorre entre três níveis de processamento: integração das unidades de micro-significados, redução do texto à sua essência e geração de uma representação mental por parte do leitor (VAN DIJK; KINTSCH, 1983; KINTSCH, 1998). Esse modelo orienta nossas discussões e análises, porque, nele, a compreensão é entendida dentro de sua complexidade, isto é, o processamento lexical é retroalimentado pelo processamento sintático, semântico e estrutural. Assim, há:

[...] um contínuo feedback entre as unidades mais e menos complexas: entender a função de uma palavra em uma oração dependerá da estrutura funcional da oração como um todo, tanto no nível sintático quanto no semântico. Isso significa que ao invés de um modelo estrutural convencional de processamento, operamos com um modelo estratégico.<sup>19</sup> (VAN DIJK; KINTSCH, 1983, p.22, tradução nossa).

Nessa perspectiva, para que a compreensão seja construída são necessários múltiplos processos que podem operar serial ou paralelamente de modo a fazer conexões entre as diferentes proposições formuladas a partir do texto. São diferentes níveis linguísticos e cognitivos que atuam em conjunto para se construir a compreensão. Nesse sentido, ocorrem processos de construção e integração a partir dos três tipos de operações citados.

De modo mais específico, o primeiro está relacionado aos elementos de significação organizados em um todo coerente e construídos a partir de uma análise semântica, que parte da construção e integração das palavras do texto e de suas relações sintáticas, formando unidades de ideias chamadas de proposições, que estão relacionadas de modo complexo em uma rede que constitui a microestrutura. Esta é a “[...] estrutura local do texto, as informações sentença por sentença, complementadas e integradas por informações da memória de longo prazo<sup>20</sup>” (KINTSCH, 1998, p.50, tradução nossa). Um modo de concretizar essas relações na microestrutura textual é a partir da anáfora correferencial, pois ela propicia essa relação entre as diferentes proposições em um texto. Ou seja, a anáfora atua em um nível de processamento micro e, de modo interativo, se relaciona com os demais processos que atuarão na construção da compreensão.

O segundo tipo de operação diz respeito a porções de significação do texto, ou seja, para além das relações entre palavras e frases, há uma estrutura global de significação, chamada de

<sup>19</sup> “[...] there is continual feedback between less complex and more complex units: Understanding the function of a word in a clause will depend on the functional structure of the clause as a whole, both at the syntactic and at the semantic level. This means that instead of a conventional structural model of processing, we operate with a *strategic* model.”

<sup>20</sup> “[...] is the local structure of the text, the sentence-by-sentence information, as supplemented by and integrated with long-term memory information.”

macroestrutura. Na macroestrutura, são formados os tópicos sob os quais as unidades que compõem a microestrutura estão organizadas. Ao se conectar os processos que possibilitam a formação da microestrutura de um texto aos de formação da macroestrutura, tem-se o que Kintsch e Rawson (2013) denominam de base textual.

A base textual traz o significado expresso pelo texto, ou seja, é o sumo, aquilo a que se chega com a interrelação das proposições presentes em um texto; podemos considerá-la a síntese do texto. Espera-se, contudo, que um leitor proficiente possa ir além da significação do texto em si, usando o conteúdo das proposições para fazer relações e inferências a partir de seu conhecimento prévio relevante e a partir dos objetivos estipulados para a leitura, ou seja, são essenciais, para a compreensão, os processos de elaboração (GAGNÉ; YEKOVICH; YEKOVICH, 1993). Ao fazer isso, constrói-se um modelo situacional acerca do que foi lido, isto é, “um modelo mental da situação descrita pelo texto” (KINTSCH; RAWSON, 2013, p. 229).

Entra, neste ponto, o terceiro tipo de operação, que é a geração mental de um novo texto a partir da memória construída por meio dos processos de compreensão. A construção desse modelo mental se dá a partir das integrações, inferências, previsões, elaborações e conclusões que ocorrem no decorrer da leitura (STERNBERG; STERNBERG, 2017). Assim, esse modelo é constantemente reconstruído conforme, durante a leitura, as hipóteses e as previsões se confirmam ou não. Esse modelo se constrói a partir do processamento mais profundo daquilo que é fornecido pelo texto, portanto, é necessário que se formule uma base textual adequada e, para que isso ocorra, como destacado, a resolução das anáforas correferenciais é uma das operações que precisam ser concretizadas no nível microestrutural, atuando em processos de integração. Como veremos no capítulo destinado ao método desta investigação, procuramos captar a anáfora correferencial em seu nível micro, investigando o tempo de processamento, mas, também, estamos interessados em saber se os participantes conseguiram formular a base textual e se o texto foi compreendido em seu sentido mais profundo e, se não foi, se as anáforas correferenciais, sobretudo o sintagma “o mesmo”, tiveram alguma influência nisso.

Nesta pesquisa, portanto, entendemos a compreensão em leitura como uma construção complexa, que envolve diferentes níveis de processamento que interagem entre si e que, de forma estratégica, se retroalimentam. Nesse sentido, os diferentes níveis atuam conjuntamente para que o objetivo estipulado para a leitura seja atingido pelo leitor. Assim, a coerência local é requerida e, para isso, o leitor precisa buscar, “da maneira mais eficaz possível, as ligações potenciais entre os fatos denotados pela proposição. [...] Uma estratégia possível, portanto, é

procurar aqueles argumentos em uma proposição que se correlaciona com um dos argumentos da proposição anterior”<sup>21</sup> (VAN DIJK; KINTSCH, 1983, p.15, tradução nossa).

Sabemos, conforme está explícito na teoria abordada nesta seção, que, durante a leitura de um texto mais longo, há uma gama de níveis de processamento que atuam conjuntamente para que o texto seja processado e a compreensão construída. Não controlamos, nesta pesquisa, todos os processos ocorridos em todos os níveis, porque a discussão tem por foco as possíveis relações entre o estabelecimento da anáfora correferencial (sobretudo aquela realizada com o sintagma “o mesmo”) e a compreensão ou não do texto. Estabelecer essas relações será possível por meio da análise do tempo de processamento, da aceitabilidade dos anafóricos, das estratégias empregadas, do comportamento diante do texto, e do nível de proficiência demonstrado na compreensão do texto lido.

Há, ainda, muitos outros aspectos cognitivos envolvidos na compreensão em leitura, aspectos relativos à memória, à atenção, à motivação, além de, como mencionado, aspectos relativos ao contexto, ou seja, às circunstâncias em que a leitura acontece e todos os fatores externos (ao texto e ao leitor) que podem incidir sobre os processos atencionais e perceptivos. Esses fatores não entraram no cerne desta investigação, mas sabemos que eles têm atuação em todo e qualquer processo de leitura. No que diz respeito ao estudo do processamento anafórico, têm especial envolvimento os componentes da memória, sobretudo a memória de trabalho, sobre a qual discorreremos na seção 2.1.5.

Com todo o apresentado, espera-se terem sido respondidas as perguntas que abrem esta seção: O que é leitura para esta tese? O que está envolvido na compreensão em leitura? Por que saber ler palavras não resulta no entendimento daquilo que é lido? Como se chega à compreensão? Discorreremos sobre o que é leitura e, a partir disso, foi possível entender que ela envolve, necessariamente, a compreensão (dentro da perspectiva deste trabalho). Há diferentes modos de entender a construção dessa compreensão; nós partimos de uma perspectiva interativa, que prevê que existem diferentes níveis que atuam conjuntamente e em paralelo de modo complexo e simultâneo em leitores proficientes. Ao considerar que a leitura se constitui por diferentes níveis de processamento, é possível entender por que ler palavras, embora seja essencial, não garante a compreensão do que foi lido. Para que essa compreensão aconteça, é preciso que ocorram os processos de níveis superiores, garantindo uma compreensão em nível inferencial, com a construção de um modelo mental do texto (GAGNÉ; YEKOVICH;

---

<sup>21</sup> the strategic establishment of local coherence requires that the language user search as effectively as possible for potential links among facts denoted by the proposition. [...] One possible strategy, therefore, is to look for those arguments in a proposition which corefer with one of the arguments of the previous proposition.



YEKOVICH, 1993; KINTSCH, 1998; KINTSCH; RAWSON, 2013; OAKHILL; CAIN; ELBRO, 2017).

Na seção 2.3, o nível do processamento da anáfora correferencial e o seu papel na compreensão em leitura será mais bem especificado e, então, entram em cena mais aspectos dos modelos de Gagné, Yekovich e Yekovich (1993) e de Kintsch e seus colaboradores (KINTSCH, 1998; KINTSCH; RAWSON, 2013), que, resguardadas suas especificidades, podem ser compreendidos como complementares. Para entender o papel do processamento da anáfora correferencial em atividades de leitura, é necessário, primeiro, discutir as noções de referência e referenciação, pois, antes de um elemento ser retomado no texto, ele é citado de modo a fazer referência a algo no mundo. Por isso, na seção 2.2, serão apresentadas as noções de referência e as formas a partir das quais ela pode ser construída e reconstruída no discurso. Antes, contudo, trazemos aspectos relacionados à memória, sistema essencial para que a compreensão em leitura ocorra.

### **2.1.5 Alguns aspectos sobre memória**

Como citado anteriormente, há diversos aspectos cognitivos envolvidos na compreensão em leitura. Um desses aspectos está relacionado à memória, que tem importante papel no processamento, na retenção e na elaboração daquilo que foi lido. De modo mais específico, faz-se ainda mais relevante conhecer os sistemas de memórias que atuam no processamento da anáfora, por isso, nesta seção, discorreremos sobre os sistemas de memória de longa duração e de memória de trabalho.

Em primeiro lugar, cabe salientar que a memória não é construída a partir de uma unidade, mas a partir de um sistema múltiplo e complexo (BADDELEY, 2011). Isso significa que não há uma memória, mas um conjunto de sistemas que atua concomitantemente para o processamento, o armazenamento e a elaboração das informações obtidas. Tal sistema tem atuação central na compreensão em leitura. De acordo com Sternberg e Sternberg (2017), foram identificadas três operações próprias da memória: codificação, armazenamento e recuperação. Na primeira, os dados sensoriais são transformados em representações mentais. No segundo, as informações que foram codificadas são mantidas na memória. Já na recuperação, as informações que foram armazenadas são acessadas.

Para resumir o que tem emergido dos estudos sobre a memória, podemos dizer que há um sistema de memória de longa duração, no qual estão armazenadas informações sobre os conhecimentos declarativos (memória declarativa), que são conhecimentos necessários para

que a leitura aconteça de modo proficiente. Tais conhecimentos dizem respeito a uma memória explícita, passível de lembrança deliberada (STERNBERG; STERNBERG, 2017). Essas lembranças podem ser de ordem semântica ou episódica. Na memória episódica, estão armazenadas lembrança de eventos específicos, que, conforme as experiências ocorrem, podem se acumular e formar o nosso conhecimento do mundo, armazenado em nossa memória semântica.

Então, a partir de eventos específicos e repetidos, as experiências são armazenadas na memória semântica e podem, assim, ser mobilizadas em outros momentos. Isso significa que a memória semântica está relacionada com o conhecimento que temos sobre o mundo. A diferença entre essas memórias reside no fato de a memória semântica tratar de um conhecimento organizado sobre o qual temos consciência (EYSENCK, 2011), enquanto a memória episódica diz respeito a lembranças de situações e eventos que ocorreram. Durante a leitura, tanto memória episódica quanto semântica são mobilizadas para que o leitor possa atribuir coerência ao lido. A depender do conhecimento adquirido acerca da temática, essas memórias serão menos ou mais acionadas.

Atuam, ainda, na memória de longa duração, os conhecimentos implícitos (STERNBERG; STERNBERG, 2017), ou seja, conhecimentos que estão armazenados e que mobilizamos de modo automático quando necessário. Na leitura, essa memória implícita auxilia em processos de nível inferior como a transformação de grafemas em fonemas, o acesso ao léxico mental, a construção sintática e a integração interfrasal.

A memória de longo prazo está relacionada à memória de trabalho, que tem atuação central nos processos de retomadas anafóricas, pois ela, sendo multicomponencial (BADDELEY, 2011), atua no armazenamento de curto período e nas “estratégias de seleção e funcionamento”, “servindo como um espaço operacional global” (BADDELEY, 2011, p.54). Assim, durante a leitura, mantemos e manipulamos em nossa memória de trabalho as informações que nos são fornecidas pelo texto. Este, por sua vez, é construído a partir de diferentes níveis linguísticos que estão relacionados. Nessas relações, atua a anáfora correferencial. Assim, para que um texto seja compreendido, as informações precisam ser retidas na memória de longa duração. Para isso, no entanto, é preciso que a memória de trabalho aja, performando diferentes operações mentais, de modo que os distintos níveis textuais possam ser integrados e relacionados adequadamente. Conforme Costa-Ferreira e Ferrari-Neto (2021), a compreensão do processamento da anáfora está intimamente relacionada à memória de trabalho, exatamente porque é essa memória a responsável por reter, manipular e relacionar as diferentes informações retomadas pelas anáforas.

A memória de trabalho está, então, envolvida na retenção temporária e na manipulação das informações. De acordo com Sternberg e Sternberg (2017, p.166), ela “retém somente a porção da memória de longo prazo mais recentemente ativada, ou consciente, e transfere os elementos ativados para dentro ou para fora do receptáculo da memória temporária”. Para entender a memória de trabalho, é necessário conhecer os componentes que a estruturam, porque, como dissemos, ela é multicomponencial. De acordo com Baddeley (2011), após muitos estudos, chegou-se a um modelo de memória de trabalho que comporta a alça fonológica, o esboço visuoespacial, o executivo central, e o *episodic buffer*.

A alça fonológica é, de acordo com Baddeley (2011, p.57), “um modelo de memória verbal de curta duração”. Então, na alça fonológica, as informações verbais ficam retidas por um tempo determinado, enquanto são manipuladas. Outro componente da memória de trabalho é o esboço visuoespacial, que é o responsável pelo armazenamento temporário de informações visuais (o quê) e espaciais (o onde). Assim, o sistema da memória de trabalho tem um componente responsável pelo armazenamento temporário de informações verbais e outro responsável pelo armazenamento temporário das informações visuais e espaciais. O executivo central é quem gerencia a memória de trabalho. Esse componente se relaciona com a atenção, é “um controlador atencional, em lugar de um sistema de memória” (BADDELEY, 2011, p.66). Baddeley (2011) propõe, ainda, um quarto componente ao sistema da memória de trabalho: o *episodic buffer*. Esse componente

é supostamente o sistema de armazenamento que consegue reter em torno de quatro segmentos de informação em um código multidimensional. Devido a essa capacidade de reter uma gama de dimensões, ele é capaz de agir como uma conexão entre vários subsistemas de memória de trabalho, e também de ligar esses subsistemas a *inputs* da memória de longa duração e da percepção. (BADDELEY, 2011, p.70).

Então, o que o *episodic buffer* faz é integrar de modo ativo informações sobre episódios e eventos combinando diferentes fontes de informações, como visuais e auditivas, ou seja, ele combina e integra as informações da alça fonológica, do esboço visuoespacial e da memória de longa duração. Esse sistema de memória é fundamental para a integração entre as partes de um texto e para que o leitor possa formar uma representação mental adequada acerca do que lê. No que tange à anáfora, como citamos anteriormente, a memória de trabalho é responsável por reter, na alça fonológica, o antecedente adequado e fazer a integração correta entre esse antecedente e o elemento anafórico. Ocorre que a anáfora precisa ser construída de uma forma a não prejudicar ou onerar esses processos de manipulação, porque, do contrário, o leitor pode

encontrar empecilhos nesse nível integrativo, que respigarão na formação de uma base textual coerente e em uma representação mental adequada.

Até este ponto do texto, foi possível vislumbrar os aspectos que envolvem leitura e compreensão em uma perspectiva cognitiva. Na seção seguinte, discorreremos sobre a construção textual da referência e da anáfora, para que seja possível compreender o papel desse elemento no texto em si. Afinal, não há leitura sem leitor, mas, também, não há leitura sem texto.

## 2.2 A NOÇÃO DE REFERÊNCIA E OS PROCESSOS DE REFERENCIAÇÃO NA LINGUÍSTICA

Nesta pesquisa, estamos estudando a anáfora em uma perspectiva do processamento em estímulos de leitura. Diante disso, precisamos compreender aspectos relativos à construção da referência em um texto escrito. Para tanto, serão mobilizadas teorias que, na linguística, dão base ao estudo da referência e da Linguística Textual. Precisamos, afinal, partir do texto, pois é lá em que referente e anáfora correferencial estão e é a partir desse estímulo linguístico que os processos cognitivos ocorrerão.

Para iniciar, discorreremos acerca da noção de referência e referenciação. Para tanto, mobilizamos a abordagem de diferentes autores que tratam dessa temática na área da linguística (HALLIDAY; HASAN, 1976; KOCH; ELIAS, 2018; MONDADA; DUBOIS, 2019). Trazemos a noção de referência, porque é a partir da sua construção em um texto que se desenham as anáforas correferenciais, foco deste trabalho. Assim, para que um item seja reativado cognitivamente (com as anáforas), ele precisa, primeiro, ser ativado, construído a partir do texto. E isso se dá por meio da referenciação.

Desde muito tempo, um dos questionamentos que envolvem diferentes áreas do conhecimento é o de como a língua refere o mundo. A filosofia clássica já refletia sobre isso e pressupunha uma relação direta entre as palavras e o mundo. Essa perspectiva compreende a língua como um espelho que reflete diretamente as coisas. No entanto, conforme apresentam Mondada e Dubois (2019), é possível considerar outra concepção, que entende os referentes como construtos culturais, construídos contextualmente. Nessa perspectiva, de acordo com as autoras, duas linhas argumentativas revisitam a noção de referenciação, a saber: a perspectiva de categorização, em que os processos cognitivos dão estabilidade ao mundo, e a perspectiva da linguística interacionista e discursiva, que considera os processos de referenciação a partir da construção de objetos do discurso.

Nessa direção, Koch e Elias (2018) apresentam a noção de referenciação como uma atividade discursiva, responsável pela “construção e reconstrução de objetos de discurso” (KOCH; ELIAS, 2018, p. 133-134). Nesse viés, “os referentes de que falamos não espelham diretamente o mundo real, não são simples rótulos para designar as coisas do mundo. Eles são construídos e reconstruídos no interior do próprio discurso [...]” (KOCH; ELIAS, 2015, p. 123). Assim, para essa perspectiva, a referenciação diz respeito aos construtos formulados no interior do discurso, a partir de operações realizadas pelo indivíduo com base em seu conhecimento prévio.

Voltando-se a uma perspectiva mais textual, é possível valer-se da noção de referência abordada por Halliday e Hasan (1976). Para os autores, a referência é um mecanismo responsável pela coesão de um texto. Nessa perspectiva, uma referência pode ser exofórica ou endofórica. É exofórica quando depende do contexto situacional para ser processada, ou seja, depende de recursos para além daqueles disponíveis no texto. Já a endofórica refere-se a algo identificado dentro do texto. Na abordagem dada por esses autores, ainda, o estudo da referência é recortado, pautando-se essencialmente nos procedimentos de anaforização, ou seja, para eles, a referência é um tipo de relação anafórica. Embora os diferentes autores citados olhem para a referência sob distintas perspectivas – textual, discursiva, filosófica –, é possível aproximá-los em alguma medida e compreender que, enquanto a referência é construída discursivamente, ela é, também, responsável pelos elos coesivos requeridos para um texto coerente.

Há que se considerar, ainda, que existem diferentes processos a partir dos quais uma referência pode ser desenvolvida, retomada e construída, a saber: dêixis, anáfora e catáfora. Embora as três andem juntas e apresentem características semelhantes, desempenham funções distintas na linguagem. Em uma compreensão que parta da perspectiva pensada por Halliday e Hasan (1976), a dêixis ocorre como referência exofórica; já a anáfora e a catáfora são referências endofóricas. De modo ampliado, a dêixis é um tipo de referenciação que depende do contexto situacional para ser compreendida, isto é, o seu referente não está explícito no texto, tudo depende do lugar, do momento e da pessoa que enuncia (APOTHÉLOZ, 2019). Assim, de acordo com Ilari e Galdi (1990, p. 67), “[...] as frases que comportam elementos dêiticos só podem ser interpretadas em estreita conexão com situações determinadas, e a informação que transmitem varia com o variar dessas situações”. Nesse sentido, o referente é construído a partir de toda uma situação que se desenvolve. A palavra “dêixis” vem do grego e significa “apontar” e tem por finalidade situar o referente em um tempo e espaço discursivo. As expressões dêiticas são formadas, frequentemente, por advérbios de tempo e de lugar, pelo tempo verbal, pelos demonstrativos e pelos pronomes pessoais de primeira e segunda pessoas. De acordo com

Apothéloz (2019, p. 68), “[..] há dêixis quando a expressão tem por objetivo deslocar o campo de atenção para um referente, e anáfora quando a expressão só faz manter o referente no campo de atenção.”. Cabe conhecer o que a literatura aborda acerca da anáfora, para se poder entender as relações e os distanciamentos entre ela e a dêixis.

O termo “anáfora” vem do grego e significa levar para trás. As construções anafóricas em um texto podem se dar de diferentes formas e sob diferentes perspectivas. Koch e Elias (2018) entendem a anáfora como um elemento responsável pela construção dos sentidos do texto, atuando na dinâmica discursiva e sendo fundamental para o processo de referenciação. Nessa perspectiva, os fatores que atuam para o processamento da anáfora têm por base o texto/discurso, mas podem implicar o conhecimento de mundo do leitor. De modo amplo, pode-se dizer que “[q]ualquer dispositivo linguístico que possa ser usado para referir a um conceito já mencionado se chama *anáfora* [...]” (KINTSCH; RAWSON, 2013, p. 233). Nesse sentido, é importante refletir sobre esses diferentes tipos de anáforas para situar a pesquisa ora desenvolvida.

A concepção clássica está relacionada com uma compreensão de anáfora direta (KOCH; ELIAS, 2018; MARCUSCHI, 2001), que indica a repetição de um sintagma já citado no texto, atuando na progressão e na coesão textual (HALLIDAY; HASAN, 1976) e buscando manter o referente, já introduzido, em foco por meio da sua reativação. Esse tipo de anáfora se dá, portanto, com a retomada de um sintagma específico explícito no texto. Assim, a partir do próprio texto, é possível identificar o antecedente da anáfora realizada. Trata-se, dessa forma, de uma anáfora correferencial, já que, como dito, antecedente e anáfora têm um mesmo referente construído contextualmente. Nesse caso, um elemento citado anteriormente é retomado, sem que haja introdução de outro referente e sem que se pretenda resumir toda uma ideia, como discutiremos mais adiante. Na perspectiva de Halliday e Hasan (1976), a anáfora é um dos elementos responsáveis por dar a um texto a sua textualidade, fornecendo coesão ao relacionar expressões que estão em diferentes partes do texto e conectando sentenças distintas. Para esses autores, a relação anafórica se estabelece quando o elemento pressuposto está explícito em período anterior e é, então, retomado na sequência.

O termo “correferencial”, por sua vez, é usado para informar que antecedente e anáfora possuem um mesmo elemento referencial. Há, nesse sentido, uma relação de identidade entre a referência do antecedente e do anafórico:

1. A cartilha trata do novo coronavírus. Ela apresenta dicas sobre os cuidados a serem tomados.

Em construções como a apresentada, temos as informações gramaticais de gênero e número fornecidas pelo pronome “ela” na segunda sentença, que deixam pistas de que tal pronome está se referindo à cartilha. O substantivo “cartilha”, por sua vez, possui um referente específico, construído textualmente, ou seja, tanto “cartilha” quanto “ela” tratam da mesma entidade, os dois possuem o mesmo referente, por isso são correferenciais. Nesta pesquisa, portanto, anáforas correferenciais se referem a essas retomadas que estão ancoradas a referentes introduzidos de modo explícito no texto.

Há, ainda, uma compreensão de anáfora que, além de considerar a complexidade existente nas relações construídas em um texto, busca dar conta de outras formas de progressão referencial. Para explicar esse tipo de relação anafórica, cabe retomar a noção de referente citada anteriormente. Apesar das diferentes formas existentes de se compreender a referenciação, considera-se, aqui, a noção de que o referente é construído a partir do contexto. Nesse sentido, quando se trata de um texto, há que se considerar que a introdução a um determinado referente pode se dar de forma não-ancorada ou de forma ancorada (KOCH; ELIAS, 2018; MARCUSCHI, 2001). Desta, fazem parte as anáforas indiretas e associativas.

Nessas construções anafóricas realizadas de modo indireto e associativo, considera-se que a progressão textual ocorre com a reativação de uma expressão citada anteriormente, ao mesmo tempo em que um novo referente é introduzido. Isso significa que o referente é novo, mas pode ser ativado de modo mais rápido, tendo em vista que o texto/discurso anterior oferece bases que apoiam esse novo referente. Encontramos isso em contextos como:

2. A fazenda tinha sido abandonada. O pasto estava dominado pelo mato.
3. O setor jurídico deve resolver o problema do contrato. Eles sempre sabem o que fazer.

“O pasto”, em 2, remete, anaforicamente, à âncora “a fazenda”. Ao mesmo tempo, é um novo referente, pois a relação com “a fazenda” se dá de modo associativo. Ou seja, quem tem um modelo mental para o frame “a fazenda”, tem “o pasto” como parte desse frame, embora não diga respeito ao mesmo referente. Trata-se, portanto, de uma anáfora indireta, construída a partir de um conhecimento de mundo ativado por meio de uma âncora de base semântica. Não existe, desse modo, uma relação correferencial entre âncora e anáfora indireta (MARCUSCHI, 2001). Já o contexto 3 traz uma anáfora associativa, ao ler o pronome “eles”, não é possível encontrar um antecedente explícito, mas, ao fazer a ancoragem com o referente “setor jurídico”, é possível compreender que o pronome remete à equipe que constitui tal setor, mais uma vez

não se tem uma relação de correferência, mas há uma construção anafórica que atua cognitivamente de modo a retomar uma informação já fornecida e de modo a fazer com que o leitor possa criar uma representação acerca do que é apresentado.

Outro tipo de anáfora pode se dar por meio do que Koch e Elias (2015) chamam de nominalização. Nesse caso, um novo referente encapsula uma informação apresentada anteriormente, como em:

4. Amor, tristeza, ansiedade. Todos esses sentimentos fazem parte da vida.

Nesse caso, a expressão “todos esses sentimentos” reativa o que foi citado anteriormente, mas introduz um novo referente, “sentimentos”. Conforme Conte (2019, p. 177), “[...] um novo referente discursivo é criado sob a base de uma informação velha [...]”. No que diz respeito à leitura, tais construções anafóricas são resolvidas a partir de processos de integração locais, que são realizados, na maioria das vezes, de modo automático.

Esse processo de nominalização pode se dar, conforme Apothéloz e Chanet (2019), por meio de um pronome demonstrativo, que condensa a ideia expressa em uma proposição. Halliday e Hasan (1976) chamam esse processo de referência estendida. Nesse caso, é guardada uma relação com a noção clássica de anáfora, salvo pelo fato de não retomar um antecedente específico, mas toda uma ideia. Além disso, é um processo que se diferencia da correferência já que não há um referente específico e individualizado:

5. As empresas tendem a demitir seus empregados diante de uma crise. Isso revela quais são as prioridades dos empresários.

Em 5 temos, então, uma anáfora encapsuladora em que o pronome “isso” retoma a ideia citada na frase anterior: “as empresas tendem a demitir seus empregados diante de uma crise” de forma condensada, mantendo-a ativa na memória do leitor, sem introduzir um novo referente no discurso.

É possível perceber que não há limites muito claros entre essas anáforas encapsuladoras, indiretas, associativas e as nominalizações. As discussões em torno das sutilezas da temática são amplas. Ressaltamos, apenas, o fato de as anáforas indiretas, associativas e as nominalizações se distanciarem do que se chama de correferência (ILARI, 2001) e se aproximarem, em alguma medida, do fenômeno entendido como dêixis (MARCUSCHI, 2001), pois, tanto para essas realizações anafóricas quanto para a dêixis, há um conhecimento do



contexto que atua para que as relações possam ser estabelecidas, ou seja, em todos esses casos é necessário ir para além do texto. No entanto, as anáforas associativas, indiretas e as nominalizações não se encaixam nem na concepção clássica de anáfora e nem na compreensão de dêixis, tendo em vista que elas não retomam um mesmo referente já citado, tampouco são resolvidas apenas por articulações extratextuais. De toda forma, tais reflexões são apresentadas no sentido de situar de modo bastante claro o ponto sobre o qual esta investigação se pauta, trabalhando com anáforas correferenciais diretas, portanto, considerando a noção estabelecida classicamente para o termo.

Para fechar esta seção, vale ponderar, ainda, sobre o fato de que, em um sentido amplo, é possível considerar que a anáfora contemple processos prospectivos e retrospectivos. Um olhar mais detido, no entanto, separa as construções anafóricas daquelas denominadas catafóricas (TEIXEIRA, 2013). Estas dizem respeito a um termo que precede o referente, ou seja, enquanto a anáfora remete para trás, a catáfora remete para frente no tecido textual, gerando processos prospectivos. Neste trabalho, abordamos a anáfora em sentido estrito, diferenciando tais processos daqueles catafóricos.

Por fim, ter em mente que existem diferentes formas de construir as retomadas anafóricas em um texto é também relevante para, mais adiante, compreendermos o que investigações têm encontrado acerca das funções para o sintagma “o mesmo”, chave na pesquisa que aqui se apresenta.

### 2.3 A ANÁFORA CORREFERENCIAL EM LEITURA

Tendo apresentado as formas de construir um referente e uma anáfora no texto, nesta seção, discorreremos sobre o estudo do processamento anafórico nas teorias de leitura. Para começar a pensar sobre as reativações cognitivas proporcionadas pelas anáforas correferenciais, é preciso, primeiro, compreender que o antecedente foi acessado e ativado pelo leitor. Essa ativação ocorre quando o antecedente é introduzido no texto. Quando se trata de um texto expositivo, é comum que seja uma ideia ou um tópico (KOCH; ELIAS, 2018) sobre o qual se discorrerá. O leitor, então, a partir do texto, mobiliza seus conhecimentos acerca daquele referente específico, manipulando-o enquanto a leitura segue e as integrações sintáticas ocorrem. Como entendemos que os processos acontecem em paralelo e de modo interativo, sabemos que o antecedente é acionado, mas, ao mesmo tempo, outros níveis linguísticos e cognitivos estão interagindo para que a compreensão do texto seja construída. Então, quando o

leitor se depara com uma anáfora correferencial, o antecedente, antes acessado, é retomado e os processos de integração entre diferentes partes do texto vão sendo concretizados.

A retomada de um referente no decorrer de um texto possibilita, então, a integração de diversas porções textuais. Desse modo, as anáforas correferenciais são um dos elementos responsáveis por esses processos de integração, que indicam ao leitor as relações entre uma e outra informação apresentada (JUST; CARPENTER, 2013). De acordo com Oakhil, Cain e Elbro (2017, p. 104),

compreendedores bem-sucedidos integram os significados de orações e frases sucessivas em um texto à medida que constroem uma representação baseada na memória de significação do texto. Esses processos de integração podem ser orientados pelo conhecimento linguístico; em particular, por anafóricos e conectivos.

Para Kintsch e Rawson (2013), o que ocorre é que as unidades de ideias que compõem o texto se inter-relacionam de modo complexo no nível microestrutural e, como vimos, uma das formas de concretizar tais relações é por meio das anáforas correferenciais. Estas auxiliam na formação de um modelo mental do lido, uma vez que atuam na integração das informações fornecidas pelo texto e são pistas que dão subsídios ao leitor para a geração de inferências.

Ao longo de um texto, diversas entidades são introduzidas e, conforme essas entidades vão sendo retomadas, tem-se uma progressão referencial (KOCH; ELIAS, 2015). A anáfora é um elemento que atua nessa progressão, formando os laços coesivos de um texto, uma vez que ela é responsável por manter um determinado referente ativo na memória de modo que o leitor possa relacionar as diferentes partes que compõem o texto e construir, dessa forma, a base textual. Assim, a relação coesiva entre um antecedente e a anáfora que o retoma auxilia na formação da textualidade (HALLIDAY; HASAN, 1976), fundamental para que o leitor possa elaborar uma representação coerente do texto lido. Dessa forma, a anáfora é um dos conhecimentos linguísticos que atua para orientar os processos de integração durante a leitura (OAKHIL; CAIN; ELBRO, 2017). Mais especificamente, quando se trata da anáfora correferencial, as informações do antecedente estão disponíveis no próprio texto. O leitor pode, então, usá-las na construção de sua compreensão.

As anáforas auxiliam, portanto, na formação de um modelo mental daquilo que está sendo lido, pois, a partir delas, é possível que o leitor conecte as diferentes partes de um texto. O fato de a anáfora atuar na retomada de referentes introduzidos no texto faz com que ela seja um recurso de progressão textual. Isso significa que o processamento e a resolução adequada das anáforas são fundamentais para a compreensão em leitura, pois não há como construir uma representação mental do que se lê sem que haja uma compreensão local e global do texto. No

entanto, a relação entre a construção de um modelo mental coerente e a resolução das anáforas textuais não é direta. Isto é, não basta apenas que um leitor consiga resolver anáforas contidas em um texto para que a compreensão se dê, mas é necessário que tais resoluções aconteçam e que sejam feitas de modo adequado, porque isso está interconectado a outros processos, como já explicado. Caso o leitor verifique algum problema nesse nível interfrasal, é necessário resolvê-lo em algum momento da leitura, pois, do contrário, a integração entre as partes do texto pode ficar prejudicada.

A integração textual é possibilitada pelo processamento dessas estruturas, dentre outras, que atuam em um nível mais local e que formam a microestrutura de um texto. Assim, a compreensão leitora é construída a partir da integração de informações que compõem níveis inferiores, mais locais, e informações que compõem níveis superiores. Nesse sentido, ao se estudar o processamento em leitura de anáforas correferenciais, há que se considerar o processamento da anáfora em nível local e, a partir disso, o seu papel no processamento global de um texto (que se dá por meio da integração proporcionada por esse recurso linguístico).

No que diz respeito ao processamento anafórico em si mesmo, trata-se de se refletir acerca de um nível local, buscando, conforme Leitão (2015, p. 47), entender “como as relações anafóricas ocorrem em termos de processos cognitivos (mentais/cerebrais)”. Para tanto, procura-se investigar “[...] como funcionam essas relações cognitivamente e quais fatores estão em jogo no momento em que ouvimos ou lemos textos que contenham essas anáforas” (LEITÃO, 2015, p. 47).

Por fim, para situarmos o entendimento que esta pesquisa tem sobre leitura e o nível em que a anáfora se encontra, trazemos o diagrama a seguir, formulado a partir de tudo o que foi percorrido até este momento, considerando as três peças do quebra-cabeça que participam do processo da leitura:



conhecer o que as pesquisas têm encontrado quando se trata de processar a anáfora em estímulos de leitura.

### **2.3.1 As investigações sobre o processamento anafórico em leitura**

Com o intuito de compreender o que tem sido evidenciado acerca dos processos envolvidos na resolução da anáfora, nesta seção do texto, são apresentados alguns dos estudos que investigam o processamento anafórico em estímulos de leitura no português brasileiro. As discussões em torno dessa temática são caras à Psicolinguística e, por conta disso, pesquisas vêm sendo desenvolvidas ao longo dos anos, com o objetivo de se conhecer o processamento da anáfora e de se obterem respostas sobre o seu funcionamento na cognição. Essas pesquisas têm se voltado para a atuação da anáfora em um nível frasal ou proposicional, buscando pelo processamento micro (local – que é o ponto em que a anáfora atua). O que ocorre é que esse processamento micro respinga nos processos de níveis superiores envolvidos na construção da compreensão de um texto<sup>22</sup>. Franzen e Souza (2020), em revisão sobre os estudos que investigam o processamento anafórico, consideram os resultados dessas pesquisas e fazem uma ponte com o ensino da leitura em sala de aula, ponderando sobre o papel desse aspecto local nos processos de integração e de progressão sequencial de um texto. Assim, ainda que os estudos desenvolvidos não toquem em processos de compreensão de texto, dão subsídios para discutir sobre esse aspecto. Nas pesquisas ora relatadas, é possível observar que fatores sintáticos, fatores semânticos, tipo de retomada e, ainda, a população estudada são levados em conta quando se pensa a atuação da anáfora cognitivamente. Há que se ressaltar, também, as diferenças entre os estudos que tratam de tal fenômeno no escopo da sentença, considerando restrições gramaticais específicas, e aqueles que se centram em um escopo textual (como o nosso), voltando-se para questões que envolvem o processamento interfrasal e considerando níveis superiores de processamento.

De modo mais específico, o objetivo desta seção é apresentar estudos que consideram aspectos relativos ao processamento anafórico, seus pontos de encontro e de desencontro, para que se possa construir o caminho trilhado até então e, dessa forma, situar a presente pesquisa dentro desse contexto investigativo. Esses estudos foram selecionados a partir de buscas realizadas na Biblioteca Digital Brasileira de Teses e Dissertações e no Portal de Periódicos Capes. Essas bases foram consultadas com a finalidade de conhecer o que vinha sendo

---

<sup>22</sup> Compreendendo o texto como um lugar em que um conjunto de ideias é construído de modo entrelaçado, formando uma teia conectada de sentidos.

investigado em torno do processamento anafórico no Brasil. Com isso, foram selecionados trabalhos que tivessem estudado o processamento anafórico em estímulos de leitura em uma perspectiva psicolinguística.

Dos estudos que buscam investigar o processamento anafórico no português-brasileiro, Cunha-Lima (2004)<sup>23</sup> investiga os processos de referenciação no que diz respeito às expressões indefinidas anafóricas. Nessa pesquisa, a anáfora é construída com a repetição do substantivo, precedido por um artigo indefinido ou definido. O objetivo primeiro do estudo é “estabelecer em quais condições uma expressão nominal indefinida pode receber leitura anafórica” (2004, p. xi). A autora realiza um experimento de leitura automonitorada a partir de um conjunto experimental com duas estruturas de retomada anafórica: uma continha pares de sentenças com frases nominais<sup>24</sup> e, outra, pares sentenças com verbo finito<sup>25</sup> no segundo elemento do par. No experimento, a leitura do verbo nas sentenças com expressões definidas se mostrou significativamente mais rápida que a leitura do verbo nas sentenças com expressões indefinidas. Tal resultado, segundo a autora, leva a pensar que ocorre, na leitura do verbo nas sentenças com indefinido, um processamento adicional por conta do estabelecimento de um referente novo. Já nas sentenças em que não havia verbo (frases nominais), a média do tempo de leitura foi menor para aquelas que contêm os indefinidos. Os resultados obtidos com o experimento reafirmam o fato de a expressão nominal não ter seu valor estabelecido somente no momento do seu processamento, mas, também, a depender do verbo e dos eventos expressos na sentença, dependendo, assim, de processos de integração.

O que a pesquisadora percebe é que o indefinido é menos informativo e muito sensível ao contexto. Além disso, outro ponto bastante relevante citado no trabalho é a relação entre a organização macro-textual dos eventos e a organização referencial de um texto. A partir do foco dado pela autora nos experimentos realizados, os resultados apontam para a presença de um forte elemento *top-down* (a dinâmica de introdução de eventos no texto), que age sobre o estabelecimento da referência. Por fim, a autora destaca que o indefinido somente será anafórico quando não for argumento de nenhum verbo finito que expresse um evento diferente daquele introduzido pelo antecedente. Esses achados podem, ainda, ser relacionados aos custos para a memória de trabalho. Ao encontrar uma expressão indefinida, aliada a um verbo, o leitor

<sup>23</sup> Sabemos haver estudos (MELO, 2003; MAIA, 1994) que já investigavam o processamento anafórico antes da primeira investigação apresentada nesta revisão. No entanto, tais estudos não retornaram com as buscas realizadas nos portais citados.

<sup>24</sup> Indefinido e frase nominal: “Meu gato caçou um rato. **Um rato** grande e gordo”. Definido e frase nominal: “Meu gato caçou um rato. **O rato** grande e gordo”. (CUNHA-LIMA, 2004, p.188).

<sup>25</sup> Indefinido e verbo finito: “Meu gato caçou um rato. **Um rato** correu porta a fora”. Definido e verbo finito: “Meu gato caçou um rato. **O rato** correu porta a fora”. (CUNHA-LIMA, 2004, p.188).

interpreta como uma informação nova (um novo evento) e, dentro do contexto das sentenças apresentadas, a informação nova pode não fazer sentido ou, então, requerer a manutenção de um conjunto maior de informações para que o sentido possa ser inferido na continuidade da leitura.

Há estudos (LEITÃO, 2005; QUEIROZ, 2009; GADELHA, 2012; TEIXEIRA, 2013; MAIA, 2013, GONDIM, 2017), ainda, que buscam pelas diferenças no tempo de processamento de distintas formas de retomadas anafóricas. Esses estudos são de ampla relevância para a presente investigação, já que olhamos, também, para o tempo de processamento de duas formas de retomada, ainda que, nesse caso, acreditemos que não haja diferenças – e é isso que esperamos encontrar.

Nos estudos encontrados, são investigadas anáforas realizadas com pronomes, nomes repetidos, retomadas nulas, hipônimos, hiperônimos. Tudo isso é feito levando em conta outros fatores envolvidos no processamento em leitura, como: posição sintática de antecedente e referente, animacidade e a relação entre as sentenças experimentais (coordenadas, justapostas, subordinadas), ou seja, são investigados se e como os fatores relativos ao texto (responsáveis por um processamento ascendente) influenciam no processamento da anáfora.

Nessa dinâmica, Leitão (2005) compara o tempo de leitura de anáforas realizadas em posição de objeto com pronomes, nomes repetidos e com objetos nulos em estruturas coordenadas. Os resultados mostraram que os pronomes são lidos de modo mais rápido que os nomes repetidos e que tanto os pronomes quanto as retomadas nulas reativam seus antecedentes. Com esses resultados, Leitão (2005) argumenta que a penalidade do nome repetido<sup>26</sup> (GORDON; GROSZ; GILLIOM, 1993) ocorre também em posição de objeto; não somente em posição de sujeito, como era sugerido inicialmente pela Teoria da Centralização ou, mesmo, pela Hipótese da Carga Informacional<sup>27</sup> (ALMOR, 1999; 2000). Além disso, embora não fosse

---

<sup>26</sup> O termo foi proposto por Gordon, Grosz e Gilliom (1993), os quais compreendiam que todo enunciado possui um centro discursivo. Nessa perspectiva, a partir de uma série de experimentos de leitura automonitorada, evidenciaram que, na língua inglesa, quando a anáfora retoma esse centro discursivo (em posição de sujeito), os pronomes são as entidades mais adequadas para realizar a correferência, ou seja, eles geram menores custos para o processamento. Já os nomes repetidos acabam por gerar mais custos, o que os teóricos denominaram de Penalidade do Nome Repetido. Resumidamente, o que os resultados dos experimentos realizados sugerem é que o processamento das formas referenciais é suscetível à posição sintática de seus antecedentes dentro de um enunciado. A comparação para a observação da penalidade foi realizada entre pronomes plenos e nomes repetidos.

<sup>27</sup> Os pronomes são processados mais rapidamente que nomes repetidos, quando o antecedente está em posição de sujeito, por conta da carga informacional contida em cada um desses elementos. Um nome, por si só, conteria mais informações semânticas a serem processadas. Já os pronomes contêm menos informações semânticas. Isso geraria custos menores para a memória de trabalho e, conseqüentemente, o processamento seria mais rápido. O mesmo ocorreria com sintagmas nominais mais gerais em relação ao antecedente e com os mais específicos com relação ao antecedente. Os sintagmas mais gerais seriam processados mais rapidamente, pois os mais específicos carregariam informações semânticas novas, o que sobrecarregaria a memória de trabalho.

o foco central do estudo, é possível destacar que os pronomes pessoais do caso reto são processados sem que haja maiores custos por estarem na posição de objeto.

Outro fator analisado por Leitão (2005) foi a animacidade e o paralelismo estrutural. Os resultados sugerem uma atuação significativa desses fatores no processamento da anáfora. Isso significa que a animacidade do antecedente e a posição sintática de antecedente e retomada podem atuar no processamento da anáfora no momento da leitura. Leitão (2005) investigou, também, o processamento de anáforas realizadas com hiperônimos e com hipônimos. O que os resultados mostram é que sintagmas mais gerais (hiperônimos) são lidos mais rapidamente que sintagmas mais específicos (hipônimos). Isso pode ocorrer pelo fato de os primeiros conterem menos traços semânticos e, assim, sobrecarregarem menos a memória de trabalho e, portanto, serem processados mais rapidamente.

Após o estudo desenvolvido por Leitão (2005), muitos outros surgiram com a finalidade de investigar o processamento dessas diferentes realizações anafóricas. Queiroz (2009) investiga, também, a eficiência do processamento da anáfora realizada com pronomes e com nomes repetidos, e com hiperônimos e hipônimos. A pesquisadora encontra resultados que mostram uma leitura mais rápida de pronomes do que de nomes repetidos nas condições investigadas, destacando uma tendência à ocorrência da penalidade do nome repetido para a posição de sujeito. Resultados como esses têm confirmado que a anáfora realizada com a repetição do antecedente tende a gerar maiores custos para o processamento. Além disso, no que diz respeito às relações de hiperonímia e de hiponímia, os resultados mostram que as retomadas feitas com hiperônimos foram lidas mais rapidamente que as retomadas feitas com hipônimos, indo em direção ao encontrado por Leitão (2005).

Outro estudo (GADELHA, 2012) buscou averiguar o processamento da anáfora realizada com pronomes e nomes repetidos em uma segunda língua (no caso desse estudo específico, o francês), testando públicos com diferentes níveis de proficiência. Segundo a pesquisa, os aprendizes iniciantes não demonstraram diferença significativa na leitura do tipo de retomada. O grupo de nível intermediário, por sua vez, mostrou diferença significativa, sendo o nome repetido lido de modo mais rápido que o pronome. Já o de nível avançado lê de modo mais rápido o pronome. Ao final da leitura de cada sentença, eram feitas perguntas para verificar se o participante havia compreendido o que leu. O índice de acerto dos grupos diante das perguntas foi alto, ou seja, o resultado *off-line* não teve muita variação, mas o processamento dos elementos anafóricos foi diferente em cada um dos níveis. Gadelha (2012) realizou, então, um segundo experimento com falantes nativos de francês e observou que eles



processaram a anáfora de modo mais rápido do que os participantes experientes do primeiro experimento. Além disso, verificou, também, a penalidade do nome repetido.

O que esses resultados sugerem é que o aumento da proficiência em uma língua pode modificar o custo de processamento e, quanto maior a proficiência, mais o processamento anafórico é facilitado com o uso de pronomes. Ao se contrastar os estudos até esse ponto reportados, percebe-se que os falantes fluentes de português e francês têm mostrado maiores custos para processar nomes repetidos, sendo os pronomes plenos elementos facilitadores. As sentenças com as quais Gadelha (2012) trabalhou possuíam antecedente na posição de sujeito e retomada também em posição de sujeito. Os resultados desses estudos mostram que, para além de como as sentenças são construídas, o nível de proficiência em uma determinada língua implica diferentes custos para esse processamento.

O estudo de Maia (2013) é outro que busca investigar o processamento do nome repetido, do pronome e do nulo como anáforas correferenciais. Os conjuntos experimentais e a forma de medição do tempo de resposta (a partir da leitura da frase toda) seguem a mesma perspectiva utilizada em estudos realizados por Lezama (2008) e por Gordon, Grosz e Gilliom (1993), distanciando-se dos conjuntos experimentais e da forma de medição de tempo de resposta usados nos estudos apresentados anteriormente<sup>28</sup>. O pesquisador busca estudar o processamento de nomes repetidos, pronomes plenos e pronomes nulos em contextos de antecedentes salientes e não salientes. A pesquisa mostrou que, nos contextos em que o antecedente está em posição de sujeito, os nulos são processados mais rapidamente. Os nomes repetidos, por sua vez, foram penalizados com relação aos pronomes nulos, mas foram favorecidos em comparação com os pronomes plenos. Houve, assim, uma penalidade do pronome pleno (ele), mas não uma penalidade do nome repetido. Quando o antecedente estava na posição de objeto (condição não saliente), os pronomes plenos e nulos se mostraram os menos favorecidos, ambos foram igualmente penalizados. Já os nomes repetidos, nesse contexto, foram as formas mais fáceis de serem processadas (tanto em comparação com pronome pleno quanto com pronome nulo). Assim, tal estudo, tendo método distinto, alcança resultados divergentes de estudos anteriores (LEITÃO, 2005; QUEIROZ, 2009; GADELHA,

---

<sup>28</sup> Exemplos de frases usadas por Maia (2013) nos diferentes experimentos: “João se encontrou com Maria. Maria/Ela/Ø achou ele triste.”; “Maria se encontrou com João. Maria/Ela/Ø achou ele triste.”; “João se encontrou com Maria. Maria/Ela o achou triste/ Achou-o triste.”; “Maria se encontrou com João. Maria/Ela o achou triste/ Achou-o triste.”; “Miguel deixou Paula no aeroporto internacional. Paula/Ela o abraçou pela última vez na vida/ Abraçou-o pela última vez na vida.”; “Paula deixou Miguel no aeroporto internacional. Paula/Ela o abraçou pela última vez na vida/ Abraçou-o pela última vez na vida.”; “Saulo deixou Ana no aeroporto internacional. Ana/Ela chorou bastante antes da despedida.”; “Ana deixou Saulo no aeroporto internacional. Ana/Ela chorou bastante antes da despedida.”

2012). Além da diferença existente no conjunto experimental, Maia (2013) mediu o tempo de leitura da sentença toda, diferentemente dos estudos já citados, que trabalham com medidas de tempo de leitura por palavra, computando o tempo de leitura específico da anáfora. Tal observação é importante, pois a forma de aferir o tempo impacta nos resultados, como veremos mais adiante em estudo realizado por Gondim (2017).

Teixeira (2013) trabalhou com as preferências sintáticas e semânticas no processamento anafórico e catafórico, diferenciando esses dois processos de acordo com o que apresentamos na seção 2.2 deste trabalho. O método usado foi o rastreamento ocular. Para investigar preferências semânticas, a pesquisadora utilizou o mesmo conjunto experimental de Leitão (2005), comparando, a partir de sentenças coordenadas, o custo de processamento da anáfora construída com hipônimos e com hiperônimos. Na investigação de Teixeira (2013), foi feita uma divisão de grupos de participantes. Um grupo leu anáforas na posição de sujeito e o outro na posição de objeto. O tempo total e o número de fixações obtidos permitem afirmar que o custo de processamento da anáfora em relação de hiperonímia com seu antecedente tende a ser mais baixo que o da anáfora em relação de hiponímia. Isso vai ao encontro dos achados de Leitão (2005) para as anáforas em posição de objeto, embora, no estudo de Teixeira (2013), a diferença entre hiperônimo e hipônimo na posição de objeto não tenha sido significativa.

Outro experimento realizado pela pesquisadora teve por foco as preferências sintáticas no processamento anafórico de pronomes plenos e nulos. A partir de estruturas ambíguas em sentenças subordinadas, o experimento buscou saber com qual elemento o participante tinha feito a correferência (se o mais saliente ou o menos saliente). No que diz respeito aos dados de movimentação ocular, foram encontrados custos menores para as condições testadas de pronomes nulos. Os testes estatísticos feitos no trabalho mostram uma interação significativa entre a condição de pronome e a condição de posição sintática. Ao fazer uma análise detalhada das medidas do tempo total de fixação e da duração média da fixação na região relativa ao sujeito na oração principal, a autora observou que os participantes fizeram, em média, três visitas à região do sujeito durante a leitura das frases. No que diz respeito à região do objeto, cada fixação ocular foi mais curta, mas os participantes parecem ter passado mais tempo nessa região, porque a soma do tempo de visitas foi um pouco mais de quatro vezes maior.

Nesse experimento, fica evidente que, em mais de 90% dos casos, os participantes relacionam o pronome nulo ao sujeito da sentença, o que corrobora os achados de Maia (2013). Já os pronomes plenos não se mostraram aptos para resolver o problema da ambiguidade apresentado no teste experimental. Nessa etapa da pesquisa, os resultados revelaram haver um

tempo maior de reação para a leitura das condições com o pronome pleno (ele) se comparadas às condições com o pronome nulo.

Outro estudo que analisa o processamento de pronomes e nomes repetidos é o de Nóbrega-Lima (2014). A autora relaciona o paralelismo<sup>29</sup> e o foco<sup>30</sup> estrutural ao processamento de pronomes e nomes repetidos, realizando um experimento em que o antecedente está em foco e em que há paralelismo, e outro experimento em que o antecedente continua em foco, mas sem paralelismo sintático entre referente e retomada. No primeiro experimento, o foco estrutural não mostrou efeito significativo, mas o tipo de retomada teve efeito significativo, sendo o pronome lido mais rapidamente. Os resultados obtidos no segundo experimento mostraram um efeito significativo do foco estrutural, ou seja, o processamento da retomada foi menos custoso quando o antecedente estava em foco. A autora esperava tal efeito, tendo em vista que, com o foco estrutural, o antecedente fica ativo na memória de trabalho de forma proeminente.

A partir, então, das constatações dos experimentos, Nóbrega-Lima (2014) questiona se pode haver hierarquia no que tange aos fatores que atuam no processamento de modo a facilitar o estabelecimento da correferência. Nesse caso, o paralelismo se sobreporia ao foco. Esse experimento indicou, também, efeito significativo para o tipo de retomada quando o antecedente estava focalizado, confirmando a penalidade do nome repetido. Quando o antecedente não estava focalizado, o efeito não foi significativo, mas houve maiores tempos de leitura para os nomes repetidos do que para os pronomes, o que corrobora a penalidade do nome repetido. Leitão (2005), em um de seus experimentos, também trabalha com o efeito de paralelismo, relacionando-o à animacidade em retomadas anafóricas feitas pelo pronome. Talvez a principal constatação feita, na mesma direção da pesquisa de Nóbrega-Lima (2014), é que fatores estruturais são atuantes no processamento anafórico, inclusive o pronominal (diferente do que trazem algumas teorias: HANKAMER; SAG, 1976; SAG; HANKAMER, 1984 – para as quais apenas fatores semânticos atuariam no processamento pronominal). Assim, a estrutura da sentença pode contribuir para que determinado referente fique ativo na memória de trabalho e seja recuperado mais facilmente.

A leitura é uma atividade cognitiva complexa que envolve diversas subtarefas, dentre elas o estabelecimento adequado da correferência, isto é, para que o leitor possa compreender

---

<sup>29</sup> “Ocorre paralelismo estrutural quando um elemento correferencial tem a mesma função e posição sintática que seu antecedente.”. (NÓBREGA-LIMA, 2014, p.36).

<sup>30</sup> A pesquisadora explica detalhadamente os diferentes conceitos de foco, mas trabalha somente com a clivagem para garantir a posição de foco estrutural.

o sentido do que está sendo lido, é primordial que, ao se deparar com uma anáfora, seja encontrado o antecedente adequado. Tudo isso deve ser feito muito rapidamente, a partir das informações que estão ativas na memória de trabalho. O que se percebe dos resultados até então apresentados é que existem elementos e estruturas que são mais favoráveis a essa ativação rápida, o que pode atuar positivamente na formação de uma representação mental coerente acerca do texto (quando se trata da leitura de textos).

Outro estudo realizado (SIMÕES, 2014) buscou pela influência da coesão e da coerência no processamento anafórico de pronomes e nomes repetidos. A pesquisadora trabalhou com as variáveis coesão (usando um conectivo) e coerência (item linguístico congruente ou incongruente). No primeiro experimento, a retomada anafórica é realizada com o pronome pleno (ele) e, no segundo, com o nome repetido. A principal observação feita está no fato de que “a combinação entre incongruência e retomada anafórica produz diferentes efeitos a depender do tipo de anáfora processada” (SIMÕES, 2014, p.86). Quando a incongruência está relacionada a uma anáfora pronominal, o processamento da anáfora não parece ter sido mais custoso. O mesmo não ocorre com a retomada anafórica com nome repetido, cujo processamento parece ser mais custoso. A autora explica tal resultado com base na carga informacional contida nas diferentes formas anafóricas. O nome repetido conteria mais informações e o seu uso, naquele contexto experimental, não se justificaria. Dessa forma, a junção de uma incongruência mais o uso anafórico do nome repetido aumentaria os custos para o processamento, exigindo mais da memória de trabalho. Tem-se, assim, mais um estudo que mostra um tempo de leitura maior para nomes repetidos do que para os pronomes, em consonância com Leitão (2005), Nóbrega-Lima (2014), Queiroz (2009), Gadelha (2012) e outros.

Tomando por base esses distintos resultados em torno da penalidade do nome repetido (NÓBREGA-LIMA, 2014; MAIA, 2013; GADELHA, 2012; QUEIROZ, 2009; LEITÃO, 2005), Gondim (2017) leva em conta as diferenças metodológicas nas pesquisas e realiza experimentos para entender em quais condições tal penalidade acontece no PB. A partir dos resultados encontrados, a pesquisadora chega à conclusão de que, para que a penalidade do nome repetido seja constatada em português brasileiro, é necessário considerar um conjunto de fatores linguísticos e metodológicos, sendo, portanto, um efeito multifatorial em que a forma de segmentação/aferição do tempo do conjunto experimental é o principal fator de influência, além da quantidade de antecedentes humanos possíveis de serem retomados. Assim, o que a pesquisadora confirma é a existência de penalidade do nome repetido em português brasileiro, quando se compara o processamento de pronome e de nome repetido, a depender da

configuração experimental e das diferenças linguísticas entre os estímulos testados. Os resultados sugerem que o custo para o processamento da anáfora varia conforme as condições apresentadas ao leitor. Isso explicaria a diferença encontrada nos trabalhos anteriores, que foram realizados a partir de material experimental e de aspectos relativos ao método distintos, e mostra a importância das escolhas experimentais no momento de realizar um estudo.

Seguindo em uma perspectiva que considera a anáfora no contexto de recepção escrita, Cavalcante (2017) investigou a influência da correferência anafórica e do humor na compreensão leitora, a partir de teste *cloze* e de teste de compreensão em leitura. De todos os estudos encontrados, esse é o único que usou porções textuais maiores (para além de apenas duas sentenças) em seu desenho experimental, trabalhando de modo mais próximo com a concepção de texto considerada nesta tese. A pesquisadora não teve por foco formas de realizações anafóricas específicas e nem considerou questões semânticas e sintáticas, que poderiam influenciar o processamento. O esperado era que textos com anáforas correferenciais e com humor favorecessem a compreensão dos participantes. No entanto, não foi possível validar tal hipótese a partir dos resultados encontrados. Isso pode ter ocorrido em razão de a forma de realização da anáfora não ter sido considerada. A pesquisadora apenas evitou o trabalho com a referência nula. O que a pesquisa concluiu foi que a compreensão textual se deu por meio da operação desses processos (correferência anafórica e humor), mas não da forma esperada e apresentada nas hipóteses. Os resultados dessa investigação abrem portas para muitas discussões em torno do papel da anáfora correferencial no processamento de um texto. Se retomarmos o que os outros estudos apresentados nesta seção trazem em seus resultados, temos mais subsídios para entender como ela é processada. O que podemos concluir é que não basta que haja relações anafóricas em um texto, é necessário que essas relações sejam realizadas de modo a favorecer seu processamento; do contrário os custos para a compreensão podem acabar se tornando maiores. Essas reflexões sobre o processamento das anáforas correferenciais e sobre sua atuação nos processos de compreensão contribuem para o questionamento que fazemos acerca da influência do uso do sintagma “o mesmo” como anáfora correferencial na compreensão em leitura. Isso porque estamos estudando um elemento que não possui característica de anáfora correferencial em sua essência, como veremos na seção 2.4.1; por outro lado, trata-se de um elemento que tem aparecido constantemente nessa função. Então, faz-se relevante entender como funciona esse processamento na leitura de um texto e como isso implica a compreensão.

Há, ainda, estudos que buscam investigar o processamento anafórico tendo como variável independente a população envolvida. Nessa perspectiva, Vasconcelos (2012) investiga

se o pronome (ele) é processado mais prontamente do que nomes repetidos em sujeitos sem queixa de linguagem e em pacientes com afasia de expressão. Além disso, verifica se sintagmas nominais hiperônimos são processados mais rapidamente do que hipônimos. Os experimentos foram realizados com a técnica de leitura automonitorada que seguiu o protocolo utilizado por Leitão (2005). Os resultados mostram que a diferença do tempo de leitura entre as condições de pronome e nome repetido foram significativas para o grupo controle, evidenciando uma penalidade do nome repetido e corroborando outros estudos realizados na área, citados anteriormente.

No que tange ao grupo experimental, não houve diferença significativa nos tempos de leitura de pronomes e de nomes repetido, não tendo sido evidenciada, portanto, a penalidade do nome repetido. De acordo com Vasconcelos (2012, p. 91, comentários nossos), tal resultado mostra que “não há uma condição em que se processe mais prontamente a correferência; ela é estabelecida de igual forma tanto para condição PR [pronome] quanto para NR [nome repetido]”. Ao comparar os tempos de leitura em cada condição entre os grupos, obteve como resultado uma diferença significativa que mostrou que o grupo experimental demora mais tempo para realizar a leitura dos segmentos críticos. No que diz respeito ao índice de acertos às perguntas realizadas ao final da leitura de cada frase, o grupo experimental teve um percentual de acertos significativamente menor do que o grupo controle. Sobre o tempo de resposta, os sujeitos do grupo experimental também demoraram mais tempo que os do grupo controle para apresentar a resposta.

Vasconcelos (2012), ainda, analisa de modo comparativo os segmentos relativos ao antecedente e à retomada por nome repetido, com a intenção de confirmar se a correferência foi, de fato, estabelecida. O esperado era que fosse obtido um menor tempo de leitura para o segmento relativo ao antecedente do que para o segmento que continha a retomada, resultado confirmado para o grupo controle, o que mostra o estabelecimento da correferência. O grupo experimental, por sua vez, não mostrou diferenças significativas no tempo de leitura do antecedente e da retomada. Esse resultado sugere que os participantes leram antecedente e retomada da mesma forma, interpretando a retomada como se fosse uma informação nova que aparecia pela primeira vez na frase.

A partir desses resultados, foi realizado o segundo experimento que buscou analisar o processamento semântico dos participantes da pesquisa. Nesse experimento, as condições testadas foram as de retomadas feitas com sintagmas em relação de hiperonímia e hiponímia com o antecedente. Os resultados mostraram um efeito principal da variável grupo e da variável tipo de retomada, sugerindo, também, efeito marginal de interação entre as variáveis. Tanto

grupo controle quanto experimental leram de modo mais rápido a condição de retomada feita com hiperônimo. O grupo controle, por sua vez, foi significativamente mais rápido na leitura de ambas as condições. Ainda que o grupo experimental também tenha lido mais rapidamente a condição de retomada feita com hiperônimo, a diferença nos tempos de leitura entre retomadas feitas com hiperônimo e com hipônimo não se mostrou significativa. A comparação da diferença nos tempos de leitura entre os grupos se mostrou significativa, sugerindo que os participantes com afasia processam de modo mais lento qualquer uma das formas de retomada.

Quanto aos dados das respostas às perguntas feitas ao final de cada frase, o grupo controle mostrou um índice de acerto significativo em ambas as condições – sendo maior na condição de retomada feita por hiperônimo –, o que evidencia o estabelecimento da correferência. O grupo experimental também apresentou um índice alto de acertos às perguntas, mostrando ter estabelecido a correferência, tanto na condição de hipônimo quanto na de hiperônimo. No grupo experimental, contudo, o índice de acertos foi maior nas respostas das frases cuja retomada era feita por um sintagma hipônimo, diferente do que ocorre no grupo controle. Com isso, mesmo que não seja uma diferença significativa, há indícios de que os participantes com afasia realizam a correferência com o sintagma mais específico. De modo geral, a pesquisadora infere que pacientes com afasia de expressão conseguem fazer as relações semânticas que possibilitam a correferência, mas isso é evidenciado somente na análise dos dados obtidos off-line, pois os dados obtidos pelas medidas on-line não mostram sensibilidade aos dois tipos de retomada.

Na pesquisa de Alves (2012), a comparação é estabelecida entre pessoas com a doença de Alzheimer e pessoas sem a doença (idosos saudáveis). O primeiro experimento mostrou que os idosos com doença de Alzheimer apresentam tempos de leitura maiores que os idosos sem a doença. Além disso, os resultados revelam uma diferença significativa no grupo experimental no que diz respeito ao tipo de retomada, sendo os nomes repetidos lidos de modo significativamente mais rápido que os pronomes plenos, ou seja, para os idosos com a doença, o nome repetido é um facilitador. Já para os idosos sem a doença de Alzheimer, a penalidade do nome repetido continua a ocorrer. Esses resultados relatados pelo pesquisador indicam uma tendência de preservação do sistema semântico (pelo menos na fase inicial da doença), mas o fato de os idosos com doença de Alzheimer não apresentarem um bom desempenho nas retomadas feitas com pronomes indica um déficit na memória de trabalho, o que foi evidenciado pelo protocolo de rastreamento com testes de memória.

O autor compara também os dados obtidos pelo grupo controle (idosos sem patologia) com o grupo de jovens universitários investigados por Leitão (2005). Ao fazer esse contraponto,

é possível perceber que os idosos demonstram precisar de um tempo maior para realizar o processamento da retomada anafórica, mesmo não tendo nenhuma patologia.

No segundo experimento, o autor trabalha com hipônimos e hiperônimos como formas de retomada anafórica. Os idosos sem patologias processaram mais rapidamente hiperônimos do que os hipônimos, seguindo o mesmo padrão apresentado nos estudos citados anteriormente. Já para os idosos com Alzheimer, os tempos de leitura encontrados para as retomadas feitas com hiperônimos também foram menores, mas as diferenças (quando considerado o intra-grupo) não foram significativas (entre hiperônimo e hipônimo), ou seja, eles não se mostraram sensíveis a nenhuma das condições. De acordo com o autor, uma possível explicação para isso seria o comprometimento da memória de trabalho.

Mais uma vez, o autor compara os resultados do experimento dos idosos sem patologia com os dados obtidos por Leitão (2005), que estudou jovens saudáveis, percebendo que ambos os grupos processam mais rapidamente os sintagmas hiperônimos, mas há uma significativa diferença nos tempos de leitura entre os diferentes grupos. Os jovens leem de modo mais rápido as duas formas de retomada (hipônimo e hiperônimo). Por fim, a partir dos experimentos realizados, o autor confirma os achados de Almor et al (1999), que relacionam o comprometimento na memória de trabalho com as diferenças no processamento anafórico, mostrando que esse sistema de memória tem atuação fundamental no processamento da anáfora.

Correia (2014), por sua vez, investiga a relação entre a memória procedimental e a linguagem em pessoas que gaguejam, trabalhando com dois grupos: um composto por pessoas que gaguejam e outro por falantes fluentes. O experimento de leitura automonitorada, em que as anáforas foram construídas com pronomes e nomes repetidos, mostrou uma ausência de efeito principal significativo para a variável grupo, o que sugere que não há diferença estatística entre os tempos médios de leitura da retomada anafórica por pessoas que gaguejam e por falantes que não gaguejam. Houve, contudo, efeito significativo para a variável tipo de retomada, mostrando que os pronomes foram processados mais rapidamente que o nome repetido. A autora explica esses resultados por conta do tipo de sentença utilizada no experimento, cuja retomada é feita na intersentença. Nessa condição, no entanto, somente parte da correferência anafórica é regida por operações e princípios estritamente gramaticais, “uma parte significativa dela é determinada por aspectos relativos à memória de trabalho e a aspectos contextuais pragmáticos” (CORREIA, 2014, p.63).



Diante desse resultado, a pesquisadora realiza outro experimento com retomadas intrassentenciais<sup>31</sup>, pois estas estariam mais expostas a princípios gramaticais que seriam os princípios que poderiam gerar um custo maior para o processamento de pessoas que gaguejam. Assim, o objetivo desse experimento foi examinar o tempo de leitura da anáfora *a si mesmo(a)* precedida por um antecedente gramatical e um agramatical, seguindo o Princípio A da Teoria da Ligação<sup>32</sup>, cuja explicação a autora baseia em Chomsky (1981, 1986). O experimento não mostrou um efeito principal da variável grupo, mas mostrou um efeito marginalmente significativo entre as variáveis grupo e tipo de sentença. Nesse sentido, os falantes que não gaguejam são mais rápidos na condição gramatical e mais lentos na condição agramatical. As pessoas que gaguejam, por sua vez, apresentam padrão oposto<sup>33</sup>. Essa pesquisa traz indícios, portanto, de que a memória procedimental é requerida para o processamento de anáforas realizadas de modo intrassentencial por conta do envolvimento gramatical necessário nesse contexto, já nas anáforas correferenciais intersentenciais, a demanda de processamento se volta para a memória de trabalho, ou seja, outros fatores atuam cognitivamente para a resolução desse tipo de retomada. Cabe ressaltar que Alves (2012), Vasconcelos (2012) e Correia (2014) trabalham com retomadas feitas na posição de objeto, partindo do conjunto experimental proposto por Leitão (2005).

Diante dos estudos apresentados, é possível dizer, de modo resumido, que há estudos que evidenciam penalidade do nome repetido em português brasileiro, quando comparado ao pronome pleno (ele), e há estudos cujos resultados permitem questionar tal penalidade, evidenciando, inclusive, uma penalidade do pronome pleno com relação ao pronome nulo, sobretudo quando o antecedente está em posição de sujeito. Essas diferenças têm aparecido por conta da metodologia e dos estímulos que os diferentes estudos têm usado. Ademais, há evidências de que anáforas construídas por meio de sintagmas hiperônimos são processadas mais rapidamente que anáforas construídas com sintagmas hipônimos, o que leva a uma discussão sobre a distância semântica entre antecedente e retomada e os custos para a memória

---

<sup>31</sup> Foram utilizadas sentenças como: Maria disse que João machucou a *si mesmo* no parque de diversão (antecedente gramatical); Maria disse que João machucou a *si mesma* no parque de diversão (antecedente agramatical).

<sup>32</sup> O princípio A da Teoria da Ligação compreende que anáforas (que são os reflexivos e recíprocos) devem estar ligadas em seu domínio de ligação (LEITÃO, 2005).

<sup>33</sup> Uma das suposições da pesquisadora é a de que, se os princípios que atuam nas correferências intrassentenciais “estão radicados na memória procedimental e essa memória está enraizada nas estruturas cerebrais que apresentam alterações neurais nas PQG [pessoas que gaguejam], podemos esperar que estas pessoas evidenciem um desempenho atípico no processamento dessa forma de correferência, já que supomos estarem esses princípios de ligação correlacionados com a gramática e a gramática com a memória procedimental.” (CORREIA, 2014, p.31-32, comentário nosso).

de trabalho. Tudo isso mostra que a forma como a anáfora é construída em um texto pode repercutir no custo de processamento, e isso pode refletir na compreensão em leitura, sendo necessário, ao leitor, lançar mão de diferentes estratégias. Há, ainda, fatores de memória envolvidos no processamento desse aspecto, uma vez que, como vimos, a anáfora é um dos recursos responsáveis pelos processos de integração, que resultam na compreensão inferencial. Assim, conforme Franzen e Souza (2020, p. 23),

não basta apenas a retomada de um antecedente para que a compreensão esteja garantida, pois, a depender de como essa retomada é construída, os custos para o processamento podem ser elevados e, conseqüentemente, podem dificultar os processos de compreensão, por terem implicações que remetem à capacidade e ao funcionamento da memória de trabalho. Ao se operar com a integração e a sumarização de um texto, é importante considerar o papel das anáforas correferenciais como pistas textuais que garantem a progressão do texto, mas é importante, também, pensar na forma que elas são realizadas e construídas.

Por conta disso, conhecer os estudos que tratam do processamento anafórico em nível local e trazer, ainda que brevemente, os resultados que têm sido encontrados são ações relevantes para que possamos pensar sobre o papel da anáfora na leitura e, ainda, para entender que o modo como uma anáfora é construída gera diferenças de custos para o processamento. O que queremos investigar é se essa diferença ocorre com o uso do sintagma “o mesmo” em posição de sujeito, retomando antecedente [-animado], contexto no qual o pronome pleno (ele) poderia ser utilizado (nossa hipótese é a de que não). De todo modo, como fica o custo do processamento com o uso de “o mesmo” quando comparado ao de “ele”? Isso reflete de alguma forma nas estratégias utilizadas durante a leitura do texto? Vamos refletir um pouco mais sobre a forma anafórica estudada para que possamos sustentar as hipóteses que guiam este estudo e, em seguida, no método, retomamos essas questões.

#### 2.4 A FORMA ANAFÓRICA ESTUDADA NESTA TESE: “O MESMO”

Como é possível verificar com os estudos apresentados na seção anterior, existem diferentes formas de realizar uma anáfora. Cada uma delas implica custos para o processamento e, também, assume objetivos diferentes no transcorrer de um texto, o que pode refletir na leitura e na compreensão. As anáforas correferenciais, sobre as quais recai o foco do presente estudo, podem ser realizadas por pronomes pessoais, pronomes demonstrativos, pela simples supressão do antecedente (nulo), pela repetição do nome, por sinônimos, por hipônimos e por hiperônimos. Conforme os estudos revelam, corroborando Kintsch e Rawson (2013), as

diferentes formas de realizar uma anáfora correferencial tendem a ter objetivos diferentes e, por isso, apresentam padrões diferentes de processamento. De acordo com os autores, “[o]s pronomes costumam ser usados em referência a conceitos que foram mencionados ou abordados recentemente, que foram introduzidos explicitamente no texto e que estão ativados na memória de trabalho.” (KINTSCH; RAWSON, 2013, p. 233). Além disso, de acordo com Garnham (2001), anáforas realizadas com pronomes pessoais tendem a ser usadas quando não há uma mudança de tópico. Já outras formas de realização da anáfora podem indicar uma mudança no tópico, inserindo um novo referente, como apresentado na seção 2.2, no caso das anáforas indiretas, associativas e nas nominalizações.

De modo geral, os pronomes pessoais são usados, de acordo com Kenedy e Othero (2018), para substituir sintagmas de valor nominal. Um sintagma de valor nominal, como tal, pode vir precedido de um elemento definidor, o que não ocorre com os pronomes. De modo mais específico, os pronomes de terceira pessoa “não têm quase nenhuma significação intrínseca e têm de adquirir seu significado com base em seus antecedentes” (OAKHIL; CAIN; ELBRO, 2017, p. 105). No português, os pronomes são marcados morfologicamente por número e gênero. Durante a leitura, qualquer dificuldade na resolução desses pronomes pode refletir em problemas na compreensão do texto ou de partes dele.

Neste trabalho, investigamos o sintagma “o mesmo”, usado como anáfora correferencial para retomar um item nominal específico, em contextos nos quais se poderia usar o pronome “ele” ou o nulo. Esse elemento estudado, embora esteja sendo usado em contextos nos quais os pronomes poderiam ser mobilizados, não possui características pronominais, o que se mostra verdadeiro ao constatar que ele sempre vem definido pelo artigo (o mesmo/a mesma). Ora, pronomes não permitem o uso de definidores. Isso mostra que “o mesmo” não se caracteriza como pronome, pelo menos não nessa construção. Trata-se de uma nominalização, precedida de um artigo definido. Assim, no nível lexical (em sua formação estrutural), esse elemento se comporta como substantivo. No nível sintático, quando ele assume a função de anáfora correferencial, comporta-se como um pronome, porque o antecedente precisa ser retomado, para o referente ser resgatado. Então, ele não é um pronome, mas tem se comportado como tal. Neste ponto, cabe trazer o que apresenta Pereira (2013, p. 100):

É claro que, diferentemente do pronome *ele* (e suas flexões), o identitivo “o mesmo” não consegue apontar deiticamente para referentes no contexto, por exemplo, apresentando apenas uma propriedade anafórica; ou seja, podemos utilizar o pronome pessoal reto *ele* para indicar um referente ainda não mencionado, mas ainda não fazemos isso com a expressão “o mesmo”. Por conta disso, essa expressão anafórica é mais adequadamente classificada como pronome demonstrativo (tal como é

classificada atualmente). Isso não exclui, porém, a possibilidade de essa forma identitiva estar se gramaticalizando em pronome pessoal de terceira pessoa [...]

Do que cita Pereira (2013), é possível destacar os distanciamentos existentes nas funções exercidas pelo sintagma “o mesmo” e pelo pronome “ele” em um texto. Como a autora ressalta, a função dêitica não é assumida pelo “o mesmo”, que se aproxima do pronome exatamente no que diz respeito ao seu papel anafórico. No que tange às aproximações entre ambos, destaca-se a hipótese levantada por Pereira (2013) sobre uma possível gramaticalização do elemento “o mesmo” em pronome de terceira pessoa, isto é, em alguns contextos, “o mesmo” estaria exercendo a função do pronome “ele”. Ou seja, textualmente, a função que o sintagma “o mesmo” exerce na língua parece estar se modificando. Em sua origem, o sintagma “mesmo” surge na língua portuguesa a partir de duas raízes distintas, uma de reforço, com equivalência de *próprio* e *em pessoa* (*ipse*) e uma de identidade (*idem*), em que necessita da presença de um artigo ou de algum demonstrativo (BECHARA, 2015; PEREIRA, 2013). De acordo com Pereira (2013), o uso de “o mesmo” como anáfora correferencial surge, então, de uma evolução da raiz de identidade. Ao considerarmos a formação lexical do item “o mesmo” e o seu uso como anáfora correferencial em textos diversos, podemos usar uma metáfora para compreender a constituição desse elemento. Ele tem uma roupagem de nome, mas, na constituição sintática, não atua com essa roupagem. É como se a sua roupa fosse um pijama e ele não pudesse usá-lo para ministrar uma aula, por exemplo, mas o faz, mesmo que digam que não é permitido. O questionamento que surge, então, é como ele funciona cognitivamente durante a leitura. É possível dizer que cognitivamente o sintagma “o mesmo” segue o padrão de processamento do pronome “ele”? E como isso reflete na leitura de textos? Esses foram os questionamentos iniciais que orientaram a formulação dos objetivos, das perguntas e das hipóteses desta pesquisa.

O presente estudo desponta, então, a partir do uso de “o mesmo” com a função do pronome “ele”. O foco, portanto, não está em toda e qualquer construção anafórica possibilitada pelo uso de “o mesmo”, mas sim em uma construção específica, que surge a partir de modificações sofridas pelo sintagma “mesmo” e que tem sido observada por alguns estudiosos da língua (MOREIRA, 2007; PEREIRA, 2013; BECHARA, 2015) e considerada inadequada por outros (BAGNO, 2012; SACCONI, 1994; ALMEIDA, 2009). Além disso, no senso comum, e mesmo entre professores e pesquisadores da área da linguagem, é frequente encontrar resistência quanto a esse uso específico.

Sabemos que as gramáticas tradicionais buscam normatizar um padrão de língua a ser utilizado na sociedade. No entanto, compreendemos que se trata apenas de uma convenção, não

refletindo o uso vivo da língua. Ainda assim, o sintagma “o mesmo” não entra na língua portuguesa com a função de anáfora correferencial, embora esteja assumindo essa função na sua evolução. Essa função foi se transformando com o uso da língua e tem sido recorrente, sobretudo, em textos escritos. De modo mais específico, para Bagno (2012), o uso de “o mesmo” ocorre em textos truncados e mal escritos. Ora, se “o mesmo” contribui para gerar textos truncados, seu uso geraria maiores custos para o seu processamento e, também, refletiria no momento da leitura, gerando impactos na compreensão enquanto processo e enquanto produto. Há que se considerar, contudo, que a crítica trazida por Bagno tem mais a ver com questões sociais do que linguísticas propriamente ditas.

Para Bechara (2015, p.175), “alguns estudiosos, por mera escolha pessoal, têm-se insurgido contra o emprego anafórico do demonstrativo *mesmo*, substantivado pelo artigo, precedido ou não de preposição, para referir-se a palavra ou declaração expressa anteriormente”. O gramático ressalta, ainda, que as razões para essa crítica não são apresentadas. Na visão de Bagno (2012, p. 966), por exemplo:

poucos fenômenos de hipercorreção caracterizam tão bem a insegurança linguística e o domínio insuficiente dos gêneros mais monitorados quanto o pronome o mesmo (e flexões). Em 111% dos casos em que aparece, é possível não empregar pronome algum ou empregar o pronome da não-pessoa ele (e flexões).

Então, para o linguista, o uso de “o mesmo” como anáfora correferencial se trata de uma hipercorreção, que representa falta de domínio da linguagem escrita. Isso significa que, embora o uso desse elemento como anafórico seja bastante frequente, mostrando-se recorrente em trabalhos acadêmicos (PEREIRA, 2013; MOREIRA, 2007), é um elemento que não tem, em sua essência, essa função correferencial como a do pronome pessoal, provocando, inclusive, polêmicas no que diz respeito ao seu uso<sup>34</sup>. Mas o que é certo, é que ele ocorre com bastante frequência, inclusive em textos de renomados pesquisadores, como apresentado na introdução. Por conta disso e tendo em vista o que os diferentes estudos apresentados na seção anterior têm levantado sobre o processamento de distintas anáforas, passamos a nos questionar sobre o processamento desse elemento específico, pois, se ele caracteriza, para alguns, uma falta de

---

<sup>34</sup> As insurgências contra esse uso de “o mesmo” são tantas que, em 2010, foi proposta uma alteração na redação de lei que dispunha sobre a fixação de avisos nas portas externas dos elevadores no Rio Grande do Sul. O aviso apresentado na Lei era o seguinte: “Aviso aos usuários: antes de entrar no elevador, certifique-se de que o mesmo encontra-se parado neste andar.” (Lei nº 11.369 de 1999 do Rio Grande do Sul), alterando-se para: “Atenção: antes de entrar, verifique se o elevador está parado neste andar.” (Lei nº 13.413 de 2010 do Rio Grande do Sul). Disponível em: <<http://www.al.rs.gov.br/FileRepository/repLegisComp/Lei%20n%C2%BA%2011.369.pdf>>. Acesso em: 03 abril 2020.

domínio na linguagem escrita, poderia, então, acarretar problemas para o processamento e a compreensão em leitura.

Ainda, sobre anáforas que violam determinados aspectos estabelecidos na língua, Garnham (2001) discorre acerca de um projeto realizado em parceria com Oakhill, em que coletaram *corpus* de anáforas realizadas de forma não esperada, levando em conta diferentes tipos de restrições. Garnham (2001, p. 63, tradução nossa) ressalta que o estudo concluiu que “[...] não há uma relação direta entre prescrições linguísticas sobre como as expressões anafóricas deveriam ser usadas e a forma como elas são interpretadas<sup>35</sup>”. É a partir desse ponto que se desenvolve este estudo, pois a prescrição diz uma coisa, mas a roupagem utilizada pelo “o mesmo” na articulação de textos tem sugerido outra, cabe, então, analisar se, em leitura a roupagem vestida por esse elemento está servindo de modo adequado diante do processamento.

Na próxima seção, para compreender um pouco melhor o sintagma “o mesmo”, são apresentados estudos que buscaram analisá-lo. Tais estudos foram encontrados a partir de buscas realizadas em bases de dados como o Portal de Periódico Capes e a Biblioteca Digital Brasileira de Teses e Dissertações. O resultado dessas buscas mostra que até aquele momento não havia estudos que investigassem o sintagma “o mesmo” com as bases epistemológicas da leitura. Os trabalhos encontrados são de base sociolinguística e eles lançam luz sobre a função anafórica de “o mesmo”.

#### **2.4.1 Os estudos sobre o sintagma “o mesmo”**

Nesta seção, é apresentada uma breve revisão de estudos que têm se debruçado sobre o sintagma “o mesmo”. As discussões feitas por tais estudos e os resultados a que eles chegam ajudam a justificar a relevância de se pensar no processamento desse elemento durante a leitura de um texto acadêmico.

Em estudo realizado por Moreira (2007), busca-se pelos contextos linguísticos que favorecem o emprego de “o mesmo” (e suas variações) como anafórico<sup>36</sup>. O estudo se dá sob a ótica da sociolinguística, tendo por foco a modalidade escrita da língua. De modo mais específico, a pesquisadora parte de textos produzidos por acadêmicos de Administração, ou seja, sua análise se pauta em textos acadêmico-científicos. No trabalho desenvolvido, o uso de “o mesmo” é compreendido como inovador na escrita formal do PB e como uso próprio da

---

<sup>35</sup> “[...] there is no straightforward relation between linguistic prescriptions about how anaphoric expressions should be used and the way they are interpreted”.

<sup>36</sup> O estudo parte da mesma definição de anáfora correferencial assumida nesta investigação.

modalidade escrita, uma vez que a pesquisadora acompanha, também, apresentações orais feitas pelos autores dos trabalhos analisados, constatando que essa forma de anáfora não aparece na fala dos indivíduos que fazem uso dela na escrita.

O estudo considera diferentes fatores de análise para refletir sobre o uso desse elemento como anafórico, a saber: tipo de retomada anafórica ([-posse] e [+posse]), tipo de antecedente ([-animado] e [+animado]), número de retomada do mesmo antecedente (1ª, 2ª ou 3ª retomada), função sintática da retomada (sujeito e complemento nominal), gênero da forma usada (masculino e feminino), número da forma usada (singular, plural). Além disso, para auxiliar na reflexão proposta, é feita, também, uma análise, nos mesmos trabalhos acadêmico-científicos, dos possessivos “dele(s)” e “seu(s)”. Isso porque uma das possibilidades levantadas é a de que o sintagma “o mesmo” estaria sendo usado na escrita para evitar uma possível ambiguidade gerada pelo “seu” e, também, para evitar o uso de “dele”, que seria mais frequente na linguagem oral.

Dentre os fatores analisados, os que se mostraram mais relevantes para o uso do sintagma “o mesmo” foram o antecedente masculino, a função sintática de sujeito e as retomadas mais próximas do antecedente. Sobre esse último aspecto, de acordo com Moreira (2007), quanto mais próxima a retomada está do antecedente, mais se emprega a forma “o mesmo” como anafórico. A pesquisadora compara esses resultados aos obtidos com os dados de “seu (s)” e “dele (s)”, usados para realizar a retomada. Nesse sentido, o ponto mais expressivo foi o fato de que essas formas têm sido empregadas de modo mais usual para retomar antecedentes femininos, reafirmando que o antecedente masculino pode estar favorecendo o uso de “o mesmo”. De acordo com a pesquisadora, esse foi um dos fatores mais relevantes encontrados pela investigação realizada. Além disso, ao concluir o trabalho, a autora destaca que o emprego de “o mesmo” como anáfora correferencial é ainda relativamente novo.

Tal estudo se distancia em muitos pontos desta investigação que estamos desenvolvendo, mas traz reflexões importantes que sustentam o trabalho empreendido. Primeiro porque mostra que o uso do sintagma “o mesmo” é próprio da escrita. Ora, se tal sintagma é empregado na escrita, é importante considerar como se dá o seu processamento durante a leitura e como ele reflete na compreensão de um texto. Além disso, os contextos encontrados por Moreira (2007) como favoráveis ao emprego desse sintagma foram considerados na produção dos materiais utilizados na coleta de dados desta investigação, levando em conta, claro, outros fatores relevantes e que foram mais bem relatados no capítulo em que apresentamos o método da pesquisa.

Outro estudo que investiga o elemento linguístico “mesmo” é o de Amorim (2009), que busca, em sua tese, “compreender e explicar os diversos usos do item *mesmo* e as relações estabelecidas por ele em textos de diferentes sincronias da língua portuguesa, pontuando regularidades e evolução”. (p. 27-28). Para isso, a pesquisadora seleciona dados escritos, a partir de cartas oficiais da Paraíba redigidas nos séculos XVIII e XIX, integrantes do *Corpus Linguístico Diacrônico da Paraíba*, e orais, a partir de entrevistas do Projeto Variação Linguística no Estado da Paraíba (VALPB) realizadas na contemporaneidade. Sua pesquisa parte de uma perspectiva sociolinguística, aborda a questão sob um viés funcionalista e faz um resgate histórico do item “mesmo”, apresentando, também, como ele é caracterizado em diferentes gramáticas e dicionários. A pesquisadora realiza a discussão dos dados a partir das funções de “mesmo” referencial<sup>37</sup> e reforço<sup>38</sup>. No que diz respeito à função anafórica, de acordo com Amorim (2009, p. 108), esse elemento integra “o complexo sistema de meios referenciais da língua como instrumento de coesão textual e de coerência conceitual”. Ao se referir à função anafórica, no entanto, a pesquisadora está tratando de anáfora em seu sentido amplo, considerando construções anafóricas como as apresentadas na seção 2.2, mas não analisando o uso de “o mesmo” como anáfora correferencial especificamente.

Para a análise dos dados, Amorim (2009) cria categorias em que encaixa as diferentes ocorrências de “mesmo”; são elas: reforço enfático ou reflexivo, reforço contextual, referência adnominal, referência nominal, circunstância, inclusão e oposição. Na categoria de referência nominal, a pesquisadora destaca que o item analisado “assume o valor de *substantivo*” (AMORIM, 2009, p. 128), sendo considerado, no estudo, uma nominalização. Seria esperado que, nessa categoria, fosse explorado o uso de “o mesmo” na função anafórica que nos interessa; no entanto, isso não ocorre. A pesquisadora faz apenas uma menção ao uso de “o mesmo” como anáfora correferencial ao citar que essa é a única forma de uso contestada pela gramática quando se trata do “padrão culto da língua”. Nos dados, ao que tudo indica, esse formato não aparece. A investigação realizada por Amorim (2009) traz algumas compreensões acerca do percurso traçado pelo “mesmo” e fornece bases que orientaram o estudo de Pereira (2013), sobre o qual se discorre na sequência.

---

<sup>37</sup> Exemplos retirados de Amorim (2009): “[...] foi me atacar com palavras injuriosas, e depois disse-me que vinha dar um recado, tornando a dizer-lhe que não viesse, tornou a dirigir-me as **mesmas** palavras, e a miaçou-me com hum páo que trazia na mão [...]”; “[...] destinados para as plantações dos Indios, que elas são suficientes para eles, huã vez, que semandem expulsar os pretendidos foreiros que nela seaxaõ sem satisfazerem ao fim para que asarrendavaõ, e que bem Longe de civilizarem os **mesmos** Indios os põem em perturbação, e dezordem por que são da intima e mais abjeta plebe[...].”

<sup>38</sup> Exemplos retirados de Amorim (2009): “Eu só me lembro que morei aqui **mesmo**.”; “Se ele queria tudo tabelado no mercado, ele num num mandava fiscal, num mandava não, ele **mesmo** ia em pessoa [...]”; “A solidão é algo estrutural. É uma coisa que : : de você para com você **mesma**, tá?”



Pereira (2013) busca descrever as funções do item linguístico “mesmo” em textos acadêmico-científicos (trabalhos de conclusão de curso), ou seja, amostras de escrita formal atual. De acordo com a autora, a opção de investigar o fenômeno em trabalhos de conclusão de curso se deu:

[...] por se tratar de um registro da escrita acadêmica, na qual estão presentes algumas características que parecem relevantes ao tratarmos do fenômeno em questão, a saber: uso de linguagem formal, modalização do discurso, uso frequente de mecanismos anafóricos, foco em um tópico discursivo específico, uso de discurso reportado, entre outras particularidades, que permitiriam um vasto uso do item *mesmo* e, conseqüentemente, um significativo número de ocorrências do objeto de estudo. (PEREIRA, 2013, p. 147).

Essa visão corrobora nossa perspectiva acerca do uso anafórico correferencial de “o mesmo” em textos acadêmico-científicos. A investigação de Pereira (2013) está, ainda, pautada em pressupostos funcionalistas, voltando-se para a gramaticalização e para a noção de referenciação, com maior ênfase nos mecanismos de retomada anafórica. A pesquisadora faz uma análise histórica sobre o item analisado, ressaltando a dupla origem vinda do latim: *idem* e *ipse*. A partir da raiz etimológica *idem*, o termo seria usado com artigo ou com outro demonstrativo, já a forma provinda da raiz *ipse* seria usada com substantivo ou pronome (BECHARA, 2015). De modo mais específico, a pesquisadora argumenta, a partir da análise realizada, que a raiz etimológica *idem* deu origem, dentre outras, a macrocategoria definida no estudo como referência/anáfora. Sobre o surgimento do item em si, uma das hipóteses levantadas pela pesquisadora é a de que ele provém desses dois vocábulos, mas passa a se comportar como um item único e polissêmico na língua portuguesa.

O trabalho traz uma discussão bastante pertinente acerca das formas como o elemento “mesmo” é classificado em diferentes gramáticas e dicionários. Percebe-se a versatilidade desse vocábulo e suas particularidades. Dentre todas as funções e classes gramaticais atribuídas ao “mesmo”, interessa-nos aquela que, de acordo com Pereira (2013, p. 151), “tal como um pronome pessoal, [...] anaforiza um sintagma nominal.”. Ou seja, aquela que diz respeito ao “mesmo” precedido de artigo e usado como anáfora correferencial. Essa ocorrência é apresentada, no estudo, como polêmica por ser compreendida por muitos manuais da contemporaneidade como errônea (SACCONI, 1994; ALMEIDA, 2009; BAGNO, 2012). Apesar da “polêmica” em torno desse uso, os resultados apresentados pela pesquisadora mostram que ele é o mais recorrente nos trabalhos analisados. Cabe salientar que sabemos que o estabelecido pelos manuais e pelas gramáticas não representa, necessariamente, o que encontramos e usamos na língua.

Como destacado, Pereira (2013) desenvolve diferentes categorias nas quais o elemento “mesmo” se encaixaria. Essa função que nos interessa de “o mesmo” é categorizada pela pesquisadora como referência nominal, que faz parte da macrocategoria denominada referência/anáfora. Essa macrocategoria é pensada com base na análise feita por Amorim (2009); no entanto Pereira (2013) faz uma reorganização com o propósito de comportar maiores especificidades do item analisado, destinando a categoria referência nominal apenas para casos de “o mesmo” como anáfora correferencial, como citado. Isso significa que o elemento “mesmo” pode assumir outras funções anafóricas, mas são diferentes dessa estudada pela presente pesquisa. Nessa categoria de referência nominal, segundo a autora, “há uma significativa neutralidade de retomada, i.e., *o mesmo* retoma um antecedente sem imprimir-lhe característica ou juízo de valor, motivo pelo qual, inclusive, ele provavelmente esteja sendo usado no discurso acadêmico e jornalístico com uma frequência considerável” (PEREIRA, 2013, p. 184). As características linguísticas do item “mesmo” usado como referência nominal são apresentadas pela pesquisadora da seguinte forma:

- **Tipo de elemento contíguo:** esse mesmo é sempre acompanhado de um artigo definido.
- **Papel sintático de mesmo:** pode exercer diversos papéis sintáticos, como sujeito, objeto direto, objeto indireto e complemento nominal.
- **Flexão de número:** flexiona-se em número, conforme o sintagma nominal ao qual está se referindo.
- **Flexão de gênero:** flexiona-se em gênero, conforme o sintagma nominal ao qual está se referindo.
- **Escopo:** é estreito, recaindo sobre o sintagma nominal ao qual está se referindo.
- **Mobilidade sintática:** apresenta alta mobilidade sintática.
- **Classificação gramatical tradicional:** pronome demonstrativo (polêmico). (PEREIRA, 2013, p. 241).

Então, com a pesquisa realizada pela autora, confirmamos que esse elemento é sempre acompanhado de um definidor e que, tal qual um pronome, ele sofre flexão de gênero e número de acordo com o sintagma nominal ao qual se refere. Quanto à classificação gramatical tradicional, a classificação de pronome demonstrativo é de “mesmo” e não de “o mesmo”. Este quase não é citado nas gramáticas tradicionais que consultamos (ALMEIDA, 2009; CUNHA; CINTRA, 2001; SACCONI, 1994;), a não ser na de Bechara (2015), que propõe uma reflexão sem julgamentos acerca desse uso.

A partir do estudo de Pereira (2013), é possível fazer algumas ponderações importantes para a investigação que estamos desenvolvendo. Primeiro, o estudo ressalta a recorrência de anáforas correferenciais construídas com o sintagma “o mesmo” em textos acadêmico-científicos, o que reafirma a relevância de investigar como esse tipo de anáfora se comporta no

processamento e na compreensão em leitura desse tipo de texto. Além disso, a partir do que Pereira (2013) discute, é possível conhecer um pouco mais sobre as especificidades dessa anáfora, reforçando o seu uso em contextos semelhantes aos quais o pronome “ele” é utilizado, ainda que, historicamente e gramaticalmente, tal elemento não tenha, na língua portuguesa, a mesma função do pronome pessoal.

Os estudos que investigam o elemento “mesmo” trazem diversas classificações para esse elemento linguístico. Dentre as classes propostas pela Nomenclatura Gramatical Brasileira, ele pode ser considerado adjetivo (A mesma testemunha foi chamada ao tribunal), substantivo (Nas aulas seguinte, ocorreu o mesmo), conjunção (Mesmo doente, continuou ajudando os outros), advérbio (Hoje mesmo falei com ela), pronome demonstrativo (Ela sempre morou na mesma casa). Isto é, ele assumiria diferentes funções na língua, nenhuma dessas classificações tradicionais, no entanto, dão conta da ocorrência frequente de “o mesmo” como anáfora correferencial empregada nos mesmos termos do pronome “ele”. Embora alguns teóricos contemporâneos tratem esse uso como demonstrativo polêmico, entende-se que, na prática, tal elemento não está exercendo o papel de demonstrativo nesse contexto, ou seja, é um elemento que tem se transformado, principalmente em contextos de escrita. Por isso, é justificado o estudo em torno do seu processamento no contexto específico investigado nesta tese e seu reflexo na compreensão em leitura.

A busca por pesquisas que investigam o elemento “mesmo” revelou estudos que têm buscado compreender e relacionar sua etimologia com as funções que ele assume na língua. O que se percebe pelos resultados que os estudos evidenciam é que há diversos usos possíveis. Um bastante comum é do elemento “mesmo” substantivado, usado como anáfora correferencial. Ou seja, tal elemento se mostra presente em textos escritos (foco de investigação dos estudos apresentados nesta seção). No entanto, não foram encontrados estudos que analisassem a leitura de tal elemento na função de anáfora correferencial.

A partir de todo o exposto até o momento, tivemos por intenção esclarecer o ponto do qual partimos e os estudos que subsidiam a pesquisa que ora realizamos. Com isso em mente, no próximo capítulo, detalhamos nossas hipóteses acerca da relação entre compreensão em leitura e o processamento do sintagma “o mesmo” enquanto anáfora correferencial. Na sequência, o método da pesquisa é detalhado.

### 3 MÉTODO

Neste capítulo, é apresentado o desenho do estudo, explicando ao leitor os passos, critérios, procedimentos, condutas e decisões para a construção da investigação. Para a realização da pesquisa, obteve-se, em 20 de dezembro de 2018, aprovação do Comitê de Ética em Pesquisa com Seres Humanos da Universidade Federal de Santa Catarina, sob o parecer n.º 3.097.074. A pesquisa está registrada no Certificado de Apresentação para Apreciação Ética (CAAE) sob o número 02803818.4.0000.0121. Inicialmente, cabe destacar que, no que tange ao delineamento da pesquisa, trata-se de um estudo cujo método é experimental com abordagem explicativa. O caráter do estudo é definido como experimental, porque, após ser determinado o objeto do estudo, foram selecionadas as variáveis sob investigação, definidos critérios de controle e observado o seu efeito no objeto investigado (GIL, 2007; KAISER, 2013; ABBUHL; GASS; MACKEY, 2013). Temos, neste estudo, duas condições sendo testadas (o pronome “ele” e o sintagma “o mesmo”) e buscamos explicar como esse tipo de anáfora, que age nas relações interfrasais, formando a microestrutura de um texto, atua na formação da base textual e reflete na compreensão desse texto. Para tanto, observamos os efeitos nas variáveis dependentes: acurácia, tempo de leitura, comportamento leitor, desempenho geral em teste de compreensão, desempenho local em teste de compreensão, aceitabilidade e estratégias empregadas. Além disso, trata-se, também, de um estudo exploratório, pois algumas das nossas perguntas buscam esclarecer questões específicas e proporcionar uma visão geral (GIL, 2007) acerca do comportamento leitor e das estratégias empregadas diante das anáforas correferenciais realizadas com o sintagma “o mesmo”. Assim, esses pontos específicos do estudo requerem exploração antes de se aventarem hipóteses.

A abordagem é explicativa. A índole da ciência Psicolinguística é essencialmente explicativa (PINTO, 2019). Buscamos por uma explicação acerca do fenômeno analisado (GIL, 2007), conhecendo e compreendendo, de modo mais específico, a influência e o papel da construção anafórica feita com “o mesmo” na compreensão em leitura. A abordagem dada ao problema é considerada mista (FLICK, 2009; CRESWELL; CLARK, 2013), envolvendo, nas discussões, elementos das abordagens quantitativas e qualitativas. Pesquisas experimentais tendem a realizar análises quantitativas (GRAY, 2012), isto é, análises que buscam checar hipóteses desenvolvidas *a priori*. Essas análises são feitas a partir de rodadas estatísticas, que revelam as características do estudo e possibilitam inferências acerca do grupo estudado. No entanto, neste estudo, a análise qualitativa é necessária, pois é com ela que atribuiremos sentido aos dados, desmembrando-os em partes menores (GRAY, 2012) para poder analisá-los de modo

mais detido. Além disso, o N amostral é pequeno, e isso requer a contemplação dos dados, de forma interdependente ao que apresentam os números.

Dessa forma, a abordagem mista é concretizada a partir da integração das diferentes técnicas utilizadas para a coleta de dados. A abordagem mista é adotada no ponto de vista da análise, em que os dados são triangulados, considerando as perspectivas quantitativa e qualitativa de modo complementar. Esse processo é compreendido, de acordo com Flick (2009, p. 43), “como a compensação complementar das deficiências e dos pontos obscuros de cada método isolado”. Assim, temos dados que permitem um olhar quantitativo e qualitativo e o que fazemos é a integração desses dados, “combinando-os (ou misturando-os) de modo sequencial, fazendo um construir o outro ou incorporando um no outro” (CRESWELL; CLARK, 2013, p.22).

Toda essa caracterização fica mais clara à medida que o design da pesquisa é explicado. Para tanto, na sequência, são retomados os objetivos, as perguntas e as hipóteses que guiam o estudo. Em seguida, são apresentados os participantes da pesquisa e a justificativa para a seleção desse público. Então, são fundamentadas e explicadas cada uma das técnicas que compõem a coleta de dados da investigação, a saber: protocolo verbal, leitura automonitorada, teste de aceitabilidade e teste de compreensão em leitura com gravação de tela. Na sequência, o texto utilizado na coleta de dados é apresentado e, nas duas seções seguintes, trata-se, respectivamente, dos procedimentos de coleta de dados e dos critérios e procedimentos de análise. Por fim, discorre-se sobre o estudo piloto realizado.

### 3.1 OBJETIVOS, PERGUNTAS E HIPÓTESES

Como já apresentado na introdução, o grande fio condutor desta pesquisa é a leitura, voltando-se para o público do ensino superior. De modo mais específico, foca-se no papel das anáforas correferenciais na compreensão em leitura, afunilando o olhar para uma retomada específica, realizada com o sintagma “o mesmo”. Por conta disso, o grande **objetivo** desta investigação está em compreender de que forma e em que medida “o mesmo” usado como anáfora correferencial influencia no processamento e na compreensão em leitura de texto acadêmico-científico. Esse objetivo é complementado pelos seguintes **objetivos específicos**:

1. Conhecer o custo de tempo de resposta diante do sintagma “o mesmo” como anáfora correferencial.

2. Explorar o tempo e a acurácia diante de anáforas correferenciais de forma a relacioná-los a estratégias utilizadas e ao desempenho em leitura de texto acadêmico-científico.
3. Identificar a aceitabilidade do uso de “o mesmo” como elemento anafórico correferencial em construções sentenciais do PB.
4. Explorar o desempenho e o comportamento do leitor diante do texto e do sintagma “o mesmo”.
5. Caracterizar e compreender as estratégias utilizadas durante a leitura de texto acadêmico-científico, no que tange à anáfora, e a relação de tais estratégias com a resolução dessa anáfora.

Embora o uso de anáforas correferenciais auxilie na construção da integração de diferentes partes do texto, sendo recurso importante para a construção de uma representação mental adequada, a compreensão em leitura não é assegurada apenas pelo fato de o texto possuir anáforas correferenciais. Isso porque, para além dos recursos de retomada, há que se pensar em outros fatores, tais como: informações sintáticas, semânticas, lexicais, conhecimentos prévios do leitor dentre tantos outros aspectos relativos ao texto, ao leitor e à situação de leitura. No que diz respeito à anáfora correferencial, há que se considerar, ainda, a forma como a retomada é realizada, a construção sintática de antecedente e anáfora, a distância entre um e outro (LEITÃO, 2005; QUEIROZ, 2009; MAIA, 2013; TEIXEIRA, 2013, dentre outros). Todos esses fatores influenciam no processamento das anáforas, implicam maiores ou menores custos para o processamento e, por sua vez, podem afetar a compreensão. O ponto deste trabalho está em compreender se “o mesmo”, usado como anáfora correferencial, pode ser uma forma de retomada que influencia a compreensão em leitura, dentro de um contexto em que posição sintática e distância entre antecedente e anáfora estão controladas. Para que seja possível analisar esse sintagma, faz-se a comparação com o pronome pessoal “ele”, já que “o mesmo” tem sido usado em contextos nos quais o pronome poderia ser usado (MOREIRA, 2007; PEREIRA, 2013).

Assim, o grande objetivo gira em torno da anáfora correferencial realizada com o sintagma “o mesmo” e, para controlar e contrastar os resultados, o pronome “ele” entra em cena. Busca-se, portanto, **responder**: de que forma e em que medida o sintagma “o mesmo”, como anáfora correferencial, atua na integração local, na formação da base textual e na compreensão em leitura de texto acadêmico-científico? A hipótese geral do estudo prevê que: “o mesmo” segue um padrão de processamento semelhante ao do pronome “ele”. Sendo assim, espera-se que os resultados em termos de processamento e de produto mostrem que “o mesmo”

pode exercer a função de anáfora correferencial sem que isso gere entraves para a compreensão em leitura. Então, ainda que o sintagma “o mesmo” não se caracterize como pronome, ele pode assumir tal função e atuar como anáfora correferencial, sem que isso gere complicações ou atrasos no processamento e na compreensão.

Como discurremos na seção anterior, a compreensão é um processo multicomponencial e complexo. Além disso, compreensão também é produto. Por conta dessa característica complexa da compreensão, a falha ou o sucesso no processamento em um nível (no caso desta pesquisa, o nível interfrasal) pode não impactar diretamente na falha ou no sucesso em outro nível (compreensão textual, no nosso caso), mas, tanto a falha quanto o sucesso em um nível aparecerá de alguma forma no decorrer da leitura (no tempo de processamento, na aceitabilidade das sentenças, nas estratégias utilizadas, no comportamento diante do texto, nas respostas a perguntas de compreensão). Então, são esses aspectos que nos permitirão analisar a relação entre esses diferentes níveis (textuais e de processamento) considerados nesta tese.

Diante disso, busca-se responder a algumas perguntas e averiguar as hipóteses levantadas a partir dessas perguntas:

- Pergunta 1: O uso de “o mesmo” como anáfora correferencial implica maior custo no tempo de resposta?
- Hipótese a: Ao considerar que o pronome “ele” é o elemento anafórico preferencial em contextos específicos (LEITÃO, 2005; QUEIROZ, 2009; GADELHA, 2012; KINTSCH; RAWSON, 2013; GONDIM, 2017) e tendo em vista o que estudos têm mostrado acerca do uso de “o mesmo” em textos acadêmico-científicos (PEREIRA, 2013; MOREIRA, 2007), entende-se que os custos de tempo de resposta não são afetados pelo uso de “o mesmo” como anáfora correferencial. Ou seja, o custo de tempo de resposta de “o mesmo” é semelhante ao custo de tempo de resposta do pronome ele.

A tarefa de leitura automonitorada dará subsídios para responder à pergunta 1 e averiguar a hipótese “a”. É esperado que a semelhança no custo do tempo de resposta seja evidenciada a partir dos tempos de leitura do segmento pós-anafórico, tendo em vista que o sintagma “o mesmo” pode atuar como outros tipos de anáfora, que não a correferencial, e isso se resolve no momento da leitura do verbo, quando, na leitura automonitorada (que foi realizada a partir da segmentação em sintagmas), ocorre a integração sintática. Além disso, o tempo total de leitura das frases da leitura automonitorada também será considerado, a fim de que se possa analisar o tempo da leitura dos participantes em sentenças com o “ele” e com “o mesmo”.

Optamos pelo trabalho com dois tempos distintos (segmento pós-anafórico e tempo total) para garantir resultados fidedignos e que não fossem apenas resultados obtidos apenas por conta dos tempos escolhidos para análise, pois, como aponta Gondim (2017), as escolhas metodológicas, dentre elas o tempo considerado para a análise, pode dar margem para a obtenção de resultados distintos.

- Pergunta 2: Considerando que o desempenho local do participante é semelhante ao desempenho geral, qual a relação entre desempenho local e tempo de resposta na leitura de “o mesmo” e de “ele” como anáforas correferenciais? Há diferenças na relação entre tempo de resposta de “o mesmo” e desempenho local e tempo de resposta de “ele” e desempenho local?
- Hipótese b: ao considerar o papel das anáforas correferenciais na integração entre sentenças, ao assumir que, nesse contexto sentencial, o uso do pronome “ele” tem apresentado custos menores para o processamento em atividade de leitura (LEITÃO, 2005; QUEIROZ, 2009; NÓBREGA-LIMA, 2014), ao considerar a aproximação no que diz respeito ao uso de “o mesmo” e “ele” em contextos de escrita (PEREIRA, 2013; MOREIRA, 2007) e, ainda, ao levar em conta a função das anáforas correferenciais para a compreensão em leitura (OAKHILL; CAIN; ELBRO, 2017; KINTSCH; RAWSON, 2013; KINTSCH, 1998), hipotetiza-se que o tempo de resposta na leitura de “o mesmo” afeta positivamente o desempenho local, assim como ocorre com o tempo de resposta na leitura de “ele”. Isto é, quanto menor o tempo de resposta na leitura desses dois anafóricos, melhor o desempenho em leitura local do texto.

A pergunta dois será respondida e a hipótese “b” averiguada com a triangulação dos resultados provindos do teste de compreensão em leitura e do teste de leitura automonitorada. O teste de compreensão fornece os dados de desempenho em leitura a partir das questões que requerem a correta integração entre partes menores do texto (integração local), já com o teste de leitura automonitorada conseguimos ter acesso ao tempo que o participante levou para ler as frases com as anáforas-alvo.

- Pergunta 3: Qual a aceitabilidade do uso de “o mesmo” como elemento anafórico correferencial em construções sentenciais do PB?



- Hipótese c: A partir do que se tem apresentado como resultado de estudos que investigam as funções do sintagma “o mesmo” em textos acadêmico-científicos (PEREIRA, 2013; MOREIRA, 2007) e a partir das corroborações feitas pelos estudos que investigam o processamento do pronome “ele” (LEITÃO, 2005; QUEIROZ, 2009; NÓBREGA-LIMA, 2014), considera-se que a aceitabilidade do uso de “o mesmo” não é diferente da aceitabilidade do uso de “ele” como anafóricos.

Para responder à pergunta 3 e testar a hipótese “c”, recorre-se aos dados obtidos com o teste de aceitabilidade, considerando a média de aceitabilidade alcançada nas diferentes sentenças realizadas com “o mesmo” e com o pronome. A partir do julgamento de aceitabilidade, os participantes revelam se, conscientemente, avaliam de modo semelhante ou não as frases construídas com “o mesmo” e com o pronome (ainda que não saibam que o foco da investigação está nesses elementos).

- Pergunta 4: Qual a relação entre o comportamento geral, o comportamento diante da anáfora correferencial construída com o sintagma “o mesmo” e o desempenho diante das retomadas anafóricas e do texto?

A questão quatro será respondida, de forma exploratória – por conta disso, não é apresentada hipótese (GRAY, 2012) – por meio dos dados obtidos com o teste de compreensão em leitura, tanto no que diz respeito ao desempenho geral e local quanto ao comportamento, obtido por meio das categorias advindas do mapeamento de tela. O comportamento dos participantes durante a realização do Teste de Compreensão Leitora fornece indícios acerca dos possíveis impasses durante a leitura e em quais pontos esses impasses ocorreram. O Teste de Compreensão Leitora, então, é mobilizado para responder a essa pergunta porque, com ele, conseguimos analisar o desempenho dos participantes na compreensão do texto e na integração local possibilitada pelas anáforas e, ainda, o comportamento estratégico adotado para conseguir responder ao solicitado.

- Pergunta 5: Que estratégias de leitura são utilizadas pelo leitor e qual a relação de tais estratégias com o desempenho diante das retomadas anafóricas e do texto? Ainda, é possível, no que diz respeito às estratégias empregadas pelos participantes, observar semelhanças frente às diferentes anáforas lidas?

A questão 5 também será respondida de forma exploratória – por isso, não se apresenta hipótese (GRAY, 2012) – a partir dos dados obtidos por meio das tarefas de protocolo verbal e do desempenho no teste de compreensão em leitura. Os protocolos verbais proveem dados das estratégias utilizadas pelos leitores e a escolha dessas estratégias revelam os passos dados pelos participantes durante a leitura do texto e, também, quais estratégias foram utilizadas em que pontos específicos. A partir da análise inferencial, é possível refletir acerca da estratégia usada e se ela revela uma dificuldade na compreensão ou não.

Com a apresentação, na seção 3.3, de cada um dos instrumentos utilizados nesta investigação, essas relações ficarão ainda mais claras.

### 3.2 PARTICIPANTES

A pesquisa foi conduzida com vinte e seis acadêmicos regularmente matriculados no curso de Engenharia de Controle e Automação da Universidade Federal de Santa Catarina (UFSC), na cidade de Florianópolis (SC). Todos os participantes participaram dos dois momentos de coleta de dados. A idade dos participantes varia de 20 a 26 anos. A maioria (88,46%) dos participantes é do sexo masculino, confirmando uma tendência dos cursos de engenharia, que, historicamente, são compostos predominantemente por acadêmicos do sexo masculino<sup>39</sup>. Assim, 23 acadêmicos e 3 acadêmicas participaram da coleta de dados. Cabe explicar por que engenharia e por que esse curso específico. Como esta pesquisa tem por foco a leitura de texto acadêmico-científico, o mais apropriado é que o trabalho seja desenvolvido com uma área específica, para delimitar um único texto de área (dessa forma, tem-se um controle experimental maior). Isso significa que, tendo em vista os objetivos da pesquisa, selecionar estudantes do ensino superior de cursos aleatórios seria contraproducente.

Diante disso, a delimitação feita considerou a área da engenharia por alguns fatores. Primeiro porque é uma área que, tradicionalmente, não apresenta disciplinas e tópicos de ementas que se voltem para questões de leitura e escrita, embora os acadêmicos precisem produzir trabalhos escritos para que possam realizar e concluir o seu curso, o que implica,

---

<sup>39</sup> Esse dado mostra uma tradição dos cursos de engenharia, que são vistos pela sociedade como predominantemente masculinos. A diferença gigantesca entre homens e mulheres na pesquisa, reflete a diferença de público masculino e feminino nos cursos de engenharia em si. São reflexos de uma sociedade que incentiva meninos, desde muito pequenos, a descobrirem e a inventarem, enquanto as meninas tendem a ser preparadas para tarefas domésticas. Ainda que isso seja estrutural e, quando se pensa na população de modo amplo, possa não atingir o nível da consciência (já que permeia todos os espaços e tempos), acaba refletindo na tendência de cursos tecnológicos receberem de forma massiva estudantes do sexo masculino. Há alguns anos, essa diferença tem diminuído, mas é perceptível que ainda há um caminho longo para ser trilhado rumo a uma sociedade efetivamente democrática, inclusiva e que respeite os direitos de todas e todos.

também, a leitura de textos acadêmico-científicos. Ademais, pesquisas (SCHLICHTING, 2021; SCHLICHTING, 2016; FISCHER; HEINIG, 2014; FRANZEN, 2012) têm mostrado a relevância de se refletir sobre a leitura e a escrita nessa área, tanto no período de formação (graduação) quanto no decorrer da vida profissional. Por conta disso, direciona-se o foco do presente estudo para a área da engenharia, popularmente conhecida pelo envolvimento em cálculos com ênfase nas demandas tecnológicas, mas que também apresenta necessidades voltadas para a leitura e a escrita. Assim, para além de todas as questões relativas ao processamento anafórico em si, tem-se a oportunidade de desenvolver um trabalho que envolva a leitura nessa área específica, podendo dar prosseguimento, em alguma medida, ao trabalho iniciado em Franzen (2012).

Com a área delimitada, entrou-se em contato com o diretor do Centro Tecnológico (CTC) da UFSC com o intuito de obter autorização para o desenvolvimento da pesquisa e, também, com o objetivo de delimitar o curso de engenharia com o qual se trabalharia. Afinal, embora sejam todos da grande área da engenharia, cada um deles possui suas delimitações e especificidades, que são relevantes ao se considerar o desenho experimental. Esse primeiro contato ocorreu em setembro de 2018. O diretor prontamente atendeu à solicitação, autorizou a realização da pesquisa e passou o contato dos coordenadores dos diferentes cursos alocados no Centro Tecnológico. Fez-se contato com esses coordenadores, e a primeira resposta obtida foi do curso de Engenharia de Controle e Automação. Como era necessário trabalhar com um curso específico já que a pesquisa trata de leitura de texto acadêmico-científico, que exige conhecimentos prévios específicos e de área, após conversa com professores responsáveis por disciplinas de estágio e de trabalho de conclusão de curso, delimitou-se o perfil dos participantes da pesquisa: acadêmicos de Engenharia de Controle e Automação.

Ainda, como a investigação trata de leitura de texto acadêmico-científico, que aborda conteúdo da área dos estudantes, o recrutamento de participantes considerou estudantes que já estivessem ao menos na metade do curso e que, ao mesmo tempo, não estivessem em fases finais. A decisão de não trabalhar com acadêmicos das fases finais se deu por sugestão dos professores do curso, porque, quando estão nos últimos três semestres, os estudantes de Engenharia de Controle e Automação já não costumam estar na Universidade com tanta frequência, visto que realizam apenas disciplinas optativas e estágios em empresas localizadas em diferentes cidades e países. Isso tornaria inviável a participação deles na pesquisa.

O referido curso desenvolve-se em dez fases, que equivalem a cinco anos. Nesse sentido, os acadêmicos recrutados para a pesquisa deviam estar cursando disciplinas do quinto, do sexto ou do sétimo período do curso. Vale a ressalva para o fato de que os acadêmicos não seguem

necessariamente a ordem estipulada pela grade curricular, cursando disciplinas de diferentes fases do curso. Todos os participantes, contudo, já estavam há, pelo menos, dois anos e meio no curso. Isso significa que já possuíam algum conhecimento de área, o que é relevante, já que leriam um texto acadêmico-científico de uma área de especialidade. O conhecimento prévio dos participantes acerca da temática tratada no texto lido não foi testado. Primeiro porque a seleção dos participantes já partiu de um ponto de controle: não ter acadêmicos novatos. Depois, porque o próprio teste de compreensão leitora nos mostraria se a compreensão do texto tivesse sido inviabilizada por conta do conteúdo em si, o que não ocorreu (como se observará no próximo capítulo, quando da análise dos dados).

O convite aos participantes se deu por meio de passagem nas salas de aula em agosto de 2019. Tal abordagem foi sugerida pelos professores do curso. Nesse momento, explicou-se a pesquisa e solicitou-se o preenchimento de um questionário, caso tivessem interesse em participar do estudo. Após essa etapa, entrou-se em contato por e-mail com todos aqueles que aceitaram o convite. Vinte e seis acadêmicos retornaram o e-mail e se dispuseram a participar de todas as etapas da coleta de dados. Dessa forma, o recrutamento dos participantes foi por convite e a participação se deu de modo voluntário.

Como retribuição pela participação no estudo, ofereceu-se uma sequência de oficinas de escrita acadêmica (seção 3.2.1). Além disso, todos que participaram receberam uma certificação que pode ser validada como AC (atividades complementares), necessárias para a integralização do curso de graduação.

A seguir, apresentamos o quadro com a caracterização dos participantes da pesquisa:

Quadro 1: Caracterização dos Participantes da Pesquisa

<b>Participante</b>	<b>Costuma ler para a graduação?</b>	<b>Quais materiais?</b>	<b>Tem dificuldades com a leitura? Quais?</b>	<b>Quantos semestres faltam para concluir o curso de graduação?</b>
P1	Sim	Artigos técnicos da área	Não	4 semestres
P2	Sim	Apostila e livros das disciplinas	Não	2 semestres <sup>40</sup>
P3	Sim	Livros, publicações, relatórios técnicos	Não	3 semestres
P4	Sim	Textos teóricos	Sim. Cita problemas em interpretar questões nas avaliações	7 semestres

<sup>40</sup> Em alguns casos, o número de semestres para conclusão do curso pode ser maior ou menor do que aquele estipulado na seleção dos participantes. Isso porque os acadêmicos não realizam o curso de modo linear. Então, ainda que o número de semestres informados possa ser distinto, todos os acadêmicos cursavam disciplinas da metade do curso e já não eram iniciantes.

P5	Sim	Textos técnicos e material didático	Não	3 semestres
P6	Sim	Livros teóricos	Não	4 semestres
P7	Sim	Textos técnicos ou livros das bibliografias das disciplinas	Com vocábulos desconhecidos	5 semestres
P8	Sim	Artigos, livros, relatórios	Não	4 semestres
P9	Sim	Livros	Tem dificuldade com textos acadêmicos	4 semestres
P10	Sim	Relatórios e, para disciplinas teóricas, há uma carga maior de leitura	Não	2 semestres
P11	Sim	Textos e livros técnicos	Não	4 semestres
P12	Sim	Livros específicos das disciplinas	Não	3 semestres
P13	Sim	Leitura de artigos para referenciar em relatórios	Não	4 semestres
P14	Sim	Livros e artigos	Não	1 semestre
P15	Sim	Artigos, livros e materiais técnicos	Sim, problemas de compreensão	5 semestres
P16	Sim	Livros e sites	Não	1 semestre
P17	Pouca leitura	-	Não	3 semestres
P18	Sim	Livros	Não	5 semestres
P19	-	Somente lê o livro quando não entende a aula	Sim. Compreender textos acadêmicos e manter o foco	4 semestres
P20	Sim	Livros das disciplinas	Não	5 semestres
P21	Sim	Livros de algumas disciplinas	Não	4 semestres
P22	Sim	Material acadêmico	Não	4 semestres
P23	Sim	Material base	Manter o foco (diz entrar em um modo automático, principalmente quando o assunto não desperta interesse)	2 semestres
P24	Sim	-	Não	3 semestres
P25	Sim	Livros das disciplinas	Não	4 semestres
P26	Sim, em algumas disciplinas	Manual	Não	6 semestres

Fonte: a autora.

Com o quadro apresentado, é possível visualizar que todos os participantes afirmam terem de ler para o seu curso de graduação, ou seja, a leitura é parte da sua vida acadêmica. Os materiais mais lidos são os artigos e livros indicados pelos professores das disciplinas. Isso significa que a leitura de textos acadêmicos-científicos está presente na realidade desses acadêmicos e é relevante, pelas respostas dadas, para a realização das disciplinas curriculares do curso. Além disso, com os dados relativos ao tempo de curso que resta aos participantes,

conseguimos inferir que todos já tem alguma experiência em seu curso, o que é relevante para um estudo que trabalha com texto e terminologia de área (SCHLICHTING, 2021).

### **3.2.1 Contrapartida aos participantes: oficinas de escrita acadêmica**

Ao se realizar uma investigação, o pesquisador pode buscar propor temáticas que contribuam de modo teórico e prático com a sua área de estudo e com a sociedade. Quando os dados analisados são obtidos por meio de pessoas, pressupõe-se a participação de cidadãos que cedem o seu tempo para a ciência, mas nem sempre são diretamente beneficiadas por aquele estudo específico. Como abordagem comum, o participante oferece seus dados e, após isso, a pesquisa segue sem a sua participação. Isso se dá por diferentes motivos que podem envolver tempo e recursos materiais.

Desde o início da investigação, tínhamos por intuito propor algo diferente, oferecendo aos participantes uma retribuição pelo tempo e pelos dados cedidos à pesquisa. Como o público participante foi composto por acadêmicos que ainda não tinham realizado o seu trabalho de conclusão de curso e como o currículo do curso desses estudantes não possui disciplinas que se voltem para a escrita acadêmica, decidimos, após consultar o interesse dos participantes, pela oferta de um ciclo de oficinas com o objetivo de auxiliá-los no processo de escrita acadêmica.

Essa contrapartida foi, inicialmente, pensada para ser presencial e para ocorrer no primeiro semestre de 2020 (abril e maio). No entanto, devido à pandemia mundial do Covid-19, ela ocorreu de modo remoto nos meses de outubro e novembro de 2020. As oficinas aconteceram no decorrer de quatro semanas. Em cada semana, eram liberados aos acadêmicos vídeos explicativos, que tinham em torno de 20 minutos, e atividades a serem feitas. Além disso, foram realizados quatro encontros síncronos de uma hora cada. Nesses encontros síncronos, o conteúdo estudado na semana era revisado, as atividades propostas eram corrigidas e dúvidas eram sanadas.

A formulação das oficinas partiu do modelo de competência em escrita, proposto por Gagné, Yekovich, Yekovich (1993), que tem por base estudos cognitivos. Assim, as oficinas foram planejadas tendo em vista os seguintes processos cognitivos: planejamento, “tradução” e revisão. Ou seja, a compreensão de escrita como processo fundamenta as oficinas (PINTO, 2019). Cabe a ressalva para o fato de que, como se trata de oficinas, o objetivo foi mesclar momentos de explicação e momentos de atividades orientadas.

Foram, então, planejadas quatro semanas, que chamamos de unidades. A primeira semana foi introduzida com um encontro síncrono, em que tratamos da organização das

oficinas, da organização do conteúdo e introduzimos a discussão acerca do planejamento da escrita, trabalhando o papel da leitura nesse processo, porque entendemos que, conforme Pinto (2018, p.11):

A escrita apoia-se, com efeito, na leitura e ela será tanto mais enriquecida por esta quanto mais informação for necessária para sustentar o que se vai redigindo. A necessidade de informação decorrente de uma investigação rigorosa ainda é mais compreensível quando se trata da escrita acadêmica [...]

Para que se fosse possível oferecer subsídios que auxiliassem no processo de leitura dos acadêmicos, foram trabalhadas algumas estratégias, voltando-se, de modo mais aprofundado, para a aquisição lexical e para a correferenciação. As estratégias, como sugere Tomitch (2012, p. 32, tradução nossa), podem ser compreendidas de modo simples como “um meio para um fim<sup>41</sup>”. Dessa forma, buscou-se oferecer essa reflexão sobre estratégias a serem usadas, pensando nas retomadas, que são importantes recursos para a progressão textual, e no léxico que compõe o texto, sobretudo, o vocabulário específico (já que se trata de leitura de área com a finalidade de uma produção escrita acadêmica). Ainda, no que diz respeito à unidade do planejamento, voltou-se para a relevância de se estabelecerem objetivos claros para o estudo a ser empreendido. Dentre as atividades dessa primeira semana, estava a solicitação para que os acadêmicos pesquisassem por um artigo científico que abordasse uma temática de interesse (de preferência, voltada para a temática que pretendiam abordar em seu projeto final de curso). Esse artigo serviria de base para outras atividades a serem realizadas.

A segunda unidade das oficinas se caracterizou como uma etapa de transição entre o planejamento e a escrita. Finalizou-se a discussão sobre as estratégias de leitura e sobre o planejamento de modo geral, retomando a reflexão sobre correferência e refletindo sobre o seu papel durante o processo de escrita. Nessa segunda semana, o texto escolhido pelos acadêmicos foi retomado em uma atividade de sumarização. A sumarização está ligada ao processo de estabelecimento do tema do texto lido, de identificação da ideia principal e de detalhes secundários (SOLÉ, 1998), é considerada um processo para a formação da macroestrutura textual (KINTSCH; RAWSON, 2013; SOLÉ, 1998), que “proporciona uma ideia global do significado do texto em um nível superior” (SOLÉ, 1998, p. 144). A sumarização reflete essa macroestrutura e mostra se o tópico central do texto foi compreendido pelo leitor, por isso fechamos o trabalho com a leitura tratando desse tema e solicitando a produção de uma sumarização.

---

<sup>41</sup> A very simple and handy definition is that of a strategy being ‘a means to an end’ [...]

Na terceira unidade, o foco foi o que Gagné, Yekovich e Yekovich (1993) chamam de “tradução”. O objetivo se voltou a uma reflexão sobre o que é um texto, de modo geral, e o que é e o que se espera de um texto acadêmico em termos linguísticos e científicos. A discussão foi proposta a partir dos processos correferenciais (estudo iniciado quando se trabalhou a leitura), discutindo a coesão e a coerência e analisando aquilo que é comum a diferentes textos acadêmicos. Além disso, foram discutidas as etapas de um texto acadêmico, propondo análises e reflexões sobre o que se espera de uma introdução. Assim, os acadêmicos receberam introduções de trabalhos realizados na sua área da Engenharia para que pudessem refletir sobre as construções efetuadas. A intenção era a de que, até esse ponto das oficinas, os acadêmicos já tivessem subsídios para iniciar um esboço de introdução para o seu trabalho de fim de curso. Essa produção foi orientada e retomada na última unidade. Além disso, para fechar a terceira unidade, foi abordado o processo de revisão (GAGNÉ, YEKOVICH, YEKOVICH, 1993). Importa dizer que a revisão foi posicionada ao final das oficinas por conta da sistematização do ensino. Ela foi, no entanto, trabalhada como constituinte de todo o processo de escrita (PINTO, 2018). Como atividade para sistematizar o que foi abordado em torno da “revisão”, os acadêmicos tiveram que avaliar uma introdução disponibilizada a eles a partir de tópicos-guia, olhando para essa introdução com olhar de revisor.

Já a quarta semana foi pensada para a construção de um esboço de introdução. Esse esboço deveria ser construído, pelos participantes, em um documento do Google para que pudesse ser acompanhado. Por fim, o encontro síncrono final foi de retomada, propondo o esclarecimento de dúvidas, revisitando os pontos trabalhados ao longo dos encontros e discutindo questões específicas da escrita de uma introdução. Não se esperava que os acadêmicos conseguissem desenvolver a introdução completa de seu trabalho, mas que conseguissem refletir sobre possíveis objetivos e formular um texto coerente em torno da temática pretendida.

Sabe-se que, durante o desenvolvimento da escrita proficiente, todos esses processos acontecem, em muitos casos, quase que simultaneamente. O escritor elabora a proposta a partir de suas leituras, inicia a escrita, revisa, replaneja, reescreve, revisa e assim por diante, em idas e vindas constantes. Essas oficinas pretenderam, portanto, abordar aspectos de compreensão em leitura para, a partir disso, chegar à produção escrita de um texto acadêmico. Tudo isso feito de modo processual, considerando as dúvidas e necessidades específicas dos acadêmicos.

As oficinas foram oferecidas a todos que participaram do estudo, tanto no momento da pilotagem quanto da coleta principal. A adesão, no entanto, foi baixa. Apenas dois acadêmicos fizeram as oficinas do início ao fim. Levantamos algumas hipóteses para a baixa adesão: 1. A



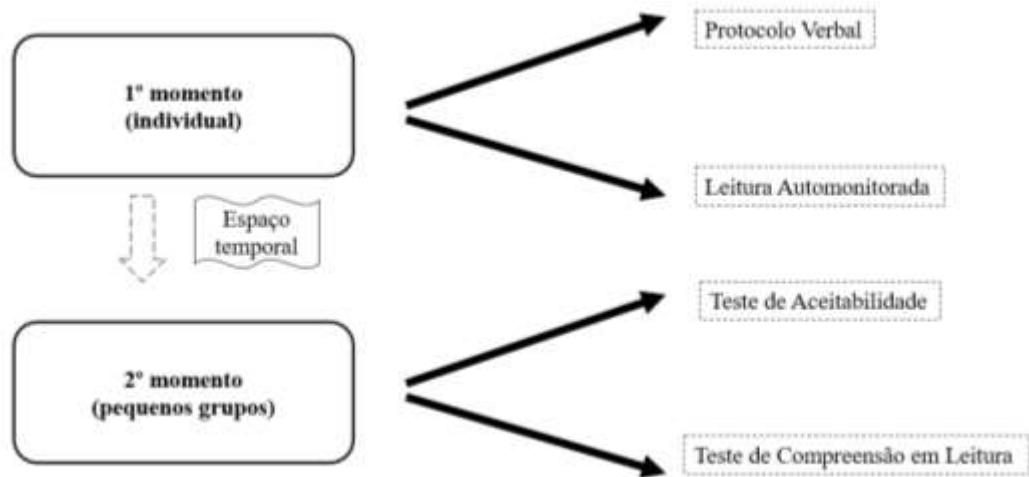
maioria havia solicitado oficinas presenciais, contudo, devido à pandemia de COVID-19, isso não foi possível; 2. A pandemia acabou sobrecarregando os acadêmicos com as atividades remotas; as oficinas foram uma atividade a mais; 3. Considerando que houve um período de suspensão do semestre letivo da UFSC por cinco meses no início da pandemia, alguns acadêmicos já haviam feito o seu trabalho de conclusão de curso quando da oferta das oficinas, então, não sentiram necessidade de participar.

### 3.3 INSTRUMENTOS E PROCEDIMENTOS DE COLETA DE DADOS

Como em toda pesquisa, é necessário fazer escolhas metodológicas que sejam as mais adequadas para se conseguir tocar no objeto investigado. Toda e qualquer técnica possui pontos fortes e pontos questionáveis. Considerando isso, optamos, nesta pesquisa, pela combinação de diferentes instrumentos, para que fosse possível contornar as limitações de um ou de outro. Em consonância com Tomitch (2008, p. 43), “a melhor conduta para se conseguir dados desejados é utilizar mais de uma ferramenta durante a coleta de dados”. Ademais, o paradigma experimental mais apropriado é aquele que mais dá conta de alcançar os objetivos pretendidos. Nesse sentido, este estudo utiliza diferentes técnicas, a fim de que seja possível testar as hipóteses levantadas e atingir os objetivos estabelecidos. Diante disso, lançamos mão de procedimentos metodológicos que auxiliam na obtenção de dados comportamentais dos participantes. As medidas comportamentais investigam os processos cognitivos a partir das reações voluntárias diante uma determinada tarefa experimental de leitura. Assim, as técnicas aqui empregadas têm por foco a compreensão como processo e a compreensão como produto.

Para que, durante a explicação dos instrumentos, o leitor desta tese possa ter um roteiro, apresentamos um esquema de como a coleta de dados foi organizada, expondo a quantidade de encontros e a ordem em que cada instrumento foi aplicado:

Figura 2: Coleta de dados



Fonte: a autora.

### 3.3.1 Protocolos Verbais de Leitura

Antes da descrição do procedimento de coleta de dados realizado por meio da técnica de protocolo verbal, é relevante trazer os fundamentos que guiam a implementação do instrumento. Assim, faz-se uma apresentação não exaustiva em torno do que vem sendo teorizado e discutido acerca dessa técnica, que tem sido usada como instrumento por diversas pesquisas que investigam leitura ao longo dos anos (AFFLERBACH; JOHNSTON, 1984; ERICSSON; SIMON, 1993; PRESSLEY; AFFLERBACH, 1995; TOMITCH, 1995; SOUZA, 2004; FINGER-KRATOCHVIL, 2010; PROCAILO, 2017).

O relato verbal é uma técnica usada para estudar processos cognitivos há bastante tempo. Platão e Aristóteles, na Grécia antiga, já incentivavam as pessoas a verbalizarem o que acontecia em sua mente (ERICSSON; SIMON, 1993; PRESSLEY; AFFLERBACH, 1995; SOUZA; RODRIGUES, 2008; ARROYO et al, 2016). Desde lá, a técnica foi aprimorada e seu uso ampliado. Em revisão realizada por Arroyo et al (2016), verifica-se que existe uma vasta tradição no uso de protocolos verbais nas mais diversas áreas da ciência (Psicologia, Sociologia, Psiquiatria, Educação, Linguística). Na área da psicologia cognitiva, é mais recente a sua validação como instrumento para obter dados de processamento, ganhando consistência a partir dos trabalhos de Ericsson e Simon (1980; 1993).

De modo geral, o protocolo verbal está amparado na introspecção. Nessa perspectiva, o sujeito examina seus pensamentos durante ou depois de realizar uma determinada tarefa cognitiva e os verbaliza. Com isso, tem-se acesso a processos mentais, que são estados internos

transformados sucessivamente pelas informações processadas (ERICSSON; SIMON, 1993). De acordo com os autores, para que esse processamento ocorra, a informação precisa ser manipulada e armazenada em memórias que apresentam diferentes capacidades e características. Uma informação recentemente recebida se encontra na memória de trabalho para que possa ser manipulada. Está, assim, disponível para que seja verbalizada (SOUZA; RODRIGUES, 2008) sem que recursos atencionais extras sejam requeridos.

Neste estudo, a atividade cognitiva com a qual se trabalha é a leitura. Nessa área, os protocolos verbais podem ser úteis para investigar uma série de processos que ocorrem durante essa atividade cognitiva. O delineamento experimental depende dos objetivos investigativos. É possível obter, dos protocolos verbais, dados provenientes de três formatos: auto-relatório, auto-revelação (retrospectivo) e auto-observação (concorrente/coocorrente) (COHEN, 1987; ERICSSON; SIMON, 1993; TOMITCH, 2007). O primeiro se volta para uma descrição geral sobre o que os leitores fazem (ou acham que fazem) durante a leitura de textos (em geral). O segundo é obtido imediatamente após a leitura de um texto específico, quando o leitor traz a sua percepção sobre o que acabou de ler. Já o terceiro, o protocolo coocorrente (de auto-observação), é obtido durante a realização da tarefa. Isso ocorre para que seja possível acessar a informação que está sendo manipulada na memória de trabalho. Nesse caso, a verbalização coocorrente é a ponte que leva o pesquisador a ter acesso aos processos ocorridos durante a atividade de leitura e fornece informações de processamento *on-line* já que os processos envolvidos na compreensão em leitura são aferidos no momento em que estão ocorrendo, ou seja, enquanto o participante está lendo.

Ainda, no que tange aos relatos verbais coocorrentes, Ericsson e Simon (1993) distinguem dois tipos de protocolos coocorrentes. Há aqueles que geram dados provenientes de *talk aloud* e outros que geram dados provenientes de *think aloud*. De acordo com Tomitch (2007), no primeiro, a informação proveniente da memória de trabalho é apenas verbalizada, enquanto no segundo (*think aloud*) ocorre uma tradução do conteúdo que está em forma não-verbal na memória de trabalho. O esperado é que a tradução não seja necessária e, sim, que os leitores possam verbalizar aquilo que está em sua memória de trabalho. Apesar dessa diferenciação feita por Simon e Ericsson (1993), o termo *think aloud* é o mais utilizado na área, mesmo quando o objetivo é obter dados sem que ocorra a tradução, conforme ressalta Tomitch (2007).

Utiliza-se, nesta pesquisa, a técnica de protocolos verbais coocorrentes, tendo em vista os objetivos traçados e os dados provindos desse formato. Para tanto, é necessário levar em conta algumas especificidades que serão consideradas nesta investigação. Para conseguir ter

acesso, durante a realização de protocolos coocorrentes, a dados de verbalização direta, é possível lançar mão de algumas técnicas, como a desenvolvida por Cavalcanti (1989) e adaptada por Tomitch (1995), a saber: o protocolo de pausa. Esse procedimento ajuda a reduzir a “tradução”, evitando que haja interferências por parte do pesquisador e, conseqüentemente, evitando elaborações por parte do participante. No protocolo de pausa, é dada a orientação de que o leitor deve parar durante a sua leitura para verbalizar os seus pensamentos. Na adaptação de Tomitch (1995), são inseridas marcações em determinados pontos do texto. Essa sinalização serve de lembrete para que o leitor verbalize seus pensamentos naqueles pontos delimitados, sem que qualquer tipo de intervenção seja necessário. Nesta pesquisa, o texto apresentado aos participantes foi dividido em parágrafos, que continham três pausas sinalizadas com círculos vermelhos, nas quais as verbalizações deveriam, necessariamente, acontecer.

Conforme Souza e Rodrigues (2008), com o uso de protocolos verbais é possível obter uma descrição dos processos cognitivos envolvidos na atividade da leitura. Nesse sentido, “[o] emprego da técnica dos protocolos verbais minimiza o efeito da inacessibilidade aos dados de processamento, uma vez que se obtêm informações não elaboradas acerca dos pensamentos dos sujeitos, enquanto ele está envolvido com a tarefa de leitura.” (SOUZA; RODRIGUES, 2008, p. 27). Os protocolos trazem, portanto, luz aos movimentos de compreensão, possibilitando a análise de como o indivíduo age para construir uma significação para o texto lido.

Quando uma tarefa de leitura é realizada, diferentes processos ocorrem. Leitores proficientes modificam constantemente tais processos, com a finalidade de se adequarem às demandas do texto lido. O resultado, conforme Pressley e Afflerbach (1995, p. 2, tradução nossa), “é um processamento complexo<sup>42</sup>”. Com o uso de protocolos verbais, busca-se acessar esses processos a partir das inferências construídas do relato feito pelos leitores. Para tanto, há que se considerar, conforme destacam Afflerbach e Johnston (1984), o fato de que o texto a ser lido não pode ser tão fácil aos leitores, pois, desse modo, não ocorre a desautomatização dos processos. Por outro lado, se o texto for extremamente difícil pode causar uma sobrecarga no sistema de processamento, fazendo com que os recursos sejam todos direcionados para a resolução dessa dificuldade, comprometendo a compreensão e os relatos verbais. Assim, é fundamental que o texto usado na tarefa seja adequado aos objetivos empreendidos. Além disso, há outros pontos, citados por Souza e Rodrigues (2008, p. 29) a serem considerados para que a tarefa de protocolo verbal traga resultados satisfatórios e que remetam aos processos cognitivos ocorridos durante a leitura. Destacam-se três deles a seguir:

---

<sup>42</sup> “The result is a complex processing.”

- 1) “O quanto o pesquisador deve ser instrutivo”: quando se trata de protocolos coocorrentes, é importante que o pesquisador solicite que o participante pense em voz alta com o intuito de que seja verbalizado tudo o que está ocorrendo na memória de trabalho durante a realização da tarefa, sem exames ou análises (cf. ERICSSON; SIMON, 1993).
- 2) “O quanto os participantes devem ser submetidos a atividades de treinamento”: é fundamental que o treinamento seja feito até que o participante compreenda como realizar a tarefa, pois o desconhecimento sobre o modo de realizar a tarefa pode afetar os dados.
- 3) “O controle das diferenças individuais em termos de espontaneidade e facilidade de expressão verbal”: essas diferenças individuais podem interferir independentemente da técnica escolhida. Para minimizar esse efeito, trabalha-se com o número de participantes considerável (nesta pesquisa, é um total de vinte e seis). Além disso, o próprio fato de serem oferecidas instruções claras e treinamento, minimiza esse efeito.

Assim, na investigação ora apresentada, todos esses fatores foram considerados para que os dados representassem, de fato, os processos cognitivos empreendidos durante a tarefa de leitura. Na subseção seguinte, serão detalhados os procedimentos da coleta de dados com a tarefa de Protocolo Verbal.

#### 3.3.1.1 Procedimento de coleta de dados

A sessão de protocolo verbal ocorreu de modo individual com cada participante. Para a realização da técnica, os participantes receberam instruções escritas (Apêndice A), que foram lidas pela pesquisadora juntamente com o participante antes da execução da tarefa. De modo geral, as instruções solicitavam que o texto fosse lido silenciosamente e que todos os pensamentos fossem verbalizados durante a leitura, sobretudo, nos pontos marcados com um círculo vermelho. Tais pontos estavam no meio e no fim de cada parágrafo (CAVALCANTI, 1989; TOMITCH, 1995).

A tarefa foi realizada por meio de notebook, 14 polegadas. O texto foi apresentado aos participantes em um arquivo do programa *Power-Point*, cada parágrafo correspondia a um *slide*. Os relatos dos participantes foram gravados em áudio em gravador de celular iPhone, modelo 6S. Apenas a sessão principal foi gravada; a de treinamento não.

Após a leitura das instruções da atividade, foi feita uma sessão de treinamento com um texto acadêmico-científico diferente daquele que seria lido na tarefa experimental. Nesse momento, a pesquisadora permaneceu ao lado do participante orientando sobre possíveis dúvidas e dificuldades acerca da tarefa. Treinar a tarefa experimental é fundamental para garantir que o participante compreenda o que deve ser feito, evitando qualquer problema relacionado ao desconhecimento ou a inexperience com o tipo de tarefa. Durante a leitura do texto experimental, a pesquisadora não interferiu, posicionando-se atrás do participante, com a intenção de não ficar visível (ERICSSON; SIMON, 1993), apenas observando os possíveis movimentos a serem realizados no texto. Com base na pilotagem, o tempo esperado para a realização da atividade era de 20 minutos. Essa etapa da coleta de dados ocorreu em uma sala da Universidade Federal de Santa Catarina, disponibilizada pelas professoras que fazem uso do espaço. Nos momentos da coleta, a sala foi de uso exclusivo para a pesquisa.

Nesse teste, as variáveis consideradas independentes são o pronome “ele” e o sintagma “o mesmo”, já as variáveis dependentes são as estratégias expostas pelos leitores durante a verbalização de seus pensamentos.

### **3.3.2 Teste de Leitura Automonitorada**

Como destacado inicialmente, em uma investigação é importante cruzar dados obtidos por meio de diferentes técnicas. Tendo em vista o objeto investigado no presente estudo, a técnica de leitura automonitorada, chamada pela literatura internacional de *self-paced reading*, se mostra bastante pertinente, pois se trata de técnica comportamental *on-line*, que capta os processos cognitivos em curso por meio de medidas de reação (KAISER, 2013; MAIA; CUNHA-LIMA, 2014; LEITÃO, 2015). De acordo com Mitchell (2004), muitas das operações cognitivas envolvidas no processamento da linguagem ocorrem em milésimos de segundos e os métodos *on-line* são importantes para alcançar esses processos. Mais especificamente, as técnicas que envolvem medidas de tempo de reação buscam conhecer quão rápido uma pessoa realiza uma tarefa linguística – neste caso, a leitura. O tempo obtido fornece indicações acerca das complexidades envolvidas no processamento, tempos de reação baixos sugerem menores custos para o processamento; já tempos de reação mais elevados assinalam uma maior carga cognitiva para o processamento de determinado aspecto.

Como o nome indica, a leitura, nesta técnica, é monitorada pelo próprio leitor. Para aplicar a técnica, um estímulo linguístico é segmentado palavra por palavra, sintagma por sintagma ou sentença por sentença. Cada segmento é exibido na tela do computador, um por

vez, a partir do acionamento de uma tecla-chave, que requisita o surgimento da próxima fração de texto. A medida fornecida diz respeito ao tempo transcorrido entre o momento em que o segmento aparece na tela até o momento de pressionar o botão.

As versões do teste podem ser formuladas de diferentes modos. Em algumas, a tela se move da esquerda para a direita, revelando a porção de texto seguinte sempre que o leitor requerer, ou seja, a tela se move, mas os segmentos são apresentados sempre na mesma posição (no centro da tela). Essa versão é conhecida como *stationary-window*. Outra versão é conhecida como *moving-window*. Nesse contexto experimental, a dinâmica funciona do seguinte modo: a sentença é apresentada como um todo e cada parte dela ocupa um espaço na tela, toda vez que o leitor pressiona a tecla-chave, uma parte é revelada. Essa versão, de acordo com Mitchell (2004), pode ser aplicada de modo cumulativo ou não-cumulativo. No primeiro, cada segmento que é revelado permanece na tela, assim, ao final, o leitor tem acesso ao estímulo inteiro. Na versão não-cumulativa, cada vez que o leitor passa para o segmento seguinte, o anterior fica encoberto (ou seja, cada segmento aparece em um momento específico, sem acumular, e não aparecem no mesmo local da tela).

No presente estudo, a versão utilizada é a *moving-window* não-cumulativa, sua vantagem está no fato de a situação de leitura ser mais próxima do natural, em que as palavras são apresentadas em suas respectivas posições na frase. Nessa versão, a abordagem não-cumulativa é interessante, porque o processamento linguístico e o comando dado pelo leitor são mais possíveis de serem aproximados. Isso porque é necessário fazer a leitura de cada fragmento no momento em que este é revelado. Por outro lado, na versão cumulativa, o leitor pode optar por revelar o estímulo todo e, depois, fazer a leitura, o que não torna possível relacionar tempo de leitura e processamento de cada segmento, impedindo a inferência de quais partes de um estímulo podem estar associadas a uma sobrecarga do processamento.

Além de permitir o acesso a dados de processamento *on-line*, a técnica de leitura automonitorada é relativamente simples e portátil. Isso traz algumas vantagens práticas, porque o pesquisador precisa, basicamente, de um programa específico instalado em seu computador para poder realizar o experimento. Esse é um fator relevante de ser considerado no momento de pensar a técnica experimental a ser utilizada. Para além disso, no entanto, no que tange às investigações relativas ao processamento anafórico, a leitura automonitorada tem sido amplamente utilizada (CUNHA-LIMA, 2004; LEITÃO, 2005; QUEIROZ, 2009; MAIA, 2013; SIMÕES, 2014; CORREIA, 2014; GONDIM, 2017) e tem-se mostrado, como ressalta Mitchell (2004), uma técnica efetiva cujos resultados encontrados têm sido corroborados por outras técnicas, como o rastreamento ocular (LEITÃO; RIBEIRO; MAIA, 2012). Nos estudos acerca

do processamento anafórico, um dos pontos investigados é a eficiência de uma ou outra forma anafórica durante a leitura. Assim, é feita a comparação entre os tempos de leitura das duas formas investigadas e aquela lida mais rapidamente dá indícios de maior eficiência para o processamento (LEITÃO, 2015).

### 3.3.2.1 Procedimentos de coleta de dados

Quando se realiza uma pesquisa, é preciso considerar alguns fatores que são relevantes para a obtenção de dados consistentes. No que tange a uma experimentação realizada com a técnica de leitura automonitorada, é importante levar em conta: o número de participantes, o número de sentenças experimentais e distratoras, a randomização das frases dentre outros fatores que serão abordados neste momento do texto.

Como apresentado na seção 3.2, o experimento foi realizado com vinte e seis participantes, o que está dentro do que se tem estabelecido como padrão em investigações envolvendo a leitura automonitorada (LEITÃO, 2005; QUEIROZ, 2009; GADELHA, 2012). Os sujeitos são acadêmicos do curso de Engenharia de Controle e Automação, ou seja, pertencem a um mesmo grupo, além disso, estão dentro de uma mesma faixa etária e possuem nível de formação bastante semelhante.

Este experimento de leitura automonitorada foi desenvolvido na versão *moving-window* não cumulativa, em que as sentenças estavam divididas em sintagmas. Foram lidas trinta e oito sentenças, sendo doze sentenças experimentais e vinte e seis distratoras. Isso porque, na metodologia experimental, o número de sentenças distratoras deve se aproximar do dobro das sentenças-alvo, com o intuito de mascarar o objetivo do teste. Como esse teste foi realizado no mesmo encontro em que ocorreram os protocolos verbais, também foi importante não ampliar muito o número de sentenças a serem lidas, para se garantir a atenção dos participantes até o final.

Dentre as sentenças experimentais, seis tinham a anáfora realizada com “a mesma/o mesmo” (três de cada), seis com “ela/ele” (três de cada). Nas frases, como é possível ter um controle maior das estruturas formuladas (diferentemente do texto, que foi adaptado de um já existente e publicado), equilibrou-se o gênero da anáfora para poder controlar o seu efeito, embora não seja objetivo da pesquisa discutir diferenças de processamento relacionadas ao gênero. Para a leitura das sentenças, utilizou-se o quadrado latino. Isso significa que cada participante foi exposto a todas as condições experimentais, ou seja, os vinte e seis participantes



leram sentenças com “o mesmo” e com “ele”, mas sem repetição de sentença. Para tanto, havia dois grupos de sentenças. Um participante leu, por exemplo, as seguintes sentenças:

O fato/ ressalta/ a crise na ciência./ O mesmo/ revela,/ com clareza,/ as fragilidades encontradas.  
O tema/ permite/ inúmeras abordagens./ Ele/ provoca/ reações diversas/ no campo de estudo.

Por sua vez, outro participante leu o seguinte:

O fato/ ressalta/ a crise na ciência./ Ele/ revela,/ com clareza,/ as fragilidades encontradas.  
O tema/ permite/ inúmeras abordagens./ O mesmo/ provoca/ reações diversas/ no campo de estudo.

Ao final das sentenças, os participantes deveriam responder a uma pergunta de controle de atenção, que, nas sentenças experimentais, servia, também, para saber se a correferência tinha sido estabelecida adequadamente. Todas as perguntas tinham como resposta “sim” ou “não”, sendo que as respostas estavam igualmente divididas, isto é, dezenove perguntas respondiam “sim” e dezenove respondiam “não”. Todas as perguntas das condições experimentais respondiam “sim”. No caso dos exemplos trazidos, as perguntas que seguem as sentenças são as seguintes:

O fato revela com clareza as fragilidades encontradas?

O tema provoca reações diversas no campo de estudo?

Respostas esperadas: Sim

Neste estudo, portanto, a variável independente é: o tipo de retomada (sintagma “o mesmo” e o pronome “ele”). Já as variáveis dependentes são: o tempo de leitura dos segmentos de interesse (o segmento pós-crítico - verbo localizado imediatamente após a anáfora realizada com o pronome e com o sintagma “o mesmo” e o tempo total de leitura da frase) e a acurácia da resposta à pergunta de controle de atenção. O processamento do tipo de retomada será explorado a partir de dois tempos de leitura, o do segmento pós-crítico e o tempo total de leitura da sentença. Isso porque, conforme apresentado na seção 2.3.1, estudos (MAIA, 2013; GONDIM, 2017) têm mostrado diferentes resultados a depender do tempo levado em conta para as análises.

Além disso, cabe esclarecer por que consideramos o tempo de leitura do segmento pós-crítico: porque o sintagma “o mesmo” pode assumir diferentes funções em um texto, como apresentado em estudos anteriormente realizados (AMORIM, 2009; PEREIRA, 2013), que foram discutidos no capítulo 2 deste trabalho. Nesse sentido, o reconhecimento desse sintagma como anafórico ocorre quando a construção sintática é acessada pelo leitor no momento da leitura do verbo. Isso porque é possível que, em um texto, ele assuma diferentes valores de significado, exatamente porque o sintagma “o mesmo” não é um pronome pessoal e, por isso, diferentemente do “ele”, não tem o valor anafórico em sua essência, embora, nesse contexto específico, pareça estar assumindo a função desse pronome pessoal.

Para rodar o experimento, utilizou-se o software DMDX<sup>43</sup>, versão 5.3.2.0, desenvolvido por Jonathan Forster e Ken Forster do Departamento de Psicologia da *University of Arizona*. Nesse software, o tempo de leitura é contabilizado em milésimos de segundos, o que gera uma medida bastante precisa.

Antes da aplicação do experimento, os participantes receberam instruções sobre o teste que realizariam (Apêndice B) e, na sequência, fizeram um treinamento com o objetivo de se habituarem com a tarefa. Assim como ocorre com os protocolos verbais, a sessão de treinamento, nesse tipo de experimento, é importante para que a operacionalização da tarefa não se torne um problema. Quanto ao procedimento, o participante deveria, primeiramente, ler as orientações que apareceriam na tela do computador, clicando na barra de espaço do teclado para passar para frente. Após ler a pergunta “Pronto para começar?”, clicava na barra de espaço pela última vez e, então, aparecia a sentença encoberta. Para revelar o primeiro segmento, o participante devia clicar na seta direita do teclado (marcada com a cor verde) e assim deveria proceder sucessivamente até ler as duas frases justapostas. Na sequência, aparecia a pergunta de controle de atenção. Caso a resposta fosse “sim”, o participante clicava na seta direita. Caso a resposta fosse não, clicava na tecla “ctrl” esquerda (marcada com a cor vermelha). As frases foram randomizadas com a intenção de evitar qualquer efeito de ordenação. Na sequência, traz-se um exemplo de como as sentenças eram apresentadas aos participantes:

“ \_\_\_\_\_ , \_\_\_\_\_ , \_\_\_\_\_ .”  
 “O fato \_\_\_\_\_ .”  
 “ \_\_\_\_\_ resalta \_\_\_\_\_ .”

<sup>43</sup> O programa está disponível para download na seguinte página: <http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm>.

“\_\_\_\_\_ a crise na ciência. \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_.”  
 “\_\_\_\_\_. O mesmo \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_.”  
 “\_\_\_\_\_ revela, \_\_\_\_\_, \_\_\_\_\_.”  
 “\_\_\_\_\_, com clareza, \_\_\_\_\_.”  
 “\_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, as fragilidades encontradas.”;  
 “O fato revela com clareza as fragilidades encontradas?”

Cabe destacar, ainda, que todas as sentenças experimentais utilizadas nesse experimento foram analisadas por três avaliadoras. Além disso, tanto sentenças experimentais quanto distratoras tiveram a sua aceitabilidade avaliada por um grupo de estudantes do curso de Letras - Português e respectiva Literatura, da Universidade Federal de Santa Catarina (UFSC), e por um grupo de estudantes do ensino médio do Instituto Federal Catarinense (IFC), que deveriam julgar, de modo escalonado, a aceitabilidade para a língua portuguesa de cada uma das sentenças. Essas avaliações foram importantes para saber se as sentenças estavam coerentes e compreensíveis para um leitor falante de língua portuguesa. A partir disso, foi possível fazer ajustes que tornassem as sentenças o mais claras e naturais possíveis, ainda que considerando todos os controles adotados. Tais avaliações serviram, também, como piloto para o teste de aceitabilidade (seção 3.3.3).

Embora se saiba que o processo que ocorre durante um experimento de leitura automonitorada não se compara ao da leitura natural, buscou-se criar condições para deixá-lo o mais natural possível. Por conta disso, a segmentação das sentenças é feita de acordo com os sintagmas que as compõem (e não por palavras). Assim, os elementos que constituem um mesmo sintagma não são separados. Essa decisão foi tomada pensando na integração das informações no momento da leitura (GAGNÉ; YEKOVICH; YEKOVICH, 1993), porque, da mesma forma que se quer a medida temporal da leitura, se quer evitar que a forma como a divisão é apresentada aumente esse tempo. Entende-se que uma segmentação que desrespeite os níveis de estruturação e organização da sentença pode afetar os tempos de leitura e, desse modo, influenciar de forma indesejada as variáveis experimentais que estão em análise.

Dito isso, vale destacar, também, como as sentenças foram pensadas e quais os controles adotados. São sentenças justapostas em que antecedente e anáfora estão na posição de sujeito (havendo, portanto, um paralelismo estrutural). Essa escolha se deu porque, conforme apontam alguns estudos realizados (LEITÃO, 2015; NÓBREGA-LIMA, 2014), o paralelismo estrutural tende a ser um facilitador para o processamento de pronomes anafóricos. Como são frases

justapostas, a interpretação da anáfora se dá por meio do contexto, já que não há vínculo sintático entre as orações. Além disso, como se trabalha também com o pronome “ele”, optou-se pela posição de sujeito para que não houvesse qualquer adição temporal por conta de estar em uma posição de objeto, especialmente objeto direto, porque se está trabalhando com universitários e com textos acadêmico-científicos, ou seja, leva-se em conta a modalidade escrita formal da língua. Nesse sentido, o uso do “ele” como objeto poderia gerar uma medida temporal de leitura maior (cf. MAIA, 2013) e o que se quer é garantir que o pronome “ele” esteja em condições favoráveis ao processamento para saber se o sintagma “o mesmo” age de modo semelhante nesse contexto, conforme hipotetizado.

A estrutura formada por frases justapostas se deu com a finalidade de seguir o mesmo padrão do texto acadêmico com o qual foi trabalhado, isto é, tanto nas sentenças experimentais quanto no texto, antecedentes e anáforas estão na mesma posição (sujeito), em sentenças distinta e sem qualquer conectivo. Cabe destacar que, embora existam resultados (LEITÃO; LIMA; CALAÇA, 2010) que indiquem um aumento no tempo de leitura do pronome “ele” em contextos de justaposição, essa construção foi a que pareceu mais natural e tem se mostrado uma posição propícia para um processamento menos custoso para o pronome quando comparado ao nome repetido (GONDIM, 2017; GORDON; GROSZ; GILLION, 1993), ou seja, o pronome acaba se mostrando um anafórico adequado em termos de processamento também em frases justapostas. Acredita-se, ainda, que, como todos os outros fatores controlados levam a uma diminuição nesse tempo, a justaposição de frases não será um problema. O que se quer dizer é que os controles aplicados procuraram garantir um contexto favorável ao processamento do pronome “ele”, pois o esperado é que o sintagma “o mesmo” seja também um anafórico favorável ao processamento nessas mesmas condições.

Tendo em vista, ainda, que não se está trabalhando com nível de decodificação (em que o grafema seria unidade fundamental) e considerando que a sílaba é a menor unidade que tem uma realidade psicológica no português brasileiro, todas as sentenças foram controladas com base no número de sílabas<sup>44</sup>. Isto é, todas as sentenças possuem de trinta e uma a trinta e cinco sílabas e a distância entre antecedente e anáfora nas frases experimentais é de onze sílabas. Essa distância está dentro do padrão usado em outros estudos realizados, que tem variado entre dez e quinze sílabas (GONDIM, 2017; NÓBREGA-LIMA, 2014; LEITÃO et al 2010). Procurou-

---

<sup>44</sup> As sílabas consideradas foram as escritas, pois o intuito foi padronizar o comprimento dos estímulos visuais para o teste experimental e, como a cadência de leitura natural não é totalmente possível (visto que as sentenças estão segmentadas em sintagmas), optou-se pela padronização da sílaba conforme escrita. Além disso, foram considerados como ditongos somente os decrescentes. Todos os possíveis crescentes foram tratados como hiatos, para garantir uniformidade e critério na análise.

se manter uma distância próxima entre antecedente e anáfora, considerando que o pronome tende a ser usado como anafórico nessas circunstâncias de proximidade (KINTSCH; RAWSON, 2013) e que o uso do sintagma “o mesmo” também tem se mostrado favorável em contextos de maior proximidade com o antecedente (MOREIRA, 2007). Ainda, as frases experimentais possuem seis antecedentes femininos e seis masculinos, todos eles possuem três sílabas (contando com o artigo que o acompanha) e são [-animados]. O segmento pós-crítico (verbo) contém, em todas as frases experimentais, três sílabas. Todas as sentenças estão divididas em sete segmentos, nas experimentais, o segmento cujo tempo de leitura é analisado é o quinto.

Ademais, considerando o tipo de texto e o contexto com o qual se está trabalhando, buscou-se usar, nas sentenças, palavras que façam parte do meio acadêmico. A estrutura da sentença também foi construída partindo daquilo que é mais comum em um texto acadêmico-científico (linguagem impessoal e objetiva). As sentenças distratoras também foram construídas seguindo esse padrão, contendo duas frases justapostas com trinta e uma a trinta e cinco sílabas, em que a primeira é introduzida por um antecedente [-animado]. No entanto, a estrutura das passagens distratoras é diversa das experimentais.

Para a rodada estatística, temos um total de 156 tempos de leitura obtidos para o pronome “ele/ela” e 156 para o sintagma “o mesmo/a mesma” (isso porque cada participante leu 6 frases com “ele/ela” e 6 frases com “o mesmo/ a mesma” e foram um total de 26 participantes).

Apresentados todos os procedimentos e cuidados tomados com o experimento de leitura automonitorada, passa-se para o teste de julgamento de aceitabilidade.

### **3.3.3 Teste de Julgamento de Aceitabilidade**

Outra técnica bastante empregada em trabalhos que se debruçam sobre o estudo do processamento anafórico é o teste de julgamento de aceitabilidade. Na presente investigação, essa técnica faz parte do desenho metodológico e será apresentada nesta seção. Antes de tratar de quais foram os procedimentos para a coleta de dados, cabe conhecer a técnica em si.

O julgamento de aceitabilidade é um instrumento que fornece dados de processamento *off-line*, pois capta a resposta reflexiva, após o processamento. Em muitos estudos, tal técnica é denominada de julgamento de gramaticalidade. No entanto, de acordo com Schütze e Sprouse (2013), o termo não é o mais apropriado, porque a gramática é um “[...] construto mental não

acessível [...]”<sup>45</sup>” (p. 28, tradução nossa), o que inviabiliza os participantes de manifestarem seu julgamento acerca da gramaticalidade de uma frase. Assim, a percepção que falantes de uma língua têm como reação espontânea a um determinado estímulo linguístico é mais bem denominada de julgamento de aceitabilidade. Trata-se, como dito, de uma percepção, que é uma medida um tanto quanto subjetiva, e não há métodos que tragam dados diretos de como essa percepção é processada na mente do participante. Há, contudo, artifícios a serem usados que podem garantir dados bastante informativos acerca do objeto estudado. Esses artifícios dizem respeito às escolhas e aos formatos a serem aplicados na construção do teste de julgamento de aceitabilidade.

De modo geral, as tarefas de julgamento podem ser divididas em dois blocos: aquelas qualitativas (não-numéricas) e as quantitativas (numéricas). A escolha a ser feita vai depender dos objetivos investigativos. Dentre as tarefas de julgamento qualitativas, há aquelas de sim-não, em que o participante responde se o estímulo recebido é aceitável (sim) ou não é aceitável (não). Há, também, em uma abordagem que se pretende qualitativa, tarefas denominadas de escolha forçada (*forced-choice*). Nestas são dadas duas ou mais opções ao participante, e ele deve escolher a mais ou a menos aceitável. Dentre as tarefas quantitativas, tem-se a de estimativa de magnitude e a de escala Likert. Na estimativa de magnitude, os participantes recebem uma sentença referência, para a qual atribuem um determinado valor de aceitabilidade, as demais sentenças são, então, classificadas a partir desse padrão estabelecido na sentença de referência (SCHÜTZE; SPROUSE, 2013).

A presente investigação trabalhou com julgamento de aceitabilidade guiado pela escala do tipo Likert, que busca quantificar a percepção dos leitores. Nesse tipo de tarefa de julgamento, a sentença, apresentada aos participantes, deve ser julgada dentro de uma escala pré-estabelecida. A escala Likert tem sua origem no ramo empresarial e foi construída com uma escala de cinco pontos. De modo geral, o mais comum é que essa escala seja construída com um número ímpar de pontos (por exemplo: de 1 a 5, de 1 a 7) porque, de acordo com Schütze e Sprouse (2013, p. 33, tradução nossa), “[...] números ímpares contêm um ponto intermediário preciso [...]”<sup>46</sup>. Isso favorece a percepção dos participantes acerca das delimitações entre o considerado não-aceitável e o aceitável, pois o ponto intermediário representa uma posição neutra.

Importa, por fim, ressaltar que, embora existam diferenças entre cada uma das tarefas, em essência, todas se pautam no mesmo tipo de tarefa cognitiva, qual seja: reportar a percepção

---

<sup>45</sup> “[...] mental constructo not accessible [...]”

<sup>46</sup> “[...] odd numbers contain a precise middle point [...]”.

da aceitabilidade de um determinado estímulo linguístico. Os dados provindos de julgamento assumem um papel relevante nas pesquisas linguísticas, pois fornecem informações sobre o *status* do fenômeno estudado, informações essas que não seriam possíveis de se obter por meio de outras técnicas. No caso do julgamento de aceitabilidade a partir de uma escala do tipo Likert, é possível obter o tamanho da diferença de percepção entre as condições investigadas.

### *3.3.3.1 Procedimento de Coleta de dados*

No teste de aceitabilidade aplicado nesta investigação, trabalhou-se, como citado, com uma escala do tipo Likert de sete pontos em que os participantes deveriam avaliar a aceitabilidade das mesmas frases que haviam lido durante o experimento de leitura automonitorada. A escala de 7 pontos foi escolhida por ter sido a escala utilizada em pesquisas como a de Maia (2013), que estuda processamento anafórico e busca avaliar a aceitabilidade de diferentes formas de retomadas. Além disso, uma escala de 7 pontos permite mapear de forma mais minuciosa a percepção dos participantes, fornecendo uma melhor discriminação das respostas e mostrando se a aceitabilidade tende a subir ou a descer, a depender da forma anafórica utilizada. Uma escala com mais pontos permite, assim, que mais nuances da percepção dos participantes sejam capturadas.

Há críticas ao uso de escalas com mais pontos, porque isso pode acabar confundindo os participantes por conta da distância pequena entre esses pontos. No entanto, estudos como o de Simms et al (2019) mostram consistência em escalas com pontos que vão de 4 a 7. Os autores buscaram investigar se há um número psicométrico mais adequado para as opções de resposta e se é melhor incluir números pares ou ímpares de opções. No que tange ao número de opções, os autores encontraram, resumidamente, que: a média de escores dos participantes tende a diminuir quando há poucas opções de respostas – o escore se mostra mais consistente quando há entre 4 e 7 opções; quanto menor o número de opções, menores foram os níveis de fidedignidade, que se mostrou prejudicada também em testes com mais de 7 opções. Assim, a partir das análises, os autores mostram que escalas com 6 ou 7 pontos tendem a ser as melhores opções. Como em toda e qualquer técnica, há escolhas a serem feitas, conhecendo os pontos fortes e fracos de cada instrumento e de cada opção feita pelo pesquisador conforme os objetivos e o contexto da pesquisa.

No que tange aos valores a serem determinados, os participantes deveriam atribuir 1 para as sentenças consideradas muito inaceitáveis e 7 para as que considerassem totalmente aceitáveis, fazendo a gradação entre esses dois valores de acordo com a percepção deles. De

modo mais específico, cada ponto da escala representou o seguinte: 1- muito inaceitável, 2- inaceitável, 3- pouco aceitável, 4- nem aceitável e nem inaceitável (ponto neutro), 5- relativamente aceitável, 6- aceitável, 7- totalmente aceitável. A instrução dada aos participantes antes da tarefa solicitou que eles lessem as frases e atribuíssem um valor, levando em conta a sua percepção de falantes da língua, sem considerar o conteúdo das sentenças, mas sim se pareciam aceitáveis para serem proferidas na língua portuguesa. Buscamos, a partir dessa instrução, fornecer uma breve contextualização da tarefa a ser realizada.

Foram doze estímulos experimentais, dos quais seis têm a anáfora correferencial realizada com “o mesmo/a mesma” e seis com o pronome “ele/ela”, e vinte e seis estímulos distratores. Sendo assim, temos um total de 156 avaliações para o pronome “ele/ela” e 156 avaliações para o sintagma “o mesmo/a mesma” (isso porque cada participante leu 6 frases com “ele/ela” e 6 frases com “o mesmo/ a mesma” e foram um total de 26 participantes).

Cabe o destaque para o fato de que houve uma distância de, pelo menos, dez dias entre o experimento de leitura automonitorada e o teste de aceitabilidade. Assim, é sabido que, por terem lido os estímulos anteriormente, há um risco de ocorrência de *priming*, já que é possível que os participantes tenham uma breve lembrança das frases. No entanto, devido ao número de sentenças, o mais provável é que o julgamento não tenha sido afetado em nosso estudo.

A realização do teste aconteceu no laboratório de informática do Centro de Comunicação e Expressão (CCE) da Universidade Federal de Santa Catarina. Cada participante estava em uma máquina. Inicialmente, foram lidas as instruções (Apêndice C) da tarefa a ser realizada, que foram, também, apresentadas na tela do computador em arquivo *Portable Document Format* (PDF), para que todos pudessem acompanhar a leitura. Na sequência, solicitou-se que os participantes maximizassem o arquivo em Excel, no qual os estímulos (Apêndice D), randomizados, foram apresentados. Dessa forma nenhum participante leu as sentenças em uma mesma ordem. Além disso, assim como na leitura automonitorada, utilizou-se o quadrado latino para a apresentação dos estímulos, o que significa que cada participante foi exposto a todas as condições experimentais, sem repetição de sentença.

Os acadêmicos foram orientados a usarem a sua percepção de falantes e leitores da língua portuguesa para julgarem as sentenças. A finalidade estava em conseguir que eles quantificassem, de modo escalonado, a aceitabilidade das construções anafóricas realizadas com o sintagma “o mesmo” e com o pronome “ele”. Com esses dados em mãos, foi possível obter indícios dos processos cognitivos que ocorrem durante a leitura de sentenças com essas retomadas anafóricas específicas.



Existem algumas críticas lançadas ao julgamento de aceitabilidade com escalas do tipo Likert. Uma delas, que atinge especialmente a essa pesquisa, está em escalas com mais pontos e que são apresentadas de modo numérico aos participantes. As críticas sugerem que pode se tornar difícil ao participante diferenciar o que representa um 2 e um 3, por exemplo, ou um 5 e um 6 dentro da escala. Assim, acaba sendo pouco intuitivo a resposta escolhida. Para amenizar esse problema, acrescentamos (verbalmente), na instrução fornecida, a informação sobre o que cada ponto representaria na escala. Ainda assim, optamos por deixar a escala numérica no arquivo ao qual os participantes tiveram acesso.

### 3.3.4 Teste de Compreensão em leitura

O teste de compreensão em leitura é um instrumento importante para se obter dados da compreensão como produto (TOMITCH, 2008). Isso significa que esse tipo de instrumento fornece dados do processamento *off-line*, permitindo a avaliação da compreensão em leitura dos participantes da pesquisa. No caso desta investigação, além do processamento *off-line*, captamos, também, o comportamento estratégico no curso das respostas às questões. Desse modo, temos, também, dados do processamento *on-line*, ou seja, dados das ações realizadas pelos participantes durante o processo de leitura do texto e de resposta às questões do teste de compreensão (SOUZA; FRANZEN; SCHLICHTING, 2019; SCHLICHTING, 2021).

Ao trabalhar com esse tipo de tarefa, é possível englobar desempenho e competência em leitura com o uso de uma única técnica. O desempenho pode ser aferido a partir do *score* que os participantes obtêm nas perguntas de compreensão; já a competência é revelada a partir do que os participantes mobilizam de habilidades, estratégias e conhecimentos para realizar o teste em si. Para que o pesquisador consiga captar esses aspectos, é necessário desenvolver um teste com critérios bem estipulados e fundamentados (CARVALHO, 2022). Isso significa dizer que todo o processo de desenvolvimento do teste, desde a escolha do texto até o desenvolvimento das questões, requer uma sustentação clara e objetiva, tanto para que a replicabilidade, tão importante em pesquisas experimentais, seja possível, quanto para que os dados obtidos tragam respostas acerca do que está sendo investigado.

Dito isso, o teste empregado neste estudo tomou por base as orientações que guiam o desenvolvimento do exame aplicado pelo Programa Internacional de Avaliação de Estudantes<sup>47</sup>,

---

<sup>47</sup> Programme for International Student Assessment.

PISA-2015<sup>48</sup>. Assim, antes de tratar de modo específico sobre o teste realizado com os participantes da investigação, é importante conhecer o PISA e como o domínio da leitura é organizado pelo Programa.

O exame do PISA é aplicado a cada três anos e busca avaliar não apenas o que os alunos sabem, mas o que sabem fazer com o conhecimento que têm (BRASIL, 2016). Como é uma avaliação aplicada internacionalmente, é possível focar nos conhecimentos e nas habilidades dos estudantes brasileiros e comparar os resultados obtidos com os de outros países que também aplicam a avaliação. A prova é formulada por especialistas dos diferentes países integrantes e convidados da Organização para Cooperação e Desenvolvimento Econômico (OCDE). O exame é elaborado de modo criterioso, cobrindo três áreas cognitivas: Leitura, Matemática e Ciências. Cada vez que ele é aplicado, embora todos os domínios sejam avaliados, um deles é tomado como principal. Em 2015, por exemplo, o foco esteve em Ciências, já em 2018 o foco foi a Leitura.

Além de a prova ser rigorosamente elaborada por especialistas que partem de concepções atualizadas acerca dos aspectos cognitivos envolvidos na avaliação, e ter uma concepção de testagem bastante criteriosa, a opção por pautar-se nos critérios do PISA se deu, também, pelo fato de a concepção de leitura que guia a avaliação ter fundamentos bem embasados e que convergem com a concepção de leitura adotada nesta tese. Isto é, o Programa compreende a leitura como um processo cognitivo pautado em “conhecimentos, habilidades e estratégias” (BRASIL, 2016), que se desenvolvem ao longo da vida, além de partir de uma perspectiva em que a compreensão em leitura é construída de modo cooperativo. Assim, o leitor

produz significado como reação ao texto, usando conhecimentos prévios e pistas textuais e situacionais que normalmente derivam de sua sociedade e de sua cultura. Ao construir significados, ele usa vários processos, habilidades e estratégias para promover, monitorar e manter o entendimento. (BRASIL, 2016, p. 92).

Como é possível perceber, a concepção adotada pelo Programa é condizente com as concepções de leitura que guiam este trabalho. Por conta disso, parte-se dos critérios adotados na prova para construir o instrumento aplicado nesta pesquisa. Isso significa que a escolha do texto e a construção das questões de compreensão não ocorreram de modo aleatório, mas orientado pelas especificações disponibilizadas nos documentos formulados pelo Programa.

---

<sup>48</sup> Como no momento da elaboração do teste de competência em leitura o relatório de último PISA aplicado em 2018 ainda não estava disponível, tomou-se por base as orientações que guiaram o PISA-2015.

Na avaliação do PISA, o domínio de leitura é representado por elementos que são considerados fundamentais para a escolha do texto e para a elaboração dos itens que compõem o teste. São eles: situação de leitura, texto e processos cognitivos. A situação de leitura diz respeito à finalidade com a qual um texto é lido. Nesse sentido, o Programa adota diferentes categorizações (pessoal, pública, educacional e ocupacional). Tal classificação é pensada para fins de avaliação, mas não se pretende exaustiva e nem taxativa. Dessa forma, há textos que podem ser compreendidos como pertencentes a mais de uma categoria, a exemplo do texto literário, que pode fazer parte de uma situação pessoal, mas também está presente na situação educacional.

O segundo elemento citado, o texto, diz respeito aos materiais a serem lidos, os quais recebem categorizações concernentes ao formato, ao meio, ao ambiente e ao tipo de texto. Quanto ao formato, há textos contínuos, não-contínuos, combinados e múltiplos. Esses formatos se referem à forma como o texto é organizado e apresentado ao leitor. Textos contínuos são aqueles organizados em frases e parágrafos, cujas informações apresentadas seguem essa linearidade, diferentemente de um texto não-contínuo, em que as informações são apresentadas de distintas formas (com o uso de imagens, gráficos, tabelas etc). Os textos combinados, como o nome sugere, combinam informações contínuas e não-contínuas. Já os textos múltiplos<sup>49</sup> são assim categorizados quando dois textos diferentes são justapostos, com o intuito de que possam ser comparados ou de que as informações apresentadas possam ser complementadas.

No que diz respeito ao meio, o texto pode ser apresentado de forma impressa ou digital. Os textos apresentados em meio digital podem, ainda, ser classificados quanto ao ambiente, isto é, os textos podem ser de autoria, quando o texto lido não pode ser modificado pelo leitor, ou baseado em mensagem, quando ocorre uma interação com o texto, podendo o leitor acrescentar ou alterar o conteúdo. Por fim, tem-se, também, a classificação quanto ao tipo de texto, quais sejam: descritivos, narrativos, expositivos, argumentativos, prescritivos ou instrutivos e interativos. Tais especificações são relevantes, porque, cognitivamente, os processos requeridos são distintos para cada formato em que um texto é apresentado. Todos esses fatores são considerados pelos avaliadores do PISA no momento de desenvolver a avaliação e de apresentar os textos que serão lidos.

Além de circunscrever as características que orientam os textos a serem apresentados na avaliação, o Programa apresenta, também, as ações que o leitor pode realizar diante do texto

---

<sup>49</sup> O formato denominado “textos múltiplos” sai de cena nas especificações trazidas na última realização do PISA em 2018.

lido, são os chamados processos cognitivos, que, até a versão aplicada em 2015, eram denominados de aspectos. De todo modo, são os diferentes processos cognitivos realizados para conseguir interagir com o texto:

1. Localizar e recuperar informações: incidem sobre a superfície textual e têm por objetivo selecionar a informação requerida e encontrar o local do texto em que tal informação está explicitada. Isso significa que as ações envolvidas nesse aspecto estão voltadas para partes específicas do texto e podem se pautar em processos de nível micro e macroestrutural.
2. Integrar e interpretar: incidem sobre diferentes partes de um texto, envolvendo uma compreensão mais completa a partir da conexão dessas diferentes partes e da construção dos significados que se estendem para além do texto e que é proporcionada pela conexão estabelecida. O leitor, a partir desses aspectos, estabelece a coerência para o texto, construindo sentidos a partir daquilo que não está explícito. Isso significa que, nesse aspecto, ocorrem elaborações e, para tanto, entram em cena as relações entre base textual e conhecimentos prévios do leitor.
3. Refletir e analisar: incidem sobre a atuação do leitor diante do texto, ou seja, o leitor usa suas próprias experiências para fazer comparações, contrastes, críticas e traçar hipóteses, fazendo elaborações para além do texto.

Todos esses aspectos guiam a elaboração da avaliação feita pelo PISA e, também, a seleção dos textos a serem abordados. Os itens que compõem o teste abrangem todos esses processos cognitivos descritos. O teste especifica, também, as escalas de letramento em leitura<sup>50</sup>, divididas em sete níveis. Na avaliação realizada em 2015, o nível mais alto foi 6 e o mais baixo 1b. Esses níveis dão a dimensão das diferentes ações requeridas aos leitores diante do texto. Assim, a questão é elaborada levando em conta esse nível e o que se espera que o leitor seja capaz de fazer para atingir a proficiência requerida em leitura. A seguir são apresentadas as características do que as tarefas relativas a cada nível utilizado no teste aplicado nesta pesquisa buscam avaliar:

Quadro 2: Descrição e percentual de estudantes nos sete níveis de proficiência em leitura – PISA 2015

---

<sup>50</sup> Para o PISA, o letramento em leitura é entendido como a capacidade de compreender, usar, interagir e se envolver com os textos escritos a fim de se atingir objetivo previamente estabelecido.

Nível	Características da Tarefa
6	Nesse nível, as tarefas normalmente requerem que o leitor faça múltiplas inferências, comparações e contrastes com precisão e detalhamento, e que demonstre a compreensão completa e detalhada de um ou mais textos, podendo envolver a integração de informações de mais de um texto. Podem exigir que o leitor lide com ideias desconhecidas, na presença de informações concorrentes relevantes, e produza categorias abstratas para interpretação. Tarefas de refletir e analisar podem solicitar que o leitor levante hipóteses sobre ou avalie criticamente um texto complexo sobre um assunto desconhecido, levando em consideração critérios ou perspectivas múltiplos e aplicando interpretações sofisticadas externas ao texto. Uma condição marcante para tarefas de localizar e recuperar nesse nível é a precisão da análise e a atenção refinada a detalhes pouco perceptíveis nos textos.
5	Nesse nível, tarefas de recuperação de informação requerem que o leitor localize e organize informações profundamente integradas, inferindo sobre quais informações no texto são relevantes. Tarefas de refletir pedem avaliação crítica ou levantamento de hipóteses, com base em conhecimento especializado. Tanto tarefas interpretativas como reflexivas exigem uma compreensão total e detalhada de texto com conteúdo ou forma não familiar. Para todos os aspectos da leitura, as tarefas nesse nível normalmente envolvem lidar com conceitos contrários às expectativas.
4	Nesse nível, tarefas de recuperação de informação requerem que o leitor localize e organize diversos fragmentos de informação integrada. Algumas tarefas nesse nível exigem interpretação do significado de nuances da linguagem em uma parte do texto, levando em consideração o texto como um todo. Outras tarefas interpretativas nesse nível exigem que o leitor use conhecimento público ou formal para levantar hipóteses ou analisar criticamente um texto. O leitor deve demonstrar uma compreensão precisa de textos longos ou complexos cujo conteúdo ou forma podem não ser conhecidos.
3	Nesse nível, as tarefas requerem que o leitor localize e, em alguns casos, reconheça a relação entre vários fragmentos de informação que devem satisfazer múltiplas condições. Tarefas interpretativas exigem que o leitor integre várias partes do texto a fim de identificar a ideia principal, entender a relação ou construir o significado de uma palavra ou oração. O leitor deve considerar muitas características textuais ao fazer comparações, diferenciações e categorizações. Em geral, a informação exigida não é relevante, há muita informação concorrente ou o texto apresenta outros obstáculos, tais como ideias contrárias à expectativa ou formuladas de maneira negativa. Tarefas reflexivas nesse nível podem solicitar correlações, comparações e explicações ou exigir que o leitor avalie uma característica do texto. Algumas exigem que o leitor demonstre uma compreensão refinada do texto em relação a conhecimentos do cotidiano. Outras tarefas não requerem uma compreensão detalhada do texto, mas pedem que o leitor explore um conhecimento menos comum.

Fonte: BRASIL, 2016, p. 101 e 102.

O nível da compreensão em leitura pode ser avaliado, ainda, a partir de diferentes formatos de perguntas (OAKHILL; CAIN; ELBRO, 2017). O PISA desenvolve itens que requerem respostas de múltipla escolha simples, respostas de múltipla escolha complexa e respostas construídas ou abertas. No primeiro formato, o participante deve selecionar uma alternativa dentre as opções oferecidas. No segundo, pode haver variação do tipo de questão; é possível contemplar respostas de verdadeiro/falso, de sim/não, de completar, de relacionar colunas dentre outros. No terceiro, as respostas são formuladas pelos participantes. Os diferentes formatos fornecem dados distintos e possuem suas vantagens e desvantagens (cf. OAKHILL; CAIN; ELBRO, 2017); por isso, sua combinação em um teste é interessante, para que seja possível avaliar o nível de compreensão do participante sob diferentes contornos.

Após conhecer as especificações que guiam a prova desenvolvida pelo PISA e que importam para esta pesquisa, é possível compreender o quão rigorosa é a condução da elaboração da prova. Por conta disso, optamos, neste trabalho, por adotar os critérios desse

exame específico para a elaboração do teste de compreensão em leitura. É necessário enfatizar que a avaliação aqui elaborada se distancia da do PISA por alguns fatores. O primeiro deles diz respeito ao público investigado. Como já apresentado, trata-se de participantes que estão cursando o ensino superior. São, portanto, adultos e com nível de escolaridade diferente dos estudantes avaliados pelo PISA, que são da educação básica, com idade de 15 a 16 anos. Isso é fundamental de ser levado em conta no momento de selecionar o texto e de desenvolver os itens que compõem o teste. Além disso, o objetivo do teste realizado na presente pesquisa, para além de alcançar os objetivos citados anteriormente (avaliar a compreensão, o comportamento e o desempenho em leitura), olha também para o foco específico do presente estudo – a anáfora correferencial realizada com o sintagma “o mesmo”.

A avaliação desenvolvida pelo PISA trabalha, ainda, com diferentes gêneros, envolvendo distintas situações de leitura. Nesta pesquisa, o trabalho se pauta em um texto específico, com foco no processamento em leitura de texto acadêmico-científico. Isso significa que os critérios e as orientações usadas para a avaliação realizada pelo PISA deram subsídios para a construção do teste aplicado, mas foram consideradas as notadas diferenças e as especificidades que um estudo experimental exige. Com o detalhamento da formulação do instrumento, será possível vislumbrar essas particularidades.

No caso da presente pesquisa, dentre as variáveis abordadas pelo PISA, a situação considerada é a educacional. Isso significa que a leitura tem por objetivo a aprendizagem acerca de um determinado conteúdo e o texto não é escolhido pelo leitor. O texto selecionado pode ser considerado acadêmico-científico, sendo mais adequadamente categorizado dentro do eixo da exposição e da argumentação. Ou seja, é um texto próprio da academia, usado para obter informações científicas acerca de um determinado aspecto estudado na área. Além disso, é um texto contínuo, e sua apresentação se deu no meio digital. A leitura foi feita na tela do computador, já que é o modo de leitura mais comum desse tipo de texto. É, ainda, considerado um texto de autoria, pois o leitor não o modifica. A seleção do texto, sua adaptação e sua constituição serão mais bem detalhadas na seção 3.4. Neste momento, atemo-nos à construção do teste em si, explicando os itens que o compõem e os procedimentos de aplicação.

Pois bem, o teste foi composto por dez itens (Apêndice F), que abrangem diferentes níveis e processos cognitivos definidos pelo PISA. Dentre essas questões, algumas requeriam a compreensão das anáforas-alvo para serem resolvidas. Nesse sentido, as questões 2, 4, 6, 8 e 9, além de compor a pontuação geral do teste de compreensão, foram, no momento da tabulação dos dados, desmembradas para que fosse possível avaliar os pontos específicos de compreensão das anáforas.

O formato das questões foi variado. São oito questões abertas, uma de múltipla escolha simples e uma de múltipla escolha complexa. No que diz respeito aos níveis das questões, partiu-se do nível 3 por conta da demanda de leitura requerida pelo texto em si. Isto é, levando em conta o texto com o qual se está trabalhando, não há como os itens estarem abaixo desse nível, pois os níveis 1b, 1a e 2 pressupõem o trabalho com textos mais curtos e simples, que tratem de assuntos conhecidos, e tenham uma dificuldade moderada. Foram elaboradas três questões de nível 3, duas questões de nível 4, duas questões de nível 5 e três questões de nível 6 (Apêndice F). Quanto aos processos cognitivos, aplicaram-se quatro itens de localizar e recuperar, três itens de integrar e interpretar e três itens de refletir e analisar.

Quadro 3: identificação do nível e do aspecto de cada questão

<b>Aspecto</b>	<b>Nível</b>	<b>Questão</b>
Refletir e Analisar	3	Questão 3
	5	Questão 7
	6	Questão 1
Integrar e Interpretar	3	Questão 6
	6	Questão 10
	6	Questão 5
Recuperar e Localizar	3	Questão 4
	4	Questão 8
	4	Questão 9
	5	Questão 2

Fonte: a autora.

O número de questões foi estipulado levando em conta os resultados do piloto realizado, que era composto de treze itens. Percebeu-se, naquela ocasião, que os acadêmicos estavam levando muito tempo para realizar o teste. Um tempo demasiadamente longo para a realização de um teste pode levar os níveis de atenção do participante a decaírem, comprometendo os resultados. Além disso, há risco de exaustão e desistência da tarefa. Desse modo, optou-se por realizar um teste mais curto do que aquele aplicado no piloto.

O teste foi elaborado de forma que fosse possível obter dados que alcançassem o nível inferencial de compreensão (GAGNÉ; YEKOVICH; YEKOVICH, 1993; KINTSCH; RAWSON, 2013), ressaltando que a intenção está, sobretudo, em investigar os processos de leitura envolvidos na compreensão de texto acadêmico-científico em que algumas das retomadas, importantes recursos de integração textual, são construídas com o sintagma “o mesmo”. Assim, o foco não está em analisar a competência em leitura de modo geral, mas sim ao que compete ao texto acadêmico-científico, perpassando pelo objeto investigado.

Ainda, no que diz respeito à construção dos itens, as questões discursivas pretendem que o participante explique pontos tratados no texto, integrando diferentes partes e relacionando as informações obtidas com o conhecimento prévio disponível. Os participantes foram orientados a construir suas respostas, sem realizar cópias explícitas do texto. Na questão de múltipla escolha simples, foram oferecidas quatro alternativas, com distratores, que poderiam responder parcialmente ao solicitado, mas que não se caracterizam como corretos. Na questão de múltipla escolha complexa, solicitou-se que os participantes analisassem afirmativas e respondessem se eram verdadeiras ou falsas de acordo com o texto. O enunciado de cada questão foi elaborado de modo a deixar claro aquilo que se requeria do participante. Por conta disso, embora não seja recomendado que o objeto testado esteja presente nos itens elaborados, optou-se por manter as anáforas nos enunciados de algumas das questões, ainda que a anáfora central de análise, realizada com “o mesmo”, não tenha aparecido nenhuma vez.

Após a etapa de desenvolvimento das questões, elas foram enviadas a quatro avaliadores especialistas: a uma doutoranda da área da linguística, a um doutorando da área de Engenharia de Controle e Automação, a um professor também da área de Engenharia de Controle e Automação e a uma engenheira civil. As avaliações recaíram, principalmente, sobre três pontos: questões que poderiam gerar dúvidas nos participantes, enunciados que não estavam muito claros e sobre a adequação do conteúdo do teste para o público-alvo. Esses pontos foram considerados e revistos para que, então, as questões fossem pilotadas – conforme será apresentado na última seção deste capítulo – com acadêmicos (diferentes daqueles que participaram da coleta principal) do curso de Engenharia de Controle e Automação.

A partir do exposto, é possível considerar que o instrumento desenvolvido para o teste de compreensão em leitura respeita parâmetros e critérios requeridos a um instrumento de avaliação confiável e válido (CARVALHO, 2022), conforme *checklist* a seguir, com base em Carvalho (2022, p.184-190):

Quadro 4: Parâmetros e Critérios seguidos no Teste de Compreensão em Leitura

PARÂMETRO	CRITÉRIO	
PLANEJAMENTO DO INSTRUMENTO		
<b>Definição dos fundamentos epistemológico-teóricos subjacentes</b>	Domínio das teorias subjacentes aos processos de compreensão leitora.	OK
<b>Discriminação dos procedimentos teóricos</b>	Definição constitutiva do construto Distinção e conceituação das dimensões (processos cognitivos) do construto Determinação e conceituação da(s) habilidade(s) a ser(em) medida(s)	OK
<b>Estabelecimento do foco da avaliação</b>	Definição e caracterização do público-alvo	OK



	Identificação das variáveis	OK
	Determinação dos objetivos da avaliação	OK
	Definição da abordagem	OK
<b>Indicação do método da avaliação</b>	Definição do instrumento, técnicas e procedimentos	OK
	Justificativa das escolhas à luz dos objetivos e dos fundamentos teóricos	OK
	Utilização de mais de um instrumento para a avaliação	OK
	Calibragem do(s) instrumentos(s)	OK
<b>DESENVOLVIMENTO DO INSTRUMENTO</b>		
<b>Seleção e adaptação do(s) texto(s)-base<sup>51</sup></b>	Definição do público-alvo	OK
	Definição da estrutura e conteúdo textual	OK
	Ponderação sobre a constituição do texto	OK
	Priorização de tópico de interesse do público-alvo	OK
	Ajustes e convalidação dos textos	OK
<b>Consideração de critérios de adequação dos itens elaborados</b>	Uso de formatos variados – múltipla-escolha simples, complexa e resposta construída e suas variantes.	OK
	Definição de escala de níveis de proficiência.	OK
	Definição do nível de proficiência a ser medido em cada item.	OK
	Equilíbrio na quantidade de formatos, processos e níveis medidos.	Parcialmente
	Verificação do grau de amplitude, equilíbrio e homogeneidade entre os itens.	OK
<b>Garantia da justiça no desenho do teste</b>	Garantia de validade das interpretações da pontuação do teste para o (s) uso (s) pretendido (s) para indivíduos de todos os subgrupos relevantes	OK
<b>Detalhamento das características do instrumento</b>	Listagem e ordenamento das especificações do teste	OK
<b>Existência de rubrica para o teste</b>	Descrição detalhada do que constitui níveis aceitáveis ou inaceitáveis de desempenho para cada um dos itens	OK
<b>VALIDAÇÃO DO INSTRUMENTO</b>		
<b>Busca de evidências de validade</b>	Busca de evidências baseadas no conteúdo	OK
	Busca de evidências baseadas no processo de resposta (mediante modelos explicativos do processamento mental que ocorre durante a resposta aos itens)	OK
	Busca de evidências baseadas na estrutura interna do teste	X
	Busca de evidências baseadas nas relações com variáveis externas	OK

<sup>51</sup> Sobre o texto falaremos de modo mais retido na seção 3.4, mas optamos por já abordar os parâmetros e critérios neste ponto, já que eles se vinculam à elaboração do teste de compreensão.

	Busca de evidências baseadas nas consequências de testagem	OK
--	--	----

Fonte: a autora com base em Carvalho (2022, p.184-190).

Alguns pontos cabem ser esclarecidos no que diz respeito ao *checklist* formulado. O critério “equilíbrio na quantidade de formatos, processos e níveis medidos” foi seguido parcialmente durante a formulação do teste de compreensão em leitura. Tendo em vista o alvo do estudo (a anáfora correferencial) e o tempo destinado ao teste, não foi possível ampliar o número de questões a fim de que houvesse paridade nos formatos, processos e níveis. Ainda assim, buscamos pelo equilíbrio, construindo um teste que atingisse diferentes processos e distintos níveis sendo avaliados a partir de diferentes formatos de questões. Além disso, o critério “busca de evidências baseadas na estrutura interna do teste” não foi seguido. Tal critério poderia ser aplicado após a realização do piloto do teste, no entanto, a amostra que obtivemos no teste piloto não nos permitiu segui-lo. Dessa forma, a análise da estrutura interna do teste foi feita de modo qualitativo.

#### 3.3.4.1 Procedimento de coleta de dados

A aplicação do teste de compreensão em leitura ocorreu em um segundo encontro com os participantes, quando ocorreu, também, o teste de aceitabilidade, no laboratório de informática do Centro de Comunicação e Expressão (CCE) da Universidade Federal de Santa Catarina (UFSC). O teste foi realizado em computador e de modo individual, embora a sessão fosse coletiva. O tempo disponibilizado foi de 60 minutos<sup>52</sup> mais o tempo usado para a instrução. Assim, a leitura das instruções (Apêndice E) foi feita em conjunto e cada participante pôde acompanhá-la em um arquivo disponível em sua máquina em *Portable Document Format* (PDF). Instruiu-se, ainda, que, após o início do teste, não seria permitido qualquer tipo de consulta extra ou de conversa, tanto entre os participantes quanto com a pesquisadora. Na sequência, os participantes deveriam ler o texto, que já era de seu conhecimento, pois foi o mesmo texto empregado na tarefa de protocolo verbal. A intenção em aplicar o mesmo texto foi a de poder contrastar os dados obtidos por meio dos protocolos com os dados obtidos com o teste de compreensão em leitura.

---

<sup>52</sup> A pilotagem indicou que este tempo seria suficiente, todavia, o tempo não foi um limitador, pois os participantes não foram interrompidos durante a testagem.

Como o teste de leitura se trata de uma medida que capta dados pós-processamento, não se entende como um problema o fato de os acadêmicos já terem lido o texto anteriormente, até porque durante o teste de compreensão, eles poderiam retomar o texto quantas vezes necessitassem para responder às questões. O texto foi disponibilizado em um arquivo *Portable Document Format* (PDF) diferente daquele onde estavam as questões. Após a leitura do texto, as questões deveriam ser respondidas, sendo que os participantes poderiam consultar o material para formular as suas respostas. O teste foi apresentado em um arquivo em formato PDF editável (PDFelement<sup>53</sup>). Dessa forma, os participantes puderam inserir suas respostas no próprio arquivo PDF.

Além dos dados *off-line* obtidos, utilizou-se, também, um software para gravar os movimentos realizados pelos participantes no momento da realização do teste. Essa gravação foi feita com o software *VLC Media Player*, reprodutor multimídia de código aberto, que comporta a funcionalidade de captura dinâmica de tela. A utilização desse software permite, conforme Souza, Franzen e Schlichting (2019, p. 3852), “conhecer o caminho que o participante fez (retomadas, saltos, marcações, anotações etc.) para chegar ao resultado final: a resposta à questão.”. Assim, são os resultados de todas essas ações que compõem os dados do teste de compreensão em leitura. Ainda, com o teste, tem-se acesso às seguintes variáveis dependentes: desempenho de compreensão em leitura e comportamento leitor.

Após conhecer todos os instrumentos que constituem a pesquisa desenvolvida, na seção seguinte, será feita a apresentação do texto selecionado para a realização dos protocolos verbais e do teste de compreensão, especificando tanto aspectos relativos à composição técnica (como número de palavras e linhas) quanto aspectos linguísticos.

### 3.4 TEXTO UTILIZADO NOS PROCEDIMENTOS: PROTOCOLO VERBAL E TESTE DE COMPREENSÃO EM LEITURA

Tanto a tarefa de protocolo verbal quanto o teste de compreensão em leitura contaram com um mesmo texto a ser lido. A opção por trabalhar com um mesmo texto foi para que fosse possível comparar e discutir os dados de compreensão como processo e como produto da leitura. Trabalhamos com um texto autêntico, isto é, um texto que estava publicado e que poderia fazer parte do ementário de leituras dos acadêmicos. Trata-se, dessa forma, de um texto acadêmico-científico, pertencente ao eixo da exposição e da argumentação. Esse tipo de texto

---

<sup>53</sup> Esse programa se encontra disponível no site: <<https://pdf.wondershare.com/pt-br/>>.

tem como propósito informar sobre um determinado conhecimento, em que estão incluídos termos e construtos específicos acerca de um determinado campo. No ensino superior, o contato com esse tipo de material é bastante frequente, visto que são usados para obtenção de conhecimentos em torno da área e isso é comprovado quando os participantes respondem ao questionário, informando o que leem em seu curso de graduação.

Os textos expositivos, conforme Moss et al (2011), costumam exigir a realização de mais inferências, tanto implícitas quanto explícitas, por parte do leitor, o que pode significar um custo maior para o seu processamento. Assim, é importante que o leitor tenha algum conhecimento prévio acerca da temática do texto. Por isso, o texto lido pelos participantes era da área de Engenharia de Controle e Automação. Ainda, de acordo com Alonso e Mateos (1985), uma das diferenças entre textos narrativos e expositivos está no fato de que a compreensão da prosa narrativa depende mais dos processos descendentes (*top-down*), enquanto a prosa expositiva depende mais dos processos que ocorrem de modo *bottom-up*. Isso não significa que os processos *top-down* não sejam requeridos. Eles o são, tanto que a seleção dos participantes considerou público-alvo específico para leitura do texto, mas, de acordo com os pesquisadores, na leitura de textos da ordem do expor, são exigidos mais processos descendentes do que em textos narrativos. Isso é relevante para o nosso estudo, tendo em vista que a compreensão e a adequada resolução da anáfora correferencial, que depende essencialmente do cotexto, estão sujeitas, portanto, a tais processos.

Entendidas essas questões, passamos para a explicação das etapas de seleção do texto propriamente dito. Inicialmente, optamos por buscar vários textos, selecionados a partir da matriz curricular do curso de Engenharia de Controle e Automação da UFSC. Tendo em vista as temáticas tratadas nas disciplinas mais específicas ao curso, buscamos por textos que, em alguma medida, contemplassem essas temáticas. Além disso, tomamos como ponto de partida o vídeo institucional do curso, no qual os acadêmicos apresentam temáticas relevantes à área. Com esses temas em mãos, buscamos em bases digitais, como o Portal de Periódicos Capes, em sites de congressos da área da engenharia (como o Congresso Brasileiro de Educação em Engenharia - COBENGE) e em livros disponíveis digitalmente, que contassem com participação de profissionais da área, sempre textos em língua portuguesa.

Dessa busca, foram selecionados seis textos. Após a primeira seleção, submetemos os textos à avaliação de três professores do curso de Engenharia de Controle e Automação. Essa avaliação foi importante para que fosse possível compreender se, de fato, os materiais abrangiam temáticas relevantes ao curso e acessíveis aos estudantes. O texto a ser selecionado não poderia ser muito simples aos acadêmicos, de modo que seus processos não fossem

desautomatizados. Não poderia, também, ser um texto com um nível de dificuldade muito elevado, pois a compreensão poderia ficar comprometida por falta de conhecimento prévio dos participantes. Tendo em vista esses aspectos, os professores deram suas sugestões acerca dos textos e, então, foi selecionado aquele intitulado “Fundamentos de detecção de defeitos em sistemas distribuídos”, com autoria de Raimundo José de Araújo Macêdo e Alirio Santos de Sá (Apêndice G), publicado no livro “Projeto de Sistemas Distribuídos e de Tempo Real para Automação”. Trata-se de um livro escrito por diferentes profissionais, dentre os quais professores da área de Engenharia de Controle e Automação da UFSC.

Tendo em vista o objetivo desta tese e o tempo para realização das duas tarefas em que o texto seria utilizado, não era possível trabalhar com o artigo todo. Além disso, era necessário fazer as adaptações e os controles experimentais. Assim, após realizar esses ajustes e controles, o texto foi submetido a cinco avaliadores especialistas: uma doutoranda da área da Linguística, um doutorando da área da Engenharia de Controle e Automação e três professores do curso de Engenharia de Controle e Automação, um deles autor de artigo constante no mesmo livro do qual o texto usado na pesquisa foi retirado. Além disso, foi realizada a pilotagem com estudantes semelhantes à população participante da pesquisa. Tudo isso foi importante para compreender se o texto estava claro e se as alterações realizadas não haviam comprometido seu conteúdo ou qualquer informação teórica, o que é essencial, sobretudo, para a realização de um teste de compreensão em leitura (CARVALHO, 2022).

Existe uma complexidade muito maior entre as relações anafóricas em um texto. Isso precisou ser considerado na escolha e na adaptação do material a ser utilizado. Não seria possível, por exemplo, inserir as anáforas correferenciais em posições muito próximas de seus antecedentes e nem a distâncias muito longas. Além disso, não seria possível construir as anáforas apenas com as variáveis investigadas, pois essa ação poderia tornar o texto muito artificial, além de colocar em evidência o elemento sob análise e esse não era o intuito. Embora seja um material controlado, não é esperado que o texto beire à artificialidade. Veremos que, no material, há a ocorrência de diferentes anáforas, mas as que nos interessam são todas anáforas que podem ser compreendidas como correferenciais. Então, a partir do artigo selecionado, foram feitas adaptações, considerando, sobretudo, introdução e conclusões a fim de se manter o foco temático do artigo selecionado. A partir disso, chegamos a um texto contínuo com cinquenta e três linhas e seiscentas e cinquenta e cinco palavras. O texto está, ainda, dividido em cinco parágrafos e contém diferentes formas de anáforas (Apêndice G).

Na sequência, analisamos cada parágrafo com o intuito de contextualizar as diferentes anáforas construídas e perceber como a progressão do texto acontece e leva, a partir desse recurso, à construção da compreensão:

Os sistemas de automação requerem a capacidade de tolerar falhas, já que defeitos ou mau funcionamento podem implicar prejuízos financeiros, ambientais ou, até mesmo, perda de vidas humanas. Por melhores que sejam as técnicas usadas na construção desses sistemas, diversas questões podem levá-los a apresentar defeitos: desgastes físicos de componentes, falhas humanas, fenômenos e/ou desastres naturais etc. Portanto, os mesmos devem possuir mecanismos adequados de tolerância a falhas. Os termos “falha”, “erro” e “defeito”, utilizados neste texto, seguem o padrão de terminologia amplamente adotado pela comunidade científica internacional da área de tolerância a falhas.

O primeiro parágrafo é iniciado com a introdução do referente pelo sintagma “os sistemas de automação”. Esse referente será retomado na sequência por uma anáfora realizada com um determinante e com parte do sintagma do antecedente; assim, ao ler “desses sistemas”, sabemos que se trata dos sistemas de automação e não de qualquer outro sistema. Em seguida, é usado o pronome em “levá-los”, uma anáfora direta. Há, ainda, nesse parágrafo, uma retomada anafórica alvo realizada com “os mesmos”. Tal anáfora está na posição de sujeito e retoma o tópico do parágrafo, que está em posição de sujeito da sentença também. Como vimos, esse antecedente já havia sido retomado outras duas vezes nesse mesmo parágrafo, depois dessas duas retomadas, está a variável de interesse desta pesquisa. Nesse mesmo primeiro parágrafo, são trazidas, ainda, informações sobre “tolerar falhas”, que serve de âncora para a nominalização apresentada na sequência “tolerância de falhas” e para “defeitos”, que pode ser considerado uma anáfora associativa e que retoma o título do texto. Esses elementos ancoram o referente introduzido no segundo parágrafo “a detecção de defeitos”:

A detecção de defeitos é um serviço utilizado para construir mecanismos de tolerância a falhas. Ela consiste na verificação do funcionamento dos dispositivos ou componentes, de modo que, quando não se comportam de acordo com objetivos especificados, ações possam ser tomadas para substituição de componentes defeituosos, recuperação ou reconfiguração do sistema. O comportamento dos componentes em caso de falhas guia a forma de implementação dos detectores de defeitos. O mesmo recebe a denominação de semântica de falhas, que determina a técnica de verificação (ou monitoramento) a ser utilizada pelo serviço de detecção de defeitos. Por exemplo, se um dispositivo do sistema pode falhar produzindo valores incorretos, testes podem ser usados para comparar as saídas do dispositivo com saídas de referência. De outro lado, se um dispositivo falha somente omitindo valores de saída (temporária ou permanentemente), temporizadores podem ser usados para detectar defeitos.

Nesse segundo parágrafo, há duas retomadas que são alvo. A primeira é realizada com o pronome “ela” e retoma o sintagma “a detecção de defeitos”. Tanto anáfora quanto antecedente estão na posição de sujeito. Esse pronome abre a sentença em que outros dois referentes são introduzidos: “dispositivos ou componentes”, que são retomados em seguida por um pronome reflexivo “se” e ancoram o referente que vem na sequência, “o comportamento dos componentes em caso de falhas”. Embora seja um referente extenso, constitui-se de apenas um sintagma, que é retomado pelo sintagma de interesse desta pesquisa, “o mesmo”. Nessa oração, é introduzido, também, o referente “semântica de falhas”, retomado por um pronome relativo. Em seguida, dentro do frame “sistema”, está ancorado o referente “um dispositivo do sistema”, que é retomado pela anáfora “do dispositivo” e, na sequência, por “um dispositivo”.

No terceiro parágrafo, não foram inseridas anáforas com as variáveis experimentais.

A natureza descentralizada de sistemas de automação modernos tem impacto no projeto de detectores de defeitos. Em alguns desses sistemas, os dispositivos de automação estão distribuídos e interconectados a partir de redes de comunicação, por exemplo: Sistemas de Controle Sobre Rede (NCS, *Networked Control Systems*); Sistemas de Controle Distribuídos (DCS, *Distributed Control Systems*), aplicações distribuídas sobre Redes Veiculares Ad Hoc (VANET, *Vehicular Ad hoc Networks*), Sistemas Ciber-Físicos (CPS, *Cyber-Physical Systems*)<sup>[6]</sup>, entre outros. A natureza descentralizada traz uma série de benefícios, tais como a possibilidade da supervisão remota, a interconexão, a adição e a remoção de novos dispositivos, a facilidade de integração entre sistemas etc. Todavia, nesses ambientes de automação distribuídos, as propriedades e o desempenho do serviço de detecção de defeitos dependem das garantias de qualidade de serviço (QoS, *Quality of Service*) entregues pelos sistemas de processamento e de comunicação utilizados.

Aqui, é trazido um novo referente, que, por certo, está ancorado aos referentes já introduzidos anteriormente, “a natureza descentralizada de sistemas de automação modernos”. Tal sintagma é quebrado e retomado por partes, com a nominalização precedida do demonstrativo “desses sistemas”, que se refere aos sistemas de automação de natureza descentralizada. Na sequência, “a natureza descentralizada” é desmembrada do sintagma inicial e repetida para que os seus benefícios sejam apresentados. Nesse ponto, o referente “benefícios” é trazido e, em seguida, retomado por “tais”. Por fim, o sintagma “nesses ambientes de automação distribuídos” é uma anáfora indireta, que retoma o referente introduzido no início do parágrafo e, também, a informação já dada sobre o fato de alguns sistemas de automação estarem “distribuídos<sup>54</sup> e interconectados”. Para fechar o parágrafo, é reativado “o serviço de detecção de defeitos”, que abrirá o parágrafo seguinte.

---

<sup>54</sup> Sistemas distribuídos e sistemas descentralizados podem ser considerados sinônimos nesse contexto.

O quarto parágrafo é, então, aberto pelo sintagma “o serviço de detecção de defeitos”, que, na oração seguinte, é retomado pelo sintagma “o mesmo”, ambos em posição de sujeito.

O serviço de detecção de defeitos, nesse sentido, pode não ser confiável em alguns ambientes. O mesmo pode acabar fornecendo informações incorretas ou apresentando altos tempos de respostas na detecção de defeitos de componentes do sistema. Isso acontece se a infraestrutura de computação subjacente possuir períodos em que não há garantias com relação aos tempos de processamento e transmissão das mensagens, o que gera a possibilidade de comprometimento no funcionamento de certos sistemas (ou subsistemas) de automação com requisitos temporais. Portanto, as hipóteses do modelo de sistema distribuído - isto é, o comportamento temporal e o comportamento de falhas dos componentes de *software*, dos dispositivos e dos canais de comunicação - são usadas para determinar as propriedades do serviço de detecção de defeitos durante sua operação. Essas propriedades ou garantias influenciam, por sua vez, as características dos protocolos usados na construção de sistemas distribuídos robustos – como, por exemplo, consenso distribuído e difusão atômica; duas técnicas usadas na construção de estratégias de replicação para tolerância a falhas.

Nesse parágrafo, há, ainda, uma anáfora encapsuladora, realizada com o demonstrativo “isso”, que retoma toda uma ideia expressa na proposição anterior. Esse encapsulamento é realizado, também, pela expressão “o que”. Depois, o texto introduz um novo referente, mas que parte, também, de outros referentes já introduzidos: “as hipóteses do modelo de sistemas distribuído” que levam o leitor ao sintagma “as propriedades do serviço de detecção de defeitos”, novo referente, mas ancorado a informações e referentes já ativados. Esse referente é retomado pelo pronome “sua” e, em seguida, pelo demonstrativo seguido de “propriedades ou garantias”. Ao final do parágrafo, é trazido o referente “sistemas distribuídos robustos” e é feita a retomada de “tolerância de falhas”.

No quinto e último parágrafo, há, ainda, duas anáforas realizadas com as variáveis experimentais.

O desempenho de um detector de defeitos é definido em termos de sua acurácia e rapidez. Ele é adequado se os seus parâmetros operacionais forem configurados de forma a atender as demandas das aplicações, mediante as condições de carga e disponibilidade dos recursos computacionais do ambiente distribuído. Detectores de defeitos consomem por si recursos computacionais para o processamento e a transmissão das mensagens de monitoramento e requerem parâmetros operacionais adequados para o seu funcionamento; por isso, eles podem degradar o desempenho da aplicação ou do próprio serviço de detecção, fornecendo informações incorretas ou tempos de detecção inadequados, se os intervalos entre verificações dos defeitos não forem definidos de forma apropriada.



A primeira anáfora possui como antecedente “o desempenho de um detector de defeitos”, que introduz o parágrafo, está em posição de sujeito e é retomado pelo pronome “ele”, também em posição de sujeito. Na sequência, parte do sintagma que introduz o parágrafo, “detector de defeitos”, é retomado com o pronome “seus”. Depois, ocorre a repetição de “detectores de defeitos”, em posição de sujeito, e esse sintagma é retomado pelo pronome “eles”, também em posição de sujeito. Todavia, antes dessa anáfora, o antecedente é retomado outras três vezes: primeiro com a anáfora reflexiva “por si”, depois com o nulo (“e  $\emptyset$  requerem”) e, na sequência, com o pronome possessivo “seu”.

Ao longo do texto, a distância entre o antecedente e a anáfora alvo não foi controlada, mas procurou-se manter uma distância que favorecesse a naturalidade do uso de pronomes como anafórico, os quais, de acordo com Kintsch e Rawson (2013, p. 233), “[...]costumam ser usados em referência a conceitos que foram mencionados ou abordados recentemente, que foram introduzidos explicitamente no texto e que estão ativados na memória de trabalho.”. Assim, quando o antecedente estava mais distante, outras retomadas foram inseridas com o objetivo de mantê-lo ativo na memória e de não comprometer a textualidade.

De modo geral, o texto contém referentes ancorados e não ancorados, que são retomados por anáforas construídas de diferentes formas no decorrer da exposição e sobre os quais gira a discussão. São eles: sistemas de automação, tolerar falhas/tolerância de falhas, defeitos ou mau funcionamento, detecção de defeitos/detector de defeitos/detectores de defeitos, serviço, dispositivos ou componentes, o comportamento dos componentes em caso de falhas, semântica de falhas, os dispositivos de automação distribuídos, benefícios, propriedades, consenso distribuído e difusão atômica; e o desempenho de um detector de defeitos.

As anáforas encontradas no texto são dos mais variados tipos. Há anáforas indiretas, associativas, encapsuladoras e correferenciais. É possível perceber como a trama do texto se constrói a partir dos usos dos diferentes tipos de retomada e da introdução de novos referentes, constituindo-se a progressão referencial do texto e articulando as suas diferentes partes de modo coerente, como discutimos no capítulo da fundamentação teórica. Essa construção é tão articulada e amarrada, que, em alguns momentos, torna-se complexo avaliar quais referentes são novos e quais são introduzidos de modo ancorado.

### 3.5 ESTUDO PILOTO

Para testar a qualidade e a adequação dos materiais e das tarefas realizadas na coleta de dados, conduzimos um estudo piloto nos meses de maio e junho de 2019 com acadêmicos do

curso de Engenharia de Controle e Automação que estavam realizando duas disciplinas específicas, a saber: Controle Multivariável e Sistemas de Controle, disciplinas optativas realizadas por acadêmicos que estão nas fases finais do curso. Assim, os participantes tinham o perfil muito semelhante ao público da pesquisa, mas nenhum participante que realizou o estudo piloto fez parte da coleta principal.

Os procedimentos de coleta de dados do estudo piloto foram praticamente os mesmos utilizados na investigação. Dos testes a serem realizados, apenas o experimento de leitura automonitorada não foi pilotado com esse público, mas as frases foram colocadas em testagem no julgamento de aceitabilidade. Após os resultados do piloto, as devidas adaptações foram realizadas.

O piloto foi, então, realizado com sete acadêmicos que participaram de todas as etapas da pesquisa, salvo a leitura automonitorada, que não foi testada nesse momento. O procedimento utilizado para recrutar os participantes foi o mesmo usado na investigação, ou seja, passamos nas salas e eles foram convidados a participar. Os interessados preencheram um questionário e, posteriormente, entramos em contato com cada um deles. O piloto aconteceu em dois momentos. O primeiro durou, em média, 60 minutos para cada participante, e os acadêmicos realizaram o teste de protocolo verbal e o julgamento de aceitabilidade. O segundo momento teve duração de 80 minutos e foi realizado o teste de compreensão em leitura.

Após os resultados fornecidos pelos acadêmicos de Engenharia, os estímulos utilizados no teste de julgamento de aceitabilidade e no teste de leitura automonitorada foram revistos, pois percebemos que era necessário reduzir o conjunto experimental e reformulá-lo. Feito isso, as frases foram enviadas a duas avaliadoras especialistas, doutoras em linguística. A partir das sugestões dadas, os estímulos foram melhorados e, então, o teste de julgamento de aceitabilidade voltou a ser pilotado. Essa fase da pilotagem aconteceu também com universitários em fase final de curso, mas, dessa vez, com acadêmicos do curso de Letras - Português e respectiva literatura da Universidade Federal de Santa Catarina. Além disso, as frases foram julgadas por estudantes do ensino médio técnico do Instituto Federal Catarinense (IFC)<sup>55</sup>. A partir dos resultados obtidos, constatamos que algumas frases experimentais estavam possibilitando leitura ambígua. Por conta disso, foram novamente revistas.

O teste de leitura automonitorada foi pilotado com um público diferente daquele participante da pesquisa. Tratava-se de estudantes do curso de Letras-Português e respectiva

---

<sup>55</sup> Como os experimentos do teste de aceitabilidade e de leitura automonitorada não requerem conhecimentos específicos de área ou conhecimentos prévios de um determinado assunto, foi possível pilotá-los com um público de área distinta daquela com a qual se trabalha nesta pesquisa.

literatura da Universidade Federal de Santa Catarina. Essa pilotagem foi realizada posteriormente, porque primeiro foram analisados os resultados obtidos pelo teste de aceitabilidade do piloto com os acadêmicos de Engenharia, de Letras e dos estudantes do ensino médio. Era importante que a aceitabilidade fosse avaliada previamente (antes de pilotar o teste de leitura automonitorada) para evitar rodar frases com construções não naturais que, por si só, seriam as responsáveis por um tempo de leitura maior.

### **3.5.1 Principais resultados da pilotagem**

Com os resultados obtidos por meio do protocolo verbal, percebemos que o texto tinha extensão adequada e que o tempo de realização da tarefa estava dentro do esperado, não passando de 20 minutos. Já o teste de julgamento de aceitabilidade teve de ser adaptado quanto à extensão, pois os participantes estavam levando mais de 25 minutos para realizá-lo. Esse tempo foi considerado demasiadamente longo. Assim, reduzimos o número de sentenças distratoras. Inicialmente, o teste contava com doze sentenças experimentais e sessenta e três distratoras. Além disso, as frases experimentais acabaram sendo reformuladas, pois, por conta da tarefa de leitura automonitorada, alguns controles foram inseridos, como número de sílabas do antecedente e do segmento pós-crítico (vide seção 3.3.2.1).

Com a realização do piloto, pudemos observar, também, que o teste de compreensão em leitura, com treze questões, estava muito longo. Os participantes levaram, em média, 75 minutos para realizar o teste, quando o esperado era que conseguissem finalizá-lo em 60 minutos para que os níveis de atenção não começassem a decair pelo cansaço inerente ao esforço cognitivo realizado. Além disso, algumas questões precisaram ser reescritas, pois constatamos que, da forma como estavam, não foram bem compreendidas pelos participantes. Outras questões permitiam respostas muito abrangentes, o que precisou ser revisto também. Por fim, algumas questões, que não eram imprescindíveis para o objetivo do teste, foram suprimidas. Dessa forma, o teste de compreensão passou a contar com 10 questões, sendo que algumas delas continham mais de uma solicitação (como as questões 02, 03, 05, 08 e 09).

No que diz respeito à leitura automonitorada, cada segmento apresentado ao participante tinha um tempo específico de exposição em tela. Assim, se o participante não desse o comando dentro do tempo, o próprio programa passava para o segmento seguinte (sem computar o tempo de leitura do participante). O piloto mostrou que era importante ampliar esse tempo para não correr o risco de o tempo de leitura do participante não ser computado, passou-se, então, de 3500 milissegundos para 5000 milissegundos de exposição. Além disso, o tempo de exposição

da pergunta de atenção foi ampliado de 5000 milissegundos para 9000 milissegundos, porque, durante o piloto, alguns participantes não conseguiram ler e responder a pergunta dentro do tempo. Outro ponto reconsiderado com a pilotagem foi a forma de apresentação do estímulo visual, os sintagmas estavam encobertos por *hashtags*, mas, durante a leitura, esse formato se mostrou inadequado, pois acabava competindo com as palavras e gerando um embaralhamento da visão. Por conta disso, as *hashtags* foram substituídas por *underscores*. Ainda, o piloto do teste de leitura automonitorada mostrou que as sentenças estavam compreensíveis e adequadas, confirmando que as alterações realizadas a partir do resultado obtido com a pilotagem do teste de aceitabilidade estavam adequadas.

A pilotagem se mostrou, portanto, fundamental para ajustar os instrumentos, treinar a pesquisadora, testar as instruções, modular a interação, verificar a adequação do espaço e organizar a sequência de aplicação dos instrumentos, minimizando, dessa forma, o risco de falhas no momento da coleta. Em uma pesquisa experimental, é essencial que as técnicas utilizadas estejam bem alinhadas com o propósito de minimizar a influência de fatores intervenientes nos resultados obtidos, de garantir a confiabilidade do estudo e as evidências de validade (CARVALHO, 2022).

### 3.6 PROCEDIMENTOS GERAIS DA COLETA DE DADOS

A coleta de dados ocorreu nos meses de outubro e novembro de 2019. Cada acadêmico precisou disponibilizar dois momentos. A primeira sessão foi individual, em que foram realizados o protocolo verbal e a leitura automonitorada. A primeira tarefa durou em torno de 30 minutos, contabilizando junto desse tempo a assinatura do Termo de Consentimento Livre e Esclarecido (TCLE), as instruções e o treinamento. A leitura automonitorada, por sua vez, durou em torno de 20 minutos, também contando o tempo para as instruções e para o treinamento. Assim, o primeiro encontro durou em torno de 50 minutos, podendo variar um pouco de acordo com o participante.

Já a segunda sessão foi realizada em grupo, em um laboratório de informática, quando foram aplicados os testes de julgamento de aceitabilidade e de compreensão em leitura. O teste de aceitabilidade levou em torno de 15 minutos, somando ao tempo da tarefa o tempo destinado à instrução. O teste de compreensão em leitura, por sua vez, durou em torno de 70 minutos, considerando aplicação do teste e instruções.

A ordem escolhida para a aplicação dos instrumentos não foi aleatória. Os protocolos verbais e o teste de leitura automonitorada ocorreram em um primeiro encontro, de forma

individual e foram marcados de acordo com a disponibilidade de cada participante. Diante disso, ocorreram do dia 29 de outubro de 2019 ao dia 08 de novembro de 2019. Os protocolos deveriam ser realizados primeiro, pois a atenção exigida para a tarefa era maior, já que os participantes precisavam fazer a leitura de um texto e verbalizar os seus pensamentos. Além disso, eram menores as chances de os participantes identificarem qual o objeto do estudo, pois a concentração estava centrada em compreender o texto. Nas frases da leitura automonitorada, seria possível levantar mais hipóteses que poderiam ser confirmadas, caso o texto fosse lido posteriormente. Como a frase é uma unidade menor, acaba por oferecer mais margem para chegar mais perto do que é o alvo. E era fundamental que os participantes não se dessem conta do objeto estudado em si, pois isso pode enviesar os dados obtidos. Assim, o teste de leitura automonitorada foi aplicado na sequência do protocolo verbal. Além disso, era importante que a tarefa de protocolo verbal acontecesse antes do teste de compreensão em leitura, já que ambas as tarefas contavam com o mesmo texto e, para a sessão de protocolo, é fundamental que o participante não conheça o texto lido. Isso aumenta a chance de desautomatização dos processos de compreensão ou de, pelo menos, alguns deles e, ainda, amplia a possibilidade de desautomatização em razão dos desafios que o texto pode apresentar.

A intenção para a segunda sessão era a de que todos os participantes realizassem em conjunto o teste de julgamento de aceitabilidade e o teste de compreensão em leitura; no entanto, devido às diferentes disponibilidades dos participantes, isso não foi possível. Assim, o segundo momento ocorreu no dia 18 de novembro de 2019 com sete participantes, no dia 20 de novembro de 2019, com cinco participantes, no dia 21 de novembro de 2019 com uma participante pela manhã e com doze participantes no período noturno e no dia 27 de novembro de 2019 com um participante. O teste de aceitabilidade foi aplicado primeiro, porque era mais rápido e com tempo de realização semelhante entre os participantes. Assim, eles não precisaram esperar por muito tempo para as instruções do teste de compreensão em leitura. Já o teste de compreensão era bastante mais demorado e com uma possível variação de tempo maior entre os participantes. Por isso, optou-se por realizá-lo por último, de sorte que eles poderiam deixar a sala, tão logo finalizassem a tarefa.

De modo resumido, o primeiro encontro aconteceu na seguinte ordem:

1. Leitura e assinatura do TCLE.
2. Instruções sobre a tarefa de Protocolo Verbal.
3. Treinamento do Protocolo Verbal.
4. Protocolo Verbal.
5. Instruções sobre o teste de Leitura Automonitorada.

6. Treinamento da Leitura Automonitorada.
7. Leitura Automonitorada.

No segundo encontro, por sua vez, adotou-se a seguinte ordem:

1. Instruções sobre o teste de Julgamento de Aceitabilidade.
2. Teste de Julgamento de Aceitabilidade.
3. Instruções sobre o teste de Compreensão em leitura.
4. Solicitação para ativar o Software VLC.
5. Teste de Compreensão em Leitura.

### 3.7 CRITÉRIOS E PROCEDIMENTOS DE ANÁLISE DOS DADOS

Nesta seção, são descritos os procedimentos de análise aplicados nos dados coletados. Os procedimentos de análise serão explicados a partir das perguntas e hipóteses específicas. Assim, mobilizaremos os aspectos quantitativos e qualitativos conjuntamente.

A análise qualitativa iniciou com a categorização das estratégias utilizadas durante os protocolos verbais e com o mapeamento dos comportamentos dos participantes durante a realização do teste de compreensão em leitura. Com isso, olhamos para as marcas de subjetividade que foram possíveis de se inferir a partir das estratégias utilizadas pelos participantes durante as sessões de protocolos verbais, a partir do comportamento estratégico adotado para responder ao teste de compreensão em leitura e a partir do produto gerado e apresentado como resposta às perguntas do teste de compreensão. Dessa forma, a análise qualitativa dá amparo aos aspectos processuais, enfatizando o ponto de vista dos participantes e dando profundidade às generalizações possibilitadas pelo olhar quantitativo (FLICK, 2009). Assim, “as descobertas qualitativas deverão facilitar a interpretação das relações existentes entre as variáveis dos conjuntos de dados quantitativos.” (FLICK, 2009, p. 40).

Para a análise estatística<sup>56</sup>, foi utilizado o programa SPSS (*Statistical Package for Social Science*) e, com o intuito de investigar as perguntas de pesquisa e testar as hipóteses que norteiam o estudo, foram adotados os seguintes procedimentos estatísticos gerais: (i) determinação do tipo de distribuição que as variáveis seguem, isso foi feito por meio de Testes de Normalidade (Kolmogorov-Smirnov e Shapiro-Wilk), inspeções visuais dos histogramas,

---

<sup>56</sup> A parte estatística da tese contou com a assessoria do professor Rafael Matielo (UFU), cujo lattes pode ser acessado no seguinte link: <http://lattes.cnpq.br/0268517166381932>.

gráficos Q-Q e diagramas de caixa para verificar a distribuição dos dados e determinar a adoção de testes paramétricos ou de alternativas não-paramétricas; (ii) testes de comparação com a finalidade de verificar possíveis diferenças ou semelhanças para algumas das variáveis aqui investigadas; e (iii) correlação entre algumas variáveis, a fim de investigar a associação existente ou não entre elas para a amostra investigada.

Para responder à **pergunta 1** e testar a **hipótese “a”**, foram consideradas duas medidas temporais obtidas por meio da leitura automonitorada, quais sejam: tempo total (TLT) das frases lidas pelos participantes e tempo do segmento pós-anafórico (TLSP). A tabulação foi feita de modo a considerar os tempos que cada participante levou em cada frase com “ele” e em cada frase com “o mesmo” e em cada segmento que era antecedido por “ele” e por “o mesmo”. Com isso em mãos, foi aplicado o teste dos Postos Sinalizados de Wilcoxon, tendo em vista que se trata da alternativa não-paramétrica adequada para comparações entre duas amostras pareadas/dependentes. No que diz respeito à acurácia, foi considerado o índice de acerto das questões de controle de atenção. O teste aplicado foi o de McNamar, pois a relação entre os dados é de pareamento e as variáveis são nominais. O olhar para a acurácia se faz essencial para responder à pergunta de pesquisa 1 e testar a hipótese “a”, pois o tempo de leitura do participante precisa refletir uma leitura atenta e que revele a compreensão acerca do lido, isso é garantido a partir da acurácia à resposta da pergunta de controle de atenção ao final de cada frase.

A **pergunta 2** foi respondida e a **hipótese “b”** testada a partir das variáveis obtidas no Teste de Compreensão em Leitura (TCL) – questões 2, 4, 6, 8 e 9 - e no Teste de Leitura Automonitorada. Os testes estatísticos aplicados foram os seguintes: Correlação de Pearson para as variáveis com distribuição aproximadamente normal (TLTLeElaMédia x TotalTCLLocal) e Coeficientes de Correlação de Postos de Spearman para as variáveis com distribuição livre (TLTOMesmoAMesmaMédia, TLSPOMesmoAMesmaMédia e TLSPEleElaMédia x TotalTCLLocal).

No que diz respeito ao Teste de Compreensão em Leitura (TCL), a pontuação de cada questão foi atribuída a partir de um padrão-resposta elaborado previamente com a ajuda de professores especialistas na área da Engenharia de Controle e Automação. Quando as respostas dos participantes diferiam daquelas previstas no padrão-resposta, elas foram analisadas por um especialista da área. A pontuação atribuída seguiu o parâmetro adotado para esta pesquisa.

Cada uma das questões valia 1,0 ponto, assim, quem gabaritasse o teste atingiria 10 pontos. Foi atribuído 1,0 ponto quando as questões descritivas foram respondidas de modo completo e de acordo com o esperado ou quando as questões objetivas foram adequadamente

respondidas. Quando a questão não foi respondida ou, então, foi respondida equivocadamente, foi atribuído 0,0. Quando as respostas às questões descritivas foram elaboradas de modo incompleto, foram atribuídas pontuações parciais de 0,5. A questão 6 e a questão 8 diferenciavam na pontuação, pois os participantes precisavam responder a mais itens. Assim, cada item de verdadeiro e falso da questão 6 valia 0,167 pontos. A questão 8, por sua vez, era composta por duas subquestões. A primeira solicitava ao participante a localização de referentes de determinadas anáforas, cada um desses referentes valia 0,1 pontos. Já a segunda questão valia 0,4 pontos. A análise e a especificação da pontuação de cada questão estão apresentadas no Apêndice F.

Para investigar a **pergunta 3** e testar a **hipótese “c”** foram utilizadas as médias dos postos da aceitabilidade das frases com o sintagma “o mesmo” e das frases com “ele”. Como o teste de aceitabilidade foi desenhado a partir de uma escala Likert de 7 pontos, as variáveis são consideradas como ordinais, por isso são comparadas as médias do ranqueamento (postos). O teste estatístico utilizado foi o Teste dos Postos Sinalizados de Wilcoxon. Buscamos, assim, examinar se a aceitabilidade do uso de “o mesmo” não apresenta diferenças com relação à aceitabilidade do uso de “ele” como anafóricos e se ambos os anafóricos podem ser considerados igualmente aceitáveis para o Português Brasileiro.

Para responder à **pergunta 4**, foram criadas categorias de comportamentos estratégicos utilizados pelos participantes. Construímos essas categorias a partir do comportamento estratégico dos participantes diante do texto quando estavam respondendo ao teste de compreensão em leitura. Chamamos as ações deles diante do texto de “comportamento estratégico” para diferenciarmos esses dados daqueles obtidos por meio dos protocolos. As categorias são:

**Retomada:** quando o participante sai do teste e volta para o texto para encontrar a resposta. Ou seja, a primeira leitura foi feita e ele está retornando ao texto para construir a resposta.

**Seleção:** quando o participante está retomando o texto e seleciona alguns trechos, mas sem marcar com o destaque.

**Busca:** quando o participante volta para o texto e utiliza a ferramenta de busca do PDF para encontrar uma palavra-chave da pergunta.

**Destaque:** quando o participante deixa destacada alguma parte do texto.

**Anotação:** quando o participante utiliza o bloco de notas para fazer alguma anotação sobre o texto. Essa anotação pode ocorrer durante a primeira leitura ou no decorrer do teste.



**Scanning:** quando o participante apenas passa rapidamente pelo texto, subindo e descendo, mostrando que não está fazendo uma leitura linear, mas que está explorando o texto.

As categorias foram criadas com base nas ações dos próprios participantes diante da leitura do texto e diante das ações tomadas para responder às questões. Com as categorias em mãos, catalogamos a frequência em que elas ocorreram e, em seguida, analisamos quais delas foram utilizadas diante da anáfora “o mesmo” – ou seja, essas categorias foram utilizadas no texto como um todo e, também, diante da anáfora. Isso foi feito para que pudéssemos estabelecer as relações entre o comportamento diante do texto como um todo e diante da anáfora “o mesmo” especificamente, depreendendo, dessa forma, possíveis dificuldades ou não.

Para a análise estatística, foram consideradas as seguintes categorias de comportamentos diante do texto: retomada, seleção, busca, destaque, anotação e *scanning*. Já no que tange ao comportamento diante da anáfora “o mesmo”, temos as seguintes variáveis: retomada, seleção, busca e destaque (que foram os comportamentos demonstrados pelos participantes). O teste de Correlação de Pearson foi usado para a correlação entre o Mapeamento da variável “retomada geral” e “retomada ‘o mesmo’” e, também, para a correlação entre os comportamentos gerais e a pontuação total obtida no Teste de Compreensão em Leitura. Já no que diz respeito às correlações entre as variáveis: “Seleção Geral”, “Busca Geral”, “Destaque Geral” e “Seleção ‘O mesmo’”, “Busca ‘O mesmo’” e “Destaque ‘O mesmo’” foi utilizado o teste de Correlação de Postos de Spearman. Além disso, os comportamentos diante da anáfora “o mesmo” foram correlacionados ao desempenho nas questões que envolviam compreensão local (2, 4, 6, 8 e 9), para isso, foi utilizado, também, o teste de Correlação de Postos de Spearman. Mais uma vez, os testes foram escolhidos tendo em vista a distribuição das variáveis e a resposta a que se buscava. Além disso, o olhar qualitativo foi importante para a análise das respostas que os participantes deram ao teste de compreensão leitora, sobretudo no que diz respeito ao resumo produzido, já que essa questão permite compreender se os participantes conseguiram formar uma base textual adequada do texto lido e isso, por sua vez, possibilita relacionar o processamento das anáforas-alvo e a compreensão do texto.

Para responder à **pergunta 5**, foram categorizadas as estratégias que os participantes adotaram durante a realização dos protocolos verbais, quais sejam: releitura, monitoramento da compreensão, conexões anafóricas intersentenciais, relato episódico, leitura em voz alta, avaliação do texto, elaboração e paráfrase. Para compreender as categorias criadas, cada uma delas é explicada na sequência:

Quadro 5: Estratégias utilizadas durante o Protocolo Verbal

<b>Estratégia</b>	<b>Explicação</b>	<b>Trecho do texto</b>	<b>Exemplo</b>
<b>Releitura</b>	Foi considerado que o participante estava fazendo uma releitura quando isso é dito pelo próprio participante (nesse caso, sabemos que se trata de uma releitura por conta do tempo de silêncio que vem depois - os participantes dizem que vão reler e, em seguida, ficam em silêncio) ou quando o participante faz a releitura em voz alta (nesse caso, a confirmação se dá pelo tempo de silêncio anterior à leitura em voz alta).	Isso acontece se a infraestrutura de computação subjacente possuir períodos em que não há garantias com relação aos tempos de processamento e transmissão das mensagens, o que gera a possibilidade de comprometimento no funcionamento de certos sistemas (ou subsistemas) de automação com requisitos temporais. Portanto, as hipóteses do modelo de sistema distribuído - isto é, o comportamento temporal e o comportamento de falhas dos componentes de <i>software</i> , dos dispositivos e dos canais de comunicação - são usadas para determinar as propriedades do serviço de detecção de defeitos durante sua operação.	Mas, como tudo isso é muito distribuído... as propriedades do serviço... [começa a reler o segundo trecho do quarto parágrafo 11:17-11:20] vão depender de uma garantia de qualidade de serviço, tá. (P13)
<b>Monitoramento da compreensão</b>	Consideramos essa estratégia quando o participante utiliza de expressões que demonstrem que está monitorando/verificando o que está lendo (com expressões como: ok! Certo!) e quando diz explicitamente que não compreendeu algo.	O comportamento dos componentes em caso de falhas guia a forma de implementação dos detectores de defeitos. O mesmo recebe a denominação de semântica de falhas, que determina a técnica de verificação (ou monitoramento) a ser utilizada pelo serviço de detecção de defeitos.	Certo, para mim, ficou um pouco confuso o... ao que o termo semântica de falhas se refere... se é o comportamento dos componentes ou a forma de implementação (P23)

<p><b>Conexões anafóricas intersentenciais</b></p>	<p>Foram computadas quando o participante relaciona alguma informação contida em uma sentença com a informação contida em outra sentença (em outra parte do texto) na qual existe uma conexão anafórica.</p>	<p>A detecção de defeitos é um serviço utilizado para construir mecanismos de tolerância a falhas. Ela consiste na verificação do funcionamento dos dispositivos ou componentes, de modo que, quando não se comportam de acordo com objetivos especificados, ações possam ser tomadas para substituição de componentes defeituosos, recuperação ou reconfiguração do sistema.</p>	<p>Um serviço utilizado pela área de tolerância de falhas é a detecção de defeitos que consiste de encontrar é... um... verificar que um componente não tá funcionando da maneira que deveria e ver qual é... qual ação pode ser tomada... (P1)</p>
<p><b>Relato episódico</b></p>	<p>Foi considerado quando é feita referência a disciplinas, a professores ou a contextos específicos que vieram à mente a partir do que está no texto.</p>	<p>A natureza descentralizada de sistemas de automação modernos tem impacto no projeto de detectores de defeitos. Em alguns desses sistemas, os dispositivos de automação estão distribuídos e interconectados a partir de redes de comunicação, por exemplo: Sistemas de Controle Sobre Rede (NCS, <i>Networked Control Systems</i>); Sistemas de Controle Distribuídos (DCS, <i>Distributed Control Systems</i>), aplicações distribuídas sobre Redes Veiculares Ad Hoc (VANET, <i>Vehicular Ad hoc Networks</i>), Sistemas Ciber-Físicos (CPS, <i>Cyber-Physical Systems</i>), entre outros.</p>	<p>e esses sistemas ciber-físicos eu já tinha ouvido falar em uma disciplina. (P13)</p>

<b>Ler em voz alta</b>	Quando o participante lê um trecho pela primeira vez em voz alta e depois silencia. A confirmação de ser uma “leitura em voz alta” e não uma releitura se dá por conta da observação dos tempos de silenciamento.	A detecção de defeitos é um serviço utilizado para construir mecanismos de tolerância a falhas. Ela consiste na verificação do funcionamento dos dispositivos ou componentes, de modo que, quando não se comportam de acordo com objetivos especificados, ações possam ser tomadas para substituição de componentes defeituosos, recuperação ou reconfiguração do sistema.	[01:52 começa a ler alto:] verificação do funcionamento dos dispositivos e componentes... de modo que [continua em silêncio até 02:15] (P18)
<b>Avaliação do texto</b>	Foi considerada avaliação do texto quando o participante concorda, discorda ou questiona algo citado pelo texto.	Os termos “falha”, “erro” e “defeito”, utilizados neste texto, seguem o padrão de terminologia amplamente adotado pela comunidade científica internacional da área de tolerância a falhas.	É... a parte sobre falha, erro e defeito é... eu sei que cada um tem a sua definição, às vezes elas se confundem entre si, mas é interessante abordar os três termos nesse texto. (P11)
<b>Elaboração e</b>	Nessa categoria entrou tudo aquilo que o participante desenvolve com base em seu conhecimento prévio (dando exemplos ou fazendo comentários) ou quando o participante cita explicitamente que algo apresentado no texto o levou a compreender aquilo de determinada forma (Inferência).	Todavia, nesses ambientes de automação distribuídos, as propriedades e o desempenho do serviço de detecção de defeitos dependem das garantias de qualidade de serviço (QoS, <i>Quality of Service</i> ) entregues pelos sistemas de processamento e de comunicação utilizados.	É... Mas, também a rede tem que ser boa, então tem que se ter uma rede confiável para, para você ter certeza que realmente o que você está rece/ que os dados trafegando naquela rede condizem com... é... os... os sensores que estão no local. (P2)

<p><b>Paráfrase</b></p>	<p>Ocorre quando é feita uma explicação, um resumo ou são repetidas partes consideradas importantes (um participante foi dizendo o que marcaria em cada trecho) do trecho lido, usando apenas outras palavras ou outras estruturas sintáticas, mas, basicamente, repetindo o que está no texto.</p>	<p>A detecção de defeitos é um serviço utilizado para construir mecanismos de tolerância a falhas. Ela consiste na verificação do funcionamento dos dispositivos ou componentes, de modo que, quando não se comportam de acordo com objetivos especificados, ações possam ser tomadas para substituição de componentes defeituosos, recuperação ou reconfiguração do sistema.</p>	<p>Hum... mais ou menos eu já comentei sobre isso eu acredito, mas é... que fala mais sobre a detecção dos defeitos em si, no caso, e como que [02:44 – 02:48 silêncio] é... e como verificar esses defeitos e, no caso, gerar alguma coisa para eles serem tratados... é... (P21)</p>
-------------------------	---	---	--

Fonte: a autora (2022).

Esses foram os critérios que orientaram a catalogação dessas estratégias, que foram criadas a partir das verbalizações dos participantes, mas, também, com o amparo do que a literatura apresenta sobre estratégias de leitura (SOLÉ, 1998; AFFLERBACH; PEARSON; PARIS, 2008; KENDEOU; BROEK; 2007).

Para responder à primeira parte da pergunta, foi aplicado o teste de Correlação de Postos de Spearman, correlacionando cada estratégia categorizada com cada uma das perguntas que envolviam compreensão local (2, 4, 6, 8 e 9) e, em seguida, o mesmo teste foi usado para correlacionar cada estratégia com o desempenho geral de cada participante no TCL. Na segunda parte da pergunta, foi aplicado o teste de Wilcoxon, a partir do qual foi feita a comparação das médias de cada estratégia aplicada diante de anáfora “o mesmo” e de cada estratégia utilizada diante da anáfora “ele”. Esse teste foi utilizado porque, nesse caso, as amostras são pareadas, pois se trata do mesmo participante aplicando estratégias diante de variáveis distintas (“o mesmo” e “ele”). Com a análise qualitativa, foi feita, ainda, a triangulação dos dados obtidos por meios do teste de compreensão em leitura, do mapeamento de tela e do protocolo verbal. Assim, as estratégias utilizadas durante a leitura (Releitura, Monitoramento da Compreensão, Conexões anafóricas intersentenciais, Relato Episódico, Leitura em voz alta, avaliação do texto, elaboração e paráfrase) serão relacionadas aos comportamentos estratégicos dos participantes (Retomada Geral, Seleção Geral, Busca Geral, Destaque Geral, Anotação e *Scanning*), principalmente no que diz respeito às anáforas-alvo.

Por fim, para poder responder ao problema investigativo e testar a hipótese geral, a análise qualitativa entra em cena porque, com ela, é possível desenvolver explicações acerca da influência do sintagma “o mesmo” na compreensão em leitura. A relevância de uma análise que parta dessa perspectiva está no estabelecimento de possíveis relações entre as diferentes técnicas utilizadas para coletar os dados. Ademais, a perspectiva qualitativa dá base à discussão de pontos que podem escapar à análise quantitativa, devido ao N da pesquisa. Dessa forma, é possível ponderar sobre desempenhos individuais e como tais desempenhos podem ter impactado nos resultados gerais.

A fim de resumir os procedimentos aplicados para a realização da análise deste estudo, apresentamos o quadro a seguir, em que estão dispostos cada variável a ser considerada em cada pergunta específica e os testes estatísticos utilizados:

**Pergunta Geral:** de que forma e em que medida o sintagma “o mesmo”, como anáfora correferencial, atua na integração local, na formação da base textual e na compreensão em leitura de texto acadêmico-científico?

**Hipótese Geral:** “o mesmo” segue um padrão de processamento semelhante ao do pronome “ele”. Sendo assim, espera-se que os resultados em termos de processamento e de produto mostrem que “o mesmo” pode exercer a função de anáfora correferencial sem que isso gere entraves para a compreensão em leitura. Então, ainda que o sintagma “o mesmo” não se caracterize como pronome, ele pode assumir tal função e atuar como anáfora correferencial, sem que isso gere complicações ou atrasos no processamento e na compreensão.

Quadro 6: Resumo das variáveis e dos testes aplicados

<b>Perguntas</b>	<b>Hipóteses</b>	<b>Instrumentos</b>	<b>Variáveis consideradas</b>	<b>Testes estatísticos aplicados</b>	<b>Análise Qualitativa</b>
Pergunta 1: O uso de “o mesmo” como anáfora correferencial implica maior custo no tempo de resposta?	Hipótese 1: Os custos de tempo de resposta não são afetados pelo uso de “o mesmo” como anáfora correferencial. Ou seja, o custo de tempo de resposta de “o mesmo” é semelhante ao custo de tempo de resposta do pronome ele.	Leitura automonitorada	Tempo total (TLT) de leitura das frases com “ele” e das frases com “o mesmo”.  Tempo de leitura do segmento pós-anafórico (TLSP) de “ele” e de “o mesmo”  Acurácia – índice de	Testes dos Postos Sinalizados de Wilcoxon  Teste de McNemar	Análise indutiva-interpretativa das marcas de subjetividade

			acertos às perguntas de controle de atenção		
Pergunta 2: Considerando que o desempenho local do participante é semelhante ao desempenho geral, qual a relação entre desempenho local e tempo de resposta na leitura de “o mesmo” e de “ele” como anáforas correferenciais? Há diferenças na relação entre tempo de resposta de “o mesmo” e desempenho local e tempo de resposta de “ele” e desempenho local?	Hipótese 2: O tempo de resposta na leitura de “o mesmo” afeta positivamente o desempenho local, assim como ocorre com o tempo de resposta na leitura de “ele”. Isto é: Quanto menor o tempo de resposta na leitura desses dois anafóricos, melhor o desempenho em leitura local do texto.	Teste de compreensão em leitura e Teste de leitura automonitorada.	Soma da pontuação obtida nas questões que medem desempenho local (TotalTCLLocal).  Média do tempo de resposta das frases com “ele” e com “o mesmo” (TLTLEleELaMédia e TLTOmesmoAMesmaMédia)  Média do tempo de resposta do segmento pós-anafórico “ele” e “o mesmo” (TLSPEleElaMédia e TLSPOMesmoAMesmaMédia)	Correlação de Pearson e Coeficientes de Correlação de Postos de Spearman	Análise indutiva-interpretativa das marcas de subjetividade do produto gerado e apresentado como resposta às perguntas do teste de compreensão
Pergunta 3: Qual a aceitabilidade do uso de “o mesmo” como elemento anafórico correferencial em construções sentenciais do PB?	Hipótese 3: A aceitabilidade do uso de “o mesmo” não é diferente da aceitabilidade do uso de “ele” como anafóricos.	Teste de aceitabilidade	Aceitabilidade de “ele” e Aceitabilidade de “O mesmo”	Testes dos Postos Sinalizados de Wilcoxon	Análise indutiva-interpretativa dos resultados estatísticos
Pergunta 4: Qual a relação entre o comportamento geral, o comportamento diante da anáfora correferencial construída com o sintagma “o mesmo” e o desempenho diante das retomadas	Exploratória	Teste de compreensão em leitura	Mapeamento do comportamento geral (Retomada Geral, Seleção Geral, Busca Geral, Destaque Geral, Anotação e <i>Scanning</i> ).  Mapeamento do comportamento diante da anáfora “o mesmo” (Retomada “o mesmo”,	Correlação de Pearson e Coeficientes de Correlação de Postos de Spearman	Análise indutiva-interpretativa das marcas de subjetividade dos comportamentos estratégicos mapeados

anafóricas e do texto?			<p>Seleção “o mesmo”, Busca “o mesmo”, Destaque “o mesmo”).</p> <p>Desempenho compreensão leitora geral (soma da pontuação obtida nas 10 questões)</p> <p>Desempenho compreensão leitora local (questões 2, 4, 6, 8 e 9)</p>		
Pergunta 5: Que estratégias são utilizadas pelo leitor e qual a relação de tais estratégias com o desempenho diante das retomadas anafóricas e do texto? Ainda, é possível, no que diz respeito às estratégias empregadas pelos participantes, observar semelhanças frente às diferentes anáforas lidas?	Exploratória	Protocolo Verbal e Teste de Compreensão em Leitura	<p>Estratégias (Releitura, Monitoramento da Compreensão, Conexões anafóricas intersentenciais, Relato Episódico, Leitura em voz alta, avaliação do texto, elaboração e paráfrase)</p> <p>Desempenho geral (soma da pontuação obtida nas 10 questões)</p> <p>Desempenho local (questões 2, 4, 6, 8 e 9)</p>	Coeficientes de Correlação de Pearson e Coeficientes de Correlação de Postos de Spearman	Análise indutiva-interpretativa das marcas de subjetividade das estratégias categorizadas

Fonte: a autora.

As análises serão organizadas de acordo com as perguntas específicas que guiam este estudo. Assim, pergunta por pergunta, os resultados estatísticos serão descritos e, na sequência, serão feitas as análises quantitativas e qualitativas, voltando nosso olhar para as subjetividades, saindo das generalizações e olhando para as relações entre o estabelecimento da anáfora e a construção da base textual a partir de uma análise indutiva, em que, por meio dos dados catalogados, são feitas inferências com base na teoria e nos resultados estatísticos obtidos. Com isso, buscamos responder à grande questão que guia esta investigação.



## 4 APRESENTAÇÃO, ANÁLISE E DISCUSSÃO DOS RESULTADOS

Este capítulo é dedicado à apresentação, análise e discussão dos dados coletados durante o processo investigativo, conforme objetivos expostos e de acordo com o método aplicado. Os movimentos analíticos qualitativos e quantitativos nos darão base para acrescentar uma peça nos estudos da leitura de textos, pensando no processamento anafórico do sintagma “o mesmo”.

### 4.1 DESCRIÇÃO E ANÁLISE DOS RESULTADOS

Cada um dos instrumentos empregados nesta investigação contribuiu com informações relativas a distintos aspectos relacionados à leitura. Mas é no conjunto, nas correlações e nas triangulações que conseguimos chegar ao nosso grande objetivo. A fim de tornar a leitura da tese mais fluida, decidimos por apresentar os dados de forma a já responder cada uma das perguntas específicas do estudo. Assim, nas próximas 5 subseções, trazemos os resultados a que chegamos durante o percurso trilhado.

Antes de apresentar os resultados estatísticos e a análise qualitativa, cabe ressaltar que, devido ao N amostral, optamos por não retirar os *outliers* da amostra para as rodadas estatísticas. Isso porque, quando retiramos um participante *outlier*, encontramos outros. Se optássemos por retirá-los, ficaríamos com um número muito reduzido de dados para realizar as análises. Além disso, manter todos os participantes com suas discrepâncias é importante para retratar a natureza dos processos desse grupo específico. Ao retirar os *outliers*, isso poderia ficar comprometido. Por conta disso, trazemos os resultados que esse grupo de participantes específico gerou, considerando suas discrepâncias e congruências.

#### 4.1.1 O uso de “o mesmo” como anáfora correferencial implica maior custo no tempo de resposta?

Para responder à pergunta que intitula esta subseção, precisamos, inicialmente, apresentar os resultados obtidos por meio do teste de leitura automonitorada. Como já especificado no capítulo do método, trabalhamos com duas medidas temporais a fim de que pudéssemos circunscrever as possíveis nuances de processamento do sintagma “o mesmo” e do pronome “ele”. Dessa forma, temos resultados obtidos com o tempo total de leitura de cada frase e, também, com o tempo de leitura do segmento pós-anafórico. Contudo, antes de trazer os resultados relativos ao tempo de leitura, é necessário verificar se a leitura que os participantes

fizeram das frases foi acurada, isto é, se leram com a devida atenção. Essa importância se dá porque precisamos considerar que o tempo de leitura gerado representa uma leitura atenta e em que a correferência tenha sido estabelecida. Para garantir isso, ao final de cada sentença, foi feita uma pergunta que tinha por objetivo central controlar a atenção dos participantes. O número total apresentado nos resultados da leitura automonitorada se dá porque consideramos as sentenças lidas pelos participantes e não o número de participantes em si. Isto é, cada um dos 26 participantes leu 6 frases com o pronome “ele” e 6 frases com o sintagma “o mesmo”, assim temos um total de 156 frases lidas para cada um dos anafóricos. São esses resultados que reportamos neste momento:

Tabela 1: Acurácia “O mesmo”

	Frequência	Percentual
<b>Sim</b>	149	95,5
<b>Não</b>	7	4,5
<b>Total</b>	156	100

Fonte: a autora.

As perguntas de controle de atenção das frases experimentais deveriam sempre responder “sim”, isto é, se o participante respondesse “sim”, acertaria a questão. De acordo com os resultados obtidos, podemos dizer que a leitura realizada pelos participantes das frases com o sintagma “o mesmo” foi atenta, pois o índice de acertos foi de mais de 95%.

A seguir reportamos os resultados descritivos acerca da acurácia na leitura das frases cujo anafórico utilizado foi o pronome:

Tabela 2: Acurácia “ele”

	Frequência	Percentual
<b>Sim</b>	149	95,5
<b>Não</b>	7	4,5
<b>Total</b>	156	100

Fonte: a autora.

O índice de acertos foi exatamente o mesmo daquele obtido com as frases cuja anáfora foi realizada com o sintagma “o mesmo”. Mais um indício de que a leitura feita na maior parte das frases foi atenta e revela a integração do anafórico com seu antecedente, mostrando que a

compreensão literal do conjunto de sentenças lidas foi atingida. Dessa forma, podemos considerar os tempos de leitura obtidos como tempos válidos.

O teste estatístico utilizado para comparar a acurácia das frases que contêm o anafórico “o mesmo” e as frases que contêm o pronome foi o de McNemar, pois os dados são nominais e a relação entre eles é de pareamento. A seguir, os resultados:

Tabela 3: Estatística do teste de acurácia

Acurácia “o mesmo” e “ele”	
<b>N</b>	156
<b>Exact Sig. (2-tailed)</b>	1000

Fonte: a autora.

Os resultados obtidos com o Teste de McNemar confirmam o que já era possível de se concluir visualmente, ou seja, não há diferenças estatisticamente significativas entre a acurácia de “o mesmo” e de “ele”,  $p > 0,05$ . A acurácia foi exatamente a mesma nos dois casos. Esses resultados nos dão, também, indícios de que os estímulos apresentados aos participantes foram bem controlados. Além disso, os resultados mostram que, ao lerem tanto o sintagma “o mesmo” quanto o pronome, os participantes estabelecem adequadamente a correferência com o antecedente, sem que haja diferença no estabelecimento da correferência construída com qualquer uma das variáveis testadas. No que diz respeito à acurácia, então, pronome e sintagma “o mesmo” parecem atuar de modo semelhante, ambos funcionando como anáfora correferencial.

Sabendo que a leitura realizada pelos participantes foi atenta e que a correferência foi estabelecida, podemos passar aos resultados relativos aos tempos de leitura. Iniciamos descrevendo os resultados do tempo total de leitura em milissegundos.

Tabela 4: Estatística Descritiva – Leitura Automonitorada – Tempo Total em milissegundos

Medida	Estatística – O mesmo	Estatística - Ele
<b>Média</b>	5152,10	5417,23
<b>Mínimo</b>	1363,06	1228,94
<b>Máximo</b>	15931,87	14893,81
<b>Desvio Padrão</b>	2455,27	2614,63

Fonte: a autora.

Quando medimos o tempo total de leitura dos participantes, temos uma média de tempo maior para as frases cuja anáfora foi realizada com o pronome “ele” se comparada à média do tempo de leitura das frases realizadas com o sintagma “o mesmo”. Para investigar se tal diferença é estatisticamente significativa ou não, realizamos testes de normalidade para escolher o melhor teste a ser utilizado.

Tabela 5: Normalidade – Teste de Leitura Automonitorada – Tempo Total em milissegundos

	Shapiro-Wilk		
	Estatística	df	Sig.
TLTEle	0,90	156	0,00
TLTOMesmo	0,90	156	0,00

Fonte: a autora.

Os resultados obtidos com os Testes de Shapiro-Wilk demonstram que os dados não seguem uma distribuição aproximadamente normal para as variáveis TLTEle e TLTOMesmo ( $p < 0,05$ ).

Tendo em vista que a distribuição dos dados não é normal e o objetivo está em comparar duas amostras dependentes, utilizamos o teste de Wilcoxon para compararmos os tempos de leitura de cada uma das anáforas-alvo:

Tabela 6: Postos do tempo total de leitura em milissegundos

		N	Média do Ranqueamento	Soma das Ranques
TLTEle	Ranques Negativos	72 <sup>a</sup>	78,50	5652,00
TLTOMesmo	Ranques Positivos	84 <sup>b</sup>	78,50	6594,00
	Ranques Iguais	0 <sup>c</sup>		
	Total	156		

<sup>a</sup>TLTEle < TLTOMesmo  
<sup>b</sup>TLTEle > TLTOMesmo  
<sup>c</sup>TLTEle = TLTOMesmo

Fonte: a autora.

O que os postos do teste mostram é que, em 72 casos, o tempo total de leitura do pronome foi menor que o tempo total de leitura do sintagma “o mesmo”. No entanto, em 84 casos, esse tempo foi maior; por isso, como vimos, a média do tempo total de leitura das frases com o pronome foi maior. Essa diferença, contudo, não se mostra estatisticamente significativa, como demonstrou o teste de Wilcoxon:

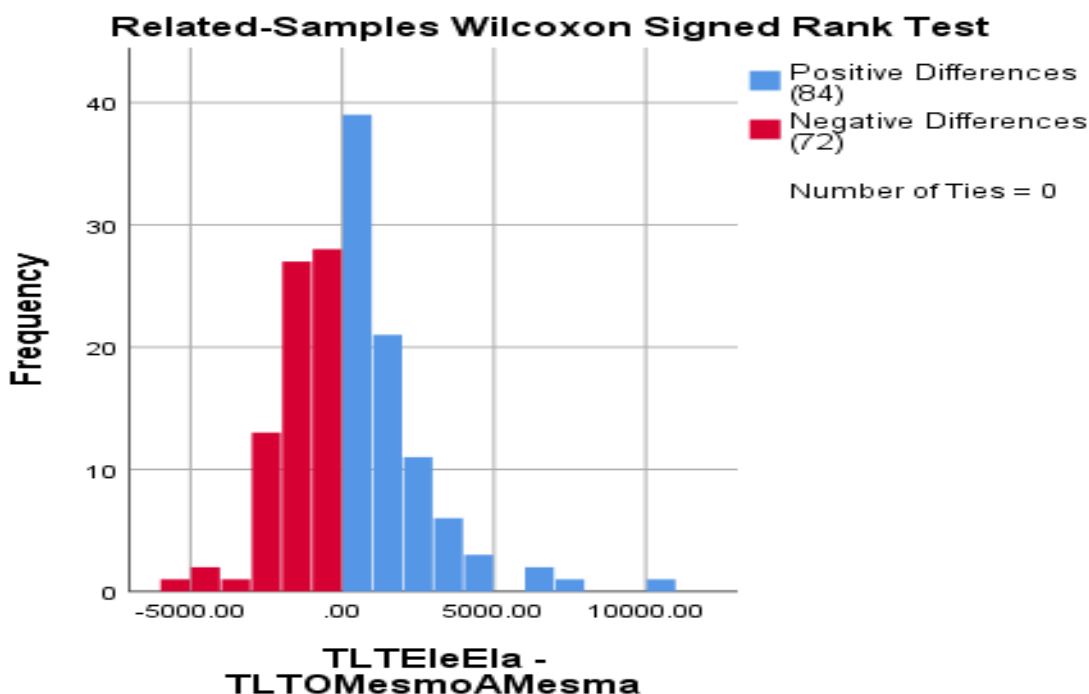
Tabela 7: Estatística do tempo total de leitura em milissegundos

	TLTEle - TLTOMesmo <sup>a</sup>
Z	-0,83 <sup>b</sup>
p	0,40
<sup>a</sup> Teste de Postos Sinalizados de Wilcoxon	
<sup>b</sup> Baseado nos ranques negativos	

Fonte: a autora.

O teste dos Postos Sinalizados de Wilcoxon indica que os postos relativos ao tempo total de leitura das frases com “o mesmo” não são estatisticamente diferentes dos postos relativos ao tempo total de leitura das frases com “ele”,  $Z = -0,83$ ,  $p > 0,05$ . Os resultados podem ser visualizados no gráfico a seguir:

Gráfico 1: teste dos Postos Sinalizados de Wilcoxon – tempo total em milissegundos



Fonte: a autora.

No gráfico, temos a distribuição das diferenças positivas e das diferenças negativas no ranqueamento. Na vertical, é possível visualizar a frequência e, na horizontal, o tempo total em milissegundos. À esquerda do gráfico (em vermelho) podemos, então, visualizar a frequência dos tempos em que a leitura do pronome foi menor que a leitura do sintagma e, à direita (em azul), conseguimos visualizar os tempos em que a leitura do pronome foi maior que a do sintagma. Como a medida utilizada foi bastante precisa (milissegundos), não houve empates nos tempos de leitura. Esses resultados até aqui reportados vão ao encontro do que previmos.

Primeiro, porque a acurácia das frases construídas com o sintagma “o mesmo” foi igual à acurácia das frases construídas com “ele”. Depois, porque a diferença no tempo de leitura entre os dois anafóricos não se mostrou estatisticamente significativa. Esse resultado demonstra que ambos os anafóricos exercem papel semelhante no processamento. De acordo com os tempos obtidos, ambos os anafóricos estão automatizados e retomam o antecedente de modo adequado, não gerando maiores custos para o processamento.

Na sequência, são apresentados os dados do tempo de leitura do segmento pós-crítico, isto é, o tempo que os participantes levaram para ler o verbo que estava após a anáfora.

Tabela 8: Estatística Descritiva – Leitura Automonitorada – Tempo em milissegundos Segmento Pós-anafórico

Medida	Estatística – O mesmo	Estatística - Ele
<b>Média</b>	460,33	486,37
<b>Mínimo</b>	23,85	9,48
<b>Máximo</b>	4195,54	3452,89
<b>Desvio Padrão</b>	412,72	471,68

Fonte: a autora.

Ao se considerar a média do tempo de leitura, temos que os segmentos após a anáfora realizada com o pronome foram lidos mais lentamente que os segmentos realizados após o sintagma “o mesmo”; é preciso, contudo, verificar se a diferença foi ou não estatisticamente significativa. Para tanto, rodamos o teste de normalidade, que nos ajudou na decisão quanto ao teste estatístico mais adequado. A seguir, reportamos os resultados acerca da normalidade dos dados:

Tabela 9: Normalidade – Teste de Leitura Automonitorada – Tempo Segmento Pós-anafórico

	Shapiro-Wilk		
	Estatística	df	Sig.
TLTEle	0,580	156	0,00
TLTOMesmo	0,540	156	0,00

Fonte: a autora.

Os resultados obtidos com os Testes de Shapiro-Wilk demonstram que os dados não seguem uma distribuição aproximadamente normal para as variáveis TLSPOMesmo e TLSPEle ( $p < 0,05$ ). Tendo em vista esse resultado e considerando a comparação de duas variáveis, usamos os testes dos postos sinalizados de Wilcoxon para compararmos os tempos de leitura:

Tabela 10: Postos do tempo de leitura do segmento pós-anafórico em milissegundos

		N	Média do Ranqueamento	Soma das Ranques
TLTEle	Ranques Negativos	79 <sup>a</sup>	78,66	6214,00
TLTOMesmo	Ranques Positivos	77 <sup>b</sup>	78,34	6032,00
	Ranques Iguais	0		
	Total	156		
<sup>a</sup> TLTEle < TLTOMesmo				
<sup>b</sup> TLTEle > TLTOMesmo				
<sup>c</sup> TLTEle = TLTOMesmo				

Fonte: a autora.

No que diz respeito aos tempos de leitura do segmento pós-anafórico, temos que, em 79 casos, o tempo de leitura do segmento que estava após o pronome foi menor do que o do segmento que estava após o sintagma “o mesmo” e, em 77 casos, o tempo de leitura do segmento que estava após o pronome foi maior. Aqui, então, temos uma proximidade grande entre o tempo de leitura do segmento que estava depois do pronome e do que estava após o sintagma “o mesmo”. A seguir, apresentamos a estatística do teste a fim de descobrirmos se a diferença reportada é ou não significativa

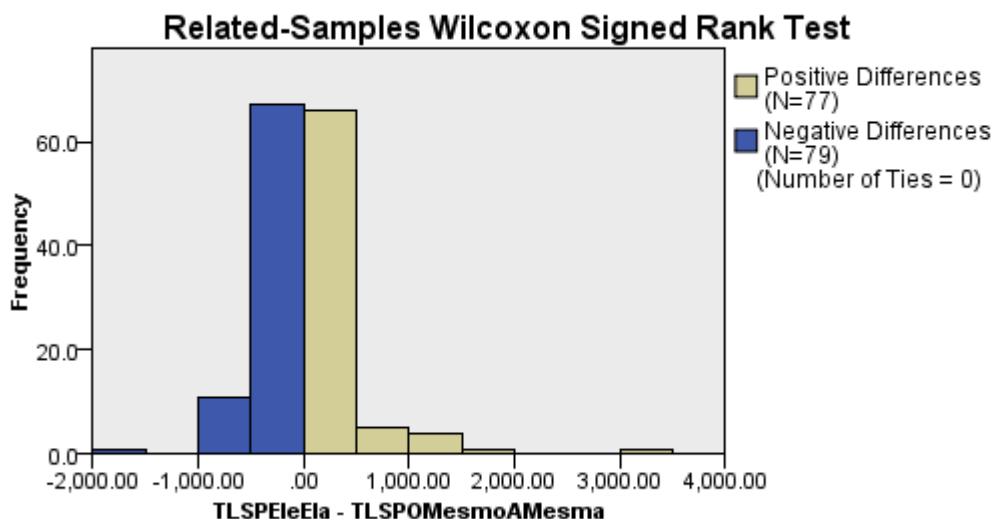
Tabela 11: Estatística do tempo de leitura do segmento pós-anafórico em milissegundos

	TLTEle - TLTOMesmo <sup>a</sup>
Z	-0,16 <sup>b</sup>
p	0,87
<sup>a</sup> Teste de Postos Sinalizados de Wilcoxon	
<sup>b</sup> Baseado nos ranques negativos	

Fonte: a autora.

O teste dos Postos Sinalizados de Wilcoxon indica, então, que os postos relativos ao tempo de leitura do segmento pós-crítico das frases com “o mesmo” não são estatisticamente diferentes dos postos relativos ao tempo de leitura do segmento pós-crítico das frases com “ele”,  $Z = -0,16$ ,  $p > 0,05$ . Os resultados podem ser visualizados no gráfico a seguir:

Gráfico 2: teste dos Postos Sinalizados de Wilcoxon – tempo segmento pós-anafórico



Fonte: a autora.

No gráfico, temos a frequência dos diferentes tempos em que o pronome foi lido mais rapidamente que o sintagma “o mesmo” (diferenças negativas – em azul) e a frequência dos tempos em que o pronome foi lido mais rapidamente que o sintagma (diferenças positivas – em bege). Assim, as duas medidas adotadas para análise neste trabalho indicam que o sintagma “o mesmo” é processado de modo semelhante ao pronome “ele”, o que vai ao encontro do que levantamos como hipótese a partir dos estudos que têm sido realizados na área do processamento anafórico e dos estudos na área da sociolinguística acerca do sintagma “o mesmo”.

O que os resultados obtidos por meio do teste de leitura automonitorada demonstram, então, é que o sintagma “o mesmo” é processado como anafórico quando está substituindo sintagmas de valor nominal, não demonstrando qualquer significação intrínseca e assumindo um significado de acordo com o antecedente que retoma (OAKHIL; CAIN; ELBRO, 2017). Assim, da mesma forma que Pereira (2013, p.184) conjecturou a partir de trabalhos escritos, o sintagma “o mesmo”, na leitura, parece cumprir o papel de retomar “um antecedente sem imprimir-lhe característica ou juízo de valor”, assim como o faz o pronome “ele”. É isso que nos sugerem os resultados obtidos por meio do teste de leitura automonitorada, a partir da análise de duas medidas temporais distintas e, também, da acurácia. Então, a partir dos resultados obtidos por meio do teste de leitura automonitorada, o sintagma “o mesmo” parece atuar como elemento de retomada construindo cadeias coesivas e fazendo com que o referente previamente introduzido seja mantido no foco do leitor, da mesma forma que ocorre com o pronome (KOCH; ELIAS, 2015).



Esse instrumento traz resultados relativos ao processamento local, ou seja, localmente, o sintagma “o mesmo” se mostra capaz de integrar sentenças. Para que possamos, contudo, confirmar o papel dessas anáforas na construção da representação local de um texto, vamos correlacionar, na sequência, os tempos obtidos na leitura automonitorada com o desempenho local dos participantes no teste de compreensão leitora.

#### **4.1.2 Qual a relação entre desempenho local e tempo de resposta na leitura de “o mesmo” e de “ele” como anáforas correferenciais? Há diferenças na relação entre tempo de resposta de “o mesmo” e desempenho local e tempo de resposta de “ele” e desempenho local?**

Para responder à pergunta 2, retomamos os resultados obtidos com o teste de leitura automonitorada e os correlacionamos com os resultados obtidos no teste de compreensão leitora nas questões que avaliavam o desempenho local dos participantes. A decisão por correlacionar com o desempenho local se deu porque a integração da anáfora ocorre localmente. Dessa forma, precisamos saber se, além do processamento, o produto, isto é, o desempenho também demonstra o estabelecimento adequado da anáfora realizada com o sintagma “o mesmo” e se o uso desse elemento no texto não inviabiliza essa integração mais local.

Para poder realizar essa correlação, trabalhamos com as médias de tempo de leitura total e de tempo de leitura do segmento pós-anafórico obtidas pelos participantes nas seis frases lidas. Optamos por trabalhar com as médias de tempo porque, para esta pergunta, precisávamos de uma relação um por um, ou seja, o tempo de leitura do participante com a sua pontuação nas questões locais do teste. Como a média indica a posição intermediária de tempo de cada participantes, optamos por ela ao invés de trabalharmos com a soma de tempos. Em seguida, essa média foi correlacionada com a pontuação total alcançada pelos participantes nas questões que exigiam um processamento local para serem respondidas. Diante disso, reportamos, na sequência, a estatística descritiva com as médias de tempo de leitura e da pontuação nas questões locais. Feito isso, apresentamos os resultados da correlação realizada.

Tabela 12: Estatística Descritiva – Leitura Automonitorada – Média em milissegundos

Medida	Estatística – O mesmo – Tempo Total	Estatística – Ele – Tempo Total	Estatística – O mesmo – Pós-anafórico	Estatística – Ele – Pós-anafórico
<b>Média</b>	5138,12	5378,16	460,38	489,43
<b>Mínimo</b>	2063,74	2098,72	174,47	190,56
<b>Máximo</b>	12041,72	11887,35	1470,02	1281,92
<b>Desvio Padrão</b>	2156,90	2186,14	261,34	229,86

Fonte: a autora.

O primeiro ponto a se considerar é que, quando trabalhamos com a média de leitura de cada participante, podemos visualizar uma proximidade entre os tempos de leitura, tanto no tempo total quanto no tempo de leitura do segmento pós-anafórico. Ainda que, para esta pergunta, não tenhamos verificado se há diferença significativa ou não, deixamos destacado essa aparente proximidade no resultado a fim de retomá-la em nossas discussões finais.

No que diz respeito ao teste de normalidade, temos o seguinte:

Tabela 13: Normalidade – Teste de Leitura Automonitorada – Média

	Shapiro-Wilk		
	Estatística	df	Sig.
TLTOMesmo	0,902	26	0,017
TLTEle	0,934	26	0,097
TLSPOMesmo	0,758	26	0,000
TLSPEle	0,867	26	0,003

Fonte: a autora.

Os resultados obtidos com os Testes de Shapiro-Wilk demonstram que os dados das variáveis TLTOMesmoMédia, TLSPOMesmoMédia e TLSPEleMédia não têm distribuição normal ( $p > 0,05$ ), ao passo que os dados da variável TLTEleMédia possuem uma distribuição próxima da normalidade. Visto isso, na sequência, reportamos os resultados obtidos a partir do desempenho local dos participantes no teste de compreensão leitora.

Tabela 14: Estatística Descritiva – Teste de Compreensão Leitora – Local

Medida	TCL - Local
<b>Média</b>	3,70
<b>Mínimo</b>	2,5
<b>Máximo</b>	4,9
<b>Desvio Padrão</b>	0,65

Fonte: a autora.

O desempenho em compreensão local dos participantes partiu de cinco questões do teste de compreensão leitora (Apêndice F). Caso acertassem às cinco, os participantes fariam cinco pontos no teste. A média ficou em 3,7, sendo que a pontuação mais alta foi 4,9. Isso significa que ninguém gabaritou as cinco questões que exigiam um bom desempenho local na leitura do texto. A partir disso, é possível levantar algumas asserções acerca de dois pontos: 1. Problemas na compreensão local por parte dos participantes; 2. Variáveis intervenientes que podem ter influenciado nesse resultado, como: falta de atenção por parte do participante, cansaço e outros fatores difíceis de serem controlados em uma pesquisa com participantes humanos. Quanto à compreensão local em si, não parece ser problemática, sendo que a média dos participantes foi alta, ou seja, ainda que ninguém tenha gabaritado, o índice de acertos demonstra que os leitores conseguiram dar conta das integrações locais necessárias para processar de modo proficiente as relações contidas nesse nível, que foram construídas a partir da localização de itens específicos no texto. Dessa forma, o fato de não terem atingido a pontuação máxima nos itens que envolviam compreensão local, não parece ter ocorrido por alguma dificuldade nesse nível de processamento. Então, existe a possibilidade de variáveis intervenientes terem influenciado no fato de nenhum dos participantes ter acertado todas as cinco questões que envolviam compreensão local.

Tabela 15: Normalidade – Teste de Leitura Automonitorada – Média

	Shapiro-Wilk		
	Estatística	df	Sig.
Total TCLLocal	0,926	26	0,063

Fonte: a autora.

Os resultados obtidos com os Testes de Shapiro-Wilk demonstram que os dados da variável TotalTCLLocal possuem distribuição próxima da normalidade ( $p > 0,05$ ).

Os testes de normalidade mostram, então, o seguinte cenário: variáveis com distribuição aproximadamente normal - TLTEleMédia x TotalTCLLocal - e variáveis sem distribuição normal - TLTOMesmoMédia, TLSPOMesmoMédia e TLSPEleMédia x TotalTCLLocal. A visualização desses cenários é importante para a escolha do teste de correlação. Diante disso, no que diz respeito às variáveis com distribuição aproximadamente normal, usaremos o teste de Pearson. Para as variáveis sem distribuição aproximadamente normal, o de Spearman.

Tabela 16: Correlações - Pearson

		TotalTCLLocal
TLTEleMédia	Correlação de Pearson	-0,148
	Significância	0,470
	N	26

Fonte: a autora.

No que diz respeito à correlação entre o tempo de leitura total das frases com “ele” e o desempenho em questões locais do teste, temos uma correlação fraca, negativa e não significativa. Uma correlação fraca e negativa mostra que, quando o tempo de leitura sobe (maior custo para o processamento), o desempenho desce, mas o fato de ser uma correlação fraca pode indicar um indício de isso ocorrer ao acaso. Além disso, essa correlação não se mostra significativa ( $N = 26, p > 0,05$ ).

Tabela 17: Correlações - Spearman

		TLTOMesmo Média	TLSPOMesmo Média	TLSPEle Média
TotalTCL Local	Correlação	-0,073	0,072	-0,050
	Significância	0,722	0,728	0,808
	N	26	26	26

Fonte: a autora.

Com os testes dos postos de Spearman, correlacionamos o desempenho em questões locais do teste de compreensão leitora com as médias dos tempo de leitura cuja distribuição não é aproximadamente normal. Como resultado, obtivemos uma correlação fraca, negativa e não significativa com a média do tempo total de leitura das frases com o sintagma “o mesmo”. Isso significa que, assim como ocorreu com a média do tempo total de leitura do pronome, a tendência é que quanto maior a pontuação nas questões locais, menor o tempo de leitura das frases construídas com o sintagma “o mesmo”.

Há, também, uma correlação fraca e negativa entre o desempenho local e a média do tempo de leitura do segmento pós-anáfora realizada com “ele”. Já a correlação entre o

desempenho local e a média do tempo de leitura do segmento pós-anáfora realizada com o sintagma “o mesmo” revelou uma correlação fraca, positiva e não significativa.

De modo geral, temos que tanto os coeficientes de correlação de Pearson quanto os coeficientes de correlação de postos de Spearman não revelaram resultados estatisticamente significativos para a associação das variáveis testadas ( $N = 26$ ,  $p > 0,05$ ).

Esses resultados nos levam a refletir acerca de alguns pontos. Primeiro, quando tratamos da correlação com a média do tempo total de leitura, percebemos uma correlação negativa em ambos os anafóricos testados. Esse resultado era o esperado, já que entendemos que, quanto mais rápida a leitura é feita, menores são os custos para o processamento e melhor é o desempenho do leitor nessas questões mais locais do texto, que requerem automatização desses processos de integração. No entanto, tal correlação foi fraca e não significativa, o que pode ter ocorrido pelo  $N$  amostral do estudo e pela natureza distinta de cada uma das variáveis testadas (média de tempo de leitura e desempenho em questões de teste de compreensão).

Outro aspecto intrigante é a correlação positiva entre o desempenho nas questões locais e a média do tempo de leitura do segmento pós-anáfora realizada com “o mesmo”. Uma correlação positiva mostra que quanto maior o tempo de leitura, maior o índice de acertos nas questões locais e que quanto menor o tempo de leitura, menor o índice de acerto. Como um resultado positivo aparece apenas quando o tempo de leitura do segmento pós-anafórico entra na correlação, é possível que tenha ocorrido devido a questões intrínsecas aos testes em si. De todo modo, precisamos de mais indícios para poder discorrer sobre o sintagma “o mesmo” como anáfora correferencial e essa integração mais local. Quando tratarmos das estratégias utilizadas nos protocolos verbais, sobretudo diante de cada forma anafórica estudada, teremos mais recursos para levantar outras asserções. Antes, contudo, vamos discorrer sobre os resultados relativos à aceitabilidade das frases com o sintagma “o mesmo” e das frases com o pronome “ele”.

#### **4.1.3 Qual a aceitabilidade do uso de “o mesmo” como elemento anafórico correferencial em construções sentenciais do PB?**

Outro objetivo do nosso estudo estava em identificar a aceitabilidade do sintagma “o mesmo” comparando com a aceitabilidade do pronome “ele”. A nossa hipótese era a de que a aceitabilidade das frases cuja anáfora foi realizada com o sintagma “o mesmo” não seria diferente da aceitabilidade das frases que continham anáforas pronominais. Como vimos discutindo ao longo do nosso trabalho, o sintagma “o mesmo” como anáfora correferencial

ainda gera questionamentos. Então, para além de investigar o processamento dessa anáfora, buscamos pela sua aceitabilidade entre os acadêmicos que participaram do estudo. Para lembrar, no caso do teste de aceitabilidade, foram 26 participantes avaliando 6 conjuntos de frases com o pronome e 6 conjuntos de frases com o sintagma, um dos participantes acabou deixando em branco uma das avaliações de uma frase com o sintagma. Para a rodada estatística, retiramos o dado faltante. A seguir, reportamos a tabela de frequências:

Tabela 18: Estatística Descritiva – Teste de Aceitabilidade

Medida	O mesmo	Ele
<b>Média</b>	5,43	5,73
<b>Mínimo</b>	2	2
<b>Máximo</b>	7	7
<b>Desvio Padrão</b>	1,49	1,47

Fonte: a autora.

A estatística descritiva revela que a média de aceitabilidade das frases realizadas com o pronome “ele” foi maior que a média de aceitabilidade das frases realizadas com o sintagma “o mesmo”. Nos resultados estatísticos, poderemos confirmar se tal diferença foi ou não significativa. O que podemos conjecturar de início é que em uma escala que ia de 1 a 7, as frases com o anafórico “o mesmo” tiveram por média 5,43 pontos de aceitabilidade e as frases realizadas com o pronome obtiveram média de 5,73. Ambas as condições ficaram acima do ponto neutro da escala (4), mas, ao que parece, as frases com “ele” estão mais próximas do 6, resultado possível de visualizar no teste de significância que será apresentado mais adiante. Para refletir um pouco mais sobre esses valores, reportamos a seguir a frequência de cada ponto da escala para cada uma das anáforas:

Tabela 19: Teste de Aceitabilidade – Frequência “O mesmo”

Escala	Frequência	Percentual
2	8	5,2
3	13	8,4
4	18	11,6
5	29	18,7
6	40	25,8
7	47	30,3

Fonte: a autora.

Tabela 20: Teste de Aceitabilidade – Frequência “Ele”

Escala	Frequência	Percentual
2	7	4,5
3	10	6,5
4	11	7,1
5	20	19,4
6	29	18,7
7	68	43,9

Fonte: a autora.

O primeiro ponto a se considerar é o fato de o índice de aceitabilidade se concentrar no topo da escala. Ou seja, tanto a avaliação das frases com o pronome quanto das frases com “o mesmo” se manteve acima do ponto neutro (4), como já havíamos verificado ao reportar a média. No entanto, nas frases com o pronome “ele”, há mais ocorrências de “totalmente aceitável”. Se considerarmos a parte superior da escala (5, 6 e 7) em conjunto, temos 127 ocorrências para as frases com “ele”, comparadas a 116 para as frases com “o mesmo”. Então, ao que parece, os acadêmicos, ainda que processem de modo mais rápido as frases construídas com o sintagma “o mesmo” (conforme verificado no teste de leitura automonitorada), preferem, em uma avaliação deliberada, as frases construídas com o pronome.

A seguir apresentamos os resultados obtidos por meio dos Testes dos Postos Sinalizados de Wilcoxon:

Tabela 21: Teste de Aceitabilidade - Postos

		N	Média dos Ranks	Soma dos Ranks
<b>Aceitabilidade Ele</b>	Ranques Negativos	43 <sup>a</sup>	55,23	2375
	Ranques Positivos	69 <sup>b</sup>	57,29	3953
	Empates	43 <sup>c</sup>		
	<b>Total</b>	155		
<b>Aceitabilidade “O mesmo”</b>				
<sup>a</sup> AceitabilidadeEleEla < AceitabilidadeOMesmoAMesma				
<sup>b</sup> AceitabilidadeEleEla > AceitabilidadeOMesmoAMesma				
<sup>c</sup> AceitabilidadeEleEla = AceitabilidadeOMesmoAMesma				

Fonte: a autora.

A aceitabilidade das frases com o pronome foi maior em 69 casos e menor em 43 casos. Dessa forma, a aceitabilidade de “ele” é menor em muito menos da metade dos casos. O que mostra que “o mesmo”, ainda que seja menos aceitável que o pronome, não é considerado inaceitável pelos participantes. Ao analisar a parte superior da escala, temos 87 ocorrências de aceitabilidade 6 e 7 (aceitável e totalmente aceitável) para “o mesmo” e 97 ocorrências para

“ele”. Para entender se tal diferença é ou não significativa, realizamos o teste estatístico de Wilcoxon:

Tabela 22: Estatística do teste de aceitabilidade - Testes dos Postos Sinalizados de Wilcoxon

	AceitabilidadeEleEla - AceitabilidadeOMesmoAMesma
Z	-2,33 <sup>a</sup>
Asymp. Sig. (2-tailed)	0,01
<sup>a</sup> Com base nos ranks negativos	

Fonte: a autora.

Os Testes dos Postos Sinalizados de Wilcoxon indicaram que os postos relativos à aceitabilidade das frases com o pronome “ele” são estatisticamente maiores do que os postos da aceitabilidade das frases com “o mesmo”,  $Z = -2,33$ ,  $p = 0,01$ . A diferença, ainda que não seja grande, é, portanto, estatisticamente significativa. Esse resultado deve, contudo, ser analisado de modo cauteloso, considerando algumas questões, explanadas na sequência.

O que os resultados obtidos por meio desse teste mostram é que a aceitabilidade de “ele” como anafórico tende a ser maior do que a aceitabilidade de “o mesmo”. Há que se considerar, no entanto, que esse teste fornece resultados de uma avaliação deliberada, isto é, a partir da reflexão do participante acerca das porções textuais a que foi exposto. Como já discutido no capítulo de fundamentos, na língua portuguesa, o pronome “ele” é o anafórico característico. De acordo com Garnham (2001), esse pronome tende a ser o usado quando não há mudança de tópico, sendo o emprego do pronome mais frequente, é compreensível que, de modo deliberado, estejamos mais habituados a ele na leitura e na oralidade. Além do mais, há que se considerar o público participante desta investigação: estudantes universitários, ou seja, participantes que já passaram por diferentes níveis de escolarização e que podem ter se deparado, ao longo de sua vida acadêmica, com o estigma que ronda a anáfora correferencial construída com “o mesmo”.

Ademais, ainda que “o mesmo” seja usado em contextos nos quais os pronomes poderiam ser mobilizados, não possui características pronominais, o que se mostra verdadeiro ao constatar que ele sempre vem definido pelo artigo (o mesmo/a mesma). Isso mostra que “o mesmo” não se caracteriza como pronome, pelo menos não nessa construção. Trata-se de uma nominalização, precedida de um artigo definido. Assim, no nível lexical (em sua formação estrutural), esse elemento se comporta como substantivo. No nível sintático, quando ele assume a função de anáfora correferencial, comporta-se como um pronome, porque o antecedente precisa ser retomado para o referente ser resgatado. Então, ele não é um pronome, mas tem se



comportado como tal neste tipo de uso. É possível que esses fatores possam ter influenciado na avaliação, pelos participantes, da aceitabilidade do sintagma.

Em um nível metacognitivo, temos indícios de que a forma anafórica preferencial é aquela realizada com “ele” e “ela”, ainda que a realização anafórica com “o mesmo” e “a mesma” não seja considerada inaceitável. Isso em uma avaliação consciente, mas vimos, com o teste de leitura automonitorada (subseção 4.1.1), que não há diferenças significativas no processamento de “ele” e de “o mesmo”. Entendemos que, como não é considerada uma construção inaceitável, é compreensível que não gere entraves no processamento. Resta refletir sobre o porquê de, numa avaliação deliberada, o pronome ser significativamente mais aceito e, no processamento, não haver diferenças significativas. Ademais, o fato de o uso de “o mesmo” como anafórico não ser considerado inaceitável e o processamento ser tão ou mais rápido que das frases com “ele” levam a crer que esse sintagma não gera entraves para a compreensão em leitura e que seja tão efetivo quanto as anáforas realizadas com ele/ela, favorecendo a integração local entre as partes do texto e a compreensão. De todo modo, a hipótese que levantamos acerca da aceitabilidade não pôde ser validada com os resultados obtidos, ainda que o sintagma “o mesmo” seja também aceito, obtivemos diferenças com relação ao pronome. Conjecturamos, contudo, que os pontos levantados anteriormente ajudem a explicar essa diferença obtida.

Para discutir pontos acerca da compreensão com mais propriedade trazemos, na sequência, os resultados do comportamento estratégico diante da leitura do texto.

#### **4.1.4 Qual a relação entre o comportamento geral, o comportamento diante da anáfora correferencial construída com o sintagma “o mesmo” e o desempenho diante das retomadas anafóricas e do texto?**

Para responder a essa pergunta, lançamos mão dos dados relativos ao teste de compreensão leitora, obtidos de duas formas: 1. a partir do desempenho dos participantes nas respostas às questões do teste; 2. do comportamento estratégico adotado diante do texto quando o objetivo estava em responder às questões. Esses resultados nos auxiliam a perceber de que forma o sintagma “o mesmo” influencia no processamento e na compreensão do texto e nos dão, ainda, indícios do nível de desempenho dos participantes no que tange a essa compreensão.

Iniciemos a análise refletindo sobre a pontuação obtida pelos participantes no teste de compreensão leitora. Temos o seguinte cenário:

Tabela 23: Estatística Descritiva – Teste de Compreensão Leitora

Medida	TCL Geral	TCL Local
<b>Média</b>	6,66	3,70
<b>Mínimo</b>	4,40	2,50
<b>Máximo</b>	8,30	4,90
<b>Desvio Padrão</b>	0,19	0,65

Fonte: a autora.

O que esses resultados apontam é que a média de acertos do teste ficou próximo de 70%, já que a pontuação total somava 10 pontos (eram 10 questões e cada uma delas valia 1 ponto) e a média ficou em 6,66. Nenhum dos participantes atingiu pontuação máxima e nenhum deixou de pontuar. O desempenho mais elevado foi de 8,3 e o mais baixo, de 4,4. Dos 26 participantes, 14 obtiveram desempenho acima da média, enquanto 10 ficaram abaixo da média e 2 na média.

Embora a pergunta 4 tenha por foco correlações, é importante que tenhamos, também, uma noção do desempenho individual desses participantes para que o leitor possa compreender melhor o perfil da amostra. Para melhor visualização, apresentamos a seguir o desempenho individual de cada um dos participantes:

Tabela 24: Desempenho Individual no Teste de Compreensão Leitora

Participante	Desempenho
<b>P01</b>	5,6
<b>P02</b>	7,8
<b>P03</b>	6
<b>P04</b>	7
<b>P05</b>	7,8
<b>P06</b>	7,3
<b>P07</b>	7
<b>P08</b>	6,2
<b>P09</b>	5,5
<b>P10</b>	7,8
<b>P11</b>	6,9
<b>P12</b>	6,5
<b>P13</b>	6,8
<b>P14</b>	8,2
<b>P15</b>	5,5
<b>P16</b>	6
<b>P17</b>	6,6
<b>P18</b>	7,9
<b>P19</b>	7,4
<b>P20</b>	8,3
<b>P21</b>	6,1
<b>P22</b>	4,4
<b>P23</b>	6,7

<b>P24</b>	5,5
<b>P25</b>	5,3
<b>P26</b>	7,3

Fonte: a autora.

Como é possível visualizar na tabela, os participantes que tiveram maior desempenho no teste foram P20, P14, P18, P02 e P05. Já os menores desempenhos foram de P22, P25, P09, P15 e P24. Podemos comparar esses resultados àqueles obtidos nas questões que tinham por foco processos locais de integração e localização. Essa comparação é relevante visto que o nosso objeto de estudo (a anáfora) atua nesse nível de processamento, que, como discutimos na fundamentação teórica, é essencial para que processos superiores possam acontecer e para que a compreensão seja construída, ainda que o processamento nesses distintos níveis não ocorra de modo linear, mas em paralelo, de modo interativo (KINTSCH, 1998; KINTSCH; RAWSON, 2013; JUST; CARPENTER, 2013; RUMELHART, 2013), ou seja, é necessário que todos os níveis linguísticos sejam processados, mas tais processos não ocorrem, necessariamente, de forma sequencial em leitores proficientes. Contudo, no caso de um teste de compreensão, direcionamos, a partir das perguntas, a atenção do leitor para um ou outro nível, a fim de conseguir obter dados específicos de seu desempenho de acordo com os objetivos do estudo.

No caso do nosso grupo amostral, a pontuação máxima a que eles chegariam nas questões locais era 5, a média ficou em 3,7, sendo que o maior desempenho chegou a 4,9 e o menor em 2,5. A seguir, apresentamos a tabela com o desempenho individual nessas questões que requeriam processos cognitivos locais para serem respondidas:

Tabela 25: Desempenho Individual no Teste de Compreensão Leitora - Local

<b>Participante</b>	<b>Desempenho</b>
<b>P01</b>	3,1
<b>P02</b>	3,8
<b>P03</b>	2,5
<b>P04</b>	4
<b>P05</b>	4,3
<b>P06</b>	4,3
<b>P07</b>	4
<b>P08</b>	3,2
<b>P09</b>	4
<b>P10</b>	4,3
<b>P11</b>	4,4
<b>P12</b>	3
<b>P13</b>	3,3
<b>P14</b>	4,2
<b>P15</b>	3

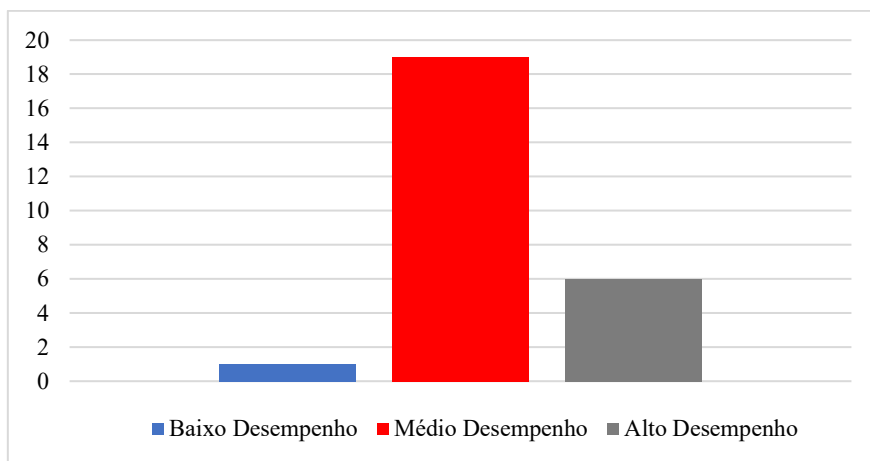
<b>P16</b>	3
<b>P17</b>	3,1
<b>P18</b>	4,9
<b>P19</b>	4,4
<b>P20</b>	4,3
<b>P21</b>	4,1
<b>P22</b>	3,4
<b>P23</b>	3,2
<b>P24</b>	2,5
<b>P25</b>	3,8
<b>P26</b>	4,3

Fonte: a autora.

Aqui, as maiores pontuações foram obtidas por P18, P11, P19 e P05, P06, P10, P20 e P26 (os cinco últimos com mesmo desempenho). Os participantes P18 e P20 se repetem, ou seja, os escores obtidos por eles, tanto nas questões locais quanto no teste todo, estiveram entre os mais altos. Por sua vez, os menores escores obtidos nas questões locais foram de P24, P03, P16, P15 e P12. Nesse caso, o único participante com menor desempenho que se repete é o P15. Isso pode significar que os demais participantes compensaram de outras formas suas dificuldades com as questões locais e acabaram melhorando seu desempenho geral no teste.

Para que seja possível visualizar o desempenho dos participantes, criamos categorias a fim de agrupar os resultados obtidos. Tais categorias partem do que Schlichting (2021) estabeleceu em seu trabalho, a saber: a) Baixo desempenho (com escores de 0,0 a 5,0); b) Médio desempenho (de 5,1 a 7,5) e; c) Alto desempenho (de 7,6 a 10,0). Partimos dessa autora, pois ela toma por base as referências do PISA, fazendo as devidas adequações às normativas de médias brasileiras, o que corresponde aos nossos delineamentos acerca do desempenho obtido pelos participantes desta pesquisa:

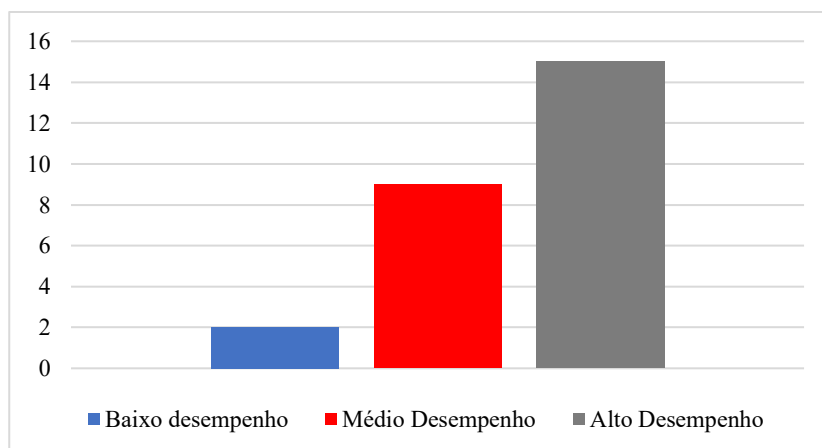
Gráfico 3: Desempenho Geral Categorizado



Fonte: a autora.

O que o teste de compreensão revela é que a maioria dos participantes atinge um nível mediano de compreensão do texto. Proporcionalmente aos valores estabelecidos para o desempenho geral no teste, estabelecemos valores para o desempenho nas questões locais, a saber: a) Baixo desempenho (com escores de 0,0 a 2,5); b) Médio desempenho (de 2,6 a 3,7) e; c) Alto desempenho (de 3,8 a 5,0).

Gráfico 4: Desempenho Local Categorizado



Fonte: a autora.

No que diz respeito ao desempenho local, é possível visualizar que a maioria dos participantes tem um desempenho alto. A partir desses resultados iniciais, podemos aventar que o fato de a maioria dos participantes ter obtido um desempenho mediano no teste não se dá por conta das questões que exigem identificação, localização e integrações no nível local, mas por conta de questões que exigem mais inferências e análises. Nesse sentido, ao que parece, o uso

da anáfora realizada com o sintagma “o mesmo” não se mostra como um complicador no desempenho dos participantes.

Para dar sequência às análises visando responder à questão relativa a essa subseção, trazemos os resultados referentes às médias de acertos dos participantes no teste de compreensão leitora, tanto geral quanto local.

Tabela 26: Estatística Descritiva – Teste de Compreensão Leitora Média

Medida	TCL Geral	TCL Local
<b>Média</b>	0,66	0,74
<b>Mínimo</b>	0,44	0,50
<b>Máximo</b>	0,83	0,98
<b>Desvio Padrão</b>	0,019	0,025

Fonte: a autora.

A partir dos resultados obtidos com as médias dos participantes, foram realizados os testes de normalidade, cujos resultados são apresentados a seguir:

Tabela 27: Normalidade – Teste de Compreensão Leitora – Média

	Shapiro-Wilk		
	Estatística	df	Sig.
TCL Geral	0,96	26	0,58
TCL Local	0,92	26	0,06

Fonte: a autora.

Os resultados obtidos com os Testes de Shapiro-Wilk demonstram que os dados seguem uma distribuição aproximadamente normal para as variáveis TCLGeral e TCLLocal.

Ainda para responder à pergunta 4, partimos de correlações feitas com as categorias do comportamento estratégico dos participantes durante a realização do teste de compreensão. Como vimos no capítulo em que abordamos o método da pesquisa, obtivemos seis categorias: retomada, seleção, busca, destaque, anotação e *scanning*. A estatística descritiva é a seguinte:

Tabela 28: Estatística Descritiva – Monitoramento Teste de Compreensão Leitora

Medida	Retomada	Seleção	Busca	Destaque	Anotação	Scanning
<b>Média</b>	87,73	8,27	1,88	1,38	0,88	0,27
<b>Mínimo</b>	32	0	0	0	0	0
<b>Máximo</b>	153	44	10	10	16	1
<b>Desvio Padrão</b>	6,68	2,08	0,52	0,57	0,63	0,89

Fonte: a autora.

Com a estatística descritiva, já é possível visualizar que o comportamento estratégico mais adotado foi o de retomada do texto, usado repetidas vezes por todos os participantes. Já todas as outras ações ocorreram em bastante menor número. A retomada frequente do texto pode ser explicada por diversos fatores, que extrapolam os objetivos desta pesquisa e, por isso, não foram confirmados, a saber: estratégia para responder às questões de modo mais propositivo (o que faz com que os participantes olhem primeiro o enunciado para, então, buscar a resposta, sem fazer leitura prévia do texto ou fazendo uma leitura prévia superficial), baixa capacidade da memória de trabalho, tendo dificuldades para manipular informações recentemente obtidas, falta de atenção etc. Todos esses são fatores que podem ter influenciado nesse comportamento estratégico específico.

No que diz respeito à normalidade dos dados, obtivemos o seguinte:

Tabela 29: Normalidade – Monitoramento Teste de Compreensão Leitora

	Shapiro-Wilk		
	Estatística	df	Sig.
Retomada	0,95	26	0,26
Seleção	0,76	26	0,00
Busca	0,70	26	0,00
Destaque	0,54	26	0,00
Anotação	0,30	26	0,00
<i>Scanning</i>	0,55	26	0,00

Fonte: a autora.

O que os resultados obtidos por meio dos Testes de Shapiro-Wilk demonstram é que os dados seguem uma distribuição aproximadamente normal para a variável “retomada”. No entanto, as variáveis “seleção”, “busca”, “destaque”, “anotação” e “*scanning*” não seguem uma distribuição aproximadamente normal.

No que diz respeito ao comportamento estratégico dos participantes diante da anáfora “o mesmo”, temos a ocorrência das seguintes categorias: retomada, seleção, busca e destaque. Ou seja, os participantes não se valem de todas as estratégias utilizadas no texto como um todo, deixando de fora anotação e *scanning*. O que é compreensível já que a retomada da anáfora correferencial sugere o uso de estratégias mais específicas, que se voltem para a integração local ou a confirmação dessa integração. A estatística descritiva é apresentada a seguir:

Tabela 30: Estatística Descritiva – Monitoramento Teste de Compreensão Leitora Local

Medida	Retomada	Seleção	Busca	Destaque
<b>Média</b>	20,5	2,69	0,54	0,15
<b>Mínimo</b>	10	0	0	0
<b>Máximo</b>	41	14	3	2
<b>Desvio Padrão</b>	1,6	0,79	0,19	0,091

Fonte: a autora.

No que tange ao comportamento diante da anáfora realizada com o sintagma “o mesmo”, temos um padrão semelhante ao que acontece no texto como um todo, sendo a “retomada” o comportamento estratégico mais recorrente. Nesse caso, anotação e *scanning* não são estratégias utilizadas. A seguir, é apresentado o teste de normalidade desses dados:

Tabela 31: Normalidade – Monitoramento Teste de Compreensão Leitora Local

	Shapiro-Wilk		
	Estatística	df	Sig.
Retomada	0,90	26	0,24
Seleção	0,71	26	0,00
Busca	0,58	26	0,00
Destaque	0,38	26	0,00

Fonte: a autora.

No caso do comportamento estratégico adotado diante da anáfora “o mesmo”, os resultados obtidos com os Testes de Shapiro-Wilk demonstram que os dados seguem uma distribuição aproximadamente normal para a variável “retomada”. Contudo, não seguem uma distribuição aproximadamente normal para todas as demais variáveis, a saber: seleção, busca e destaque.

A partir, então, desses resultados, rodamos os testes de correlação a fim de relacionar o comportamento estratégico dos participantes diante do texto e, especificamente, diante da anáfora construída com o sintagma “o mesmo”. O teste de Pearson foi adotado para correlacionar a variável “retomada geral” e “retomada ‘o mesmo’”, já que ela apresenta uma distribuição normal. A seguir, apresentamos os resultados:



Tabela 32: Correlações de Pearson do Mapeamento do Comportamento Geral x Mapeamento do Comportamento “O Mesmo”

		<b>Retomada “O mesmo”</b>
<b>Retomada Geral</b>	Correlação de Pearson	0,74
	Sig.	0,00
	N	26

Fonte: a autora.

Os Coeficientes de Correlação de Pearson encontraram, para retomada geral x retomada “o mesmo”, uma correlação positiva, forte e estatisticamente significativa ( $N = 26$ ,  $p = 0,00$ ). Os valores obtidos sinalizam um tamanho de efeito de 0,55, o que mostra que há 55% de variabilidade dos dados na correlação das duas variáveis. Isso significa dizer que, quando o comportamento estratégico de “retomada” sobe para o texto como um todo, sobe, também, diante da anáfora “o mesmo”. Esse resultado sinaliza que o uso da retomada diante do sintagma “o mesmo” pode não ser por conta de uma dificuldade inerente à anáfora. Além disso, algumas das questões que solicitavam integração local, que envolviam a anáfora realizada com “o mesmo”, eram questões de localização, o que requeria que os participantes voltassem ao texto.

Vejamos, a partir das Correlações de Postos de Spearman, os resultados obtidos para as demais variáveis, cuja distribuição não é normalmente distribuída:

Tabela 33: Correlações de Postos de Spearman do Mapeamento do Comportamento Geral x Mapeamento do Comportamento “O Mesmo”

		<b>Seleção Geral</b>	<b>Busca Geral</b>	<b>Destaque Geral</b>
<b>Seleção “O mesmo”</b>	Correlação	0,87	0,10	-0,21
	Sig.	0,00	0,62	0,29
	N	26	26	26
<b>Busca “O mesmo”</b>	Correlação	0,10	0,66	0,16
	Sig.	0,60	0,00	0,43
	N	26	26	26
<b>Destaque “O mesmo”</b>	Correlação	0,19	0,22	0,58
	Sig.	0,33	0,27	0,002
	N	26	26	26

Fonte: a autora.

As Correlações dos Postos de Spearman mostram uma relação forte, positiva e estatisticamente significativa entre a seleção do sintagma “o mesmo” e a seleção geral. Esse resultado indica que, conforme a seleção do texto sobe, a seleção da anáfora também sobe. O que, mais uma vez, assim como na retomada, sinaliza para um comportamento estratégico que

não diz respeito a uma estratégia inerente ao sintagma “o mesmo”, pois é algo que ocorre de modo generalizado no texto. Já no que diz respeito à busca, temos, também, uma correlação positiva, estatisticamente significativa e moderada (quase forte). Então, mais uma vez, quando os participantes utilizam a ferramenta de busca do pdf, utilizam de modo correlato para o texto e para o sintagma “o mesmo”. Quanto ao destaque, esse padrão se mantém.

Tudo isso traz subsídios que corroboram o fato de o sintagma “o mesmo” atuar como anáfora correferencial sem que isso demande mais estratégias dos participantes ou estratégias que sejam diferentes daquelas utilizadas na leitura do texto como um todo. Além disso, vimos, no capítulo de fundamentação teórica, que leitores proficientes tendem a ser ativos e estratégicos, então, ao que parece, os leitores usam de comportamentos estratégicos que lhes auxiliam a alcançar o objetivo com a leitura do texto – responder ao teste de compreensão – e essas estratégias não parecem refletir problemas com a anáfora em si.

Na sequência, são apresentados os resultados das correlações entre o comportamento estratégico diante do texto como um todo e o desempenho dos participantes no teste de compreensão:

Tabela 34: Correlações de Pearson do Mapeamento do Comportamento Geral x Desempenho Compreensão Leitora Geral

		<b>Retomada</b>	<b>Seleção Geral</b>	<b>Busca Geral</b>	<b>Destaque Geral</b>	<b>Anotação Geral</b>	<b>Scanning</b>
<b>Total TCL</b>	Cor.	0,036	-0,14	-0,004	0,093	0,18	-0,32
	Sig.	0,82	0,47	0,98	0,65	0,36	0,11
	N	26	26	26	26	26	26

Fonte: a autora.

Os Coeficientes de Correlação de Pearson não demonstraram nenhuma correlação estatisticamente significativa entre o mapeamento do comportamento estratégico geral e o desempenho geral na compreensão leitora para nenhuma das variáveis investigadas (N = 26, p > 0,05).

Tabela 35: Correlações de Postos de Spearman do Mapeamento do Comportamento “O Mesmo” x Desempenho Compreensão Leitora Local

		<b>Retomada</b>	<b>Seleção</b>	<b>Busca</b>	<b>Destaque</b>
<b>Questão 2</b>	Cor.	-0,28	-0,45	0,02	-0,17
	Sig.	0,15	0,02	0,89	0,38
	N	26	26	26	26
<b>Questão 4</b>	Cor.	-0,18	-0,19	0,27	-0,25
	Sig.	0,35	0,34	0,17	0,21
	N	26	26	26	26
<b>Questão 6</b>	Cor.	-0,22	-0,11	0,01	-0,001
	Sig.	0,27	0,57	0,95	0,99
	N	26	26	26	26
<b>Questão 8</b>	Cor.	0,24	0,20	0,30	-0,008
	Sig.	0,22	0,30	0,13	0,97
	N	26	26	26	26
<b>Questão 9</b>	Cor.	-0,048	0,09	0,51	0,28
	Sig.	0,81	0,64	0,07	0,15
	N	26	26	26	26

Fonte: a autora.

Os Coeficientes de Correlação de Postos de Spearman revelaram que: há uma correlação estatisticamente significativa, negativa e moderada para a associação entre as variáveis **seleção do sintagma “o mesmo” e questão 2** ( $N = 26$ ,  $p < 0,05$ ), com um tamanho de efeito de 0,20 relativo à covariância dos dados. Na questão 2, o participante precisava integrar duas partes do texto. Tratava-se de uma integração local em que deveria ser feita uma correlação adequada entre a anáfora realizada com o sintagma “o mesmo” e seu antecedente. Então, o resultado que obtivemos mostra o seguinte: quanto maior o desempenho do participante na questão 2, menor é o emprego de seleção do sintagma “o mesmo”. Diante disso, selecionar o sintagma que fazia a integração local não se mostrou uma estratégia adequada para conseguir responder à questão de modo adequado. Há, ainda, 20% de variabilidade dos dados na correlação das duas variáveis.

Ademais, os Coeficientes de Correlação de Postos de Spearman revelaram uma correlação estatisticamente significativa, positiva e moderada para as variáveis **busca do sintagma “o mesmo” e questão 9** ( $N = 26$ ,  $p < 0,05$ ), com um tamanho de efeito de 0,20, que se relaciona à covariância dos dados. Ou seja, quanto mais é feita a busca pelo sintagma, maior a nota da questão 9. Os resultados dos testes também demonstraram que não há nenhuma correlação estatisticamente significativa para a associação entre as demais variáveis.

Pensando nos resultados obtidos, vale refletir sobre o solicitado pela questão 2 do teste de compreensão. Nessa questão, os participantes precisavam voltar para o texto e integrar o

referente e a anáfora realizada com o sintagma “o mesmo” para que pudessem explicar a “semântica de falhas”, ponto solicitado pela questão. Assim, é possível aventar que, para confirmar a integração entre as partes do texto, os participantes selecionavam a referida anáfora. No entanto, a correlação se mostra negativa, o que revela um fato interessante, isto é, quanto mais selecionavam, menor o escore na questão e quanto menos selecionavam, maior o escore na questão. Dessa forma, a localização da anáfora era importante, a identificação adequada do referente era essencial, mas, feito isso, o participante precisava conseguir integrar as duas sentenças para poder definir “semântica de falhas”. Assim, a estratégia de selecionar o sintagma não se mostra efetiva para conseguir integrar as diferentes partes do texto e poder definir a expressão solicitada de modo adequado.

Já a questão 9 solicitava, explicitamente, pela busca das anáforas realizadas com o sintagma “o mesmo”; então, os participantes se valiam da ferramenta de busca para poder encontrá-las mais rapidamente e responder ao solicitado. A correlação positiva aqui demonstra que a estratégia funcionou, pois quanto mais a ferramenta de busca foi utilizada, maior o escore na questão e quanto menos utilizada, menor o escore. Dessa forma, tendo em vista que essa questão solicitava, explicitamente, ao participante que retomasse duas anáforas realizadas com o sintagma “o mesmo” e que explicasse as informações apresentadas naquele trecho, era esperado que a ferramenta de busca fosse utilizada e que isso aumentasse o índice de acerto à questão.

Para fechar nossas reflexões acerca dessa relação entre nível micro (integração local por meio das anáforas) e macro (formação de uma base textual coerente) do texto, trazemos, a seguir, os resultados acerca das estratégias que os leitores utilizaram durante as sessões de protocolo verbal.

#### **4.1.5 Que estratégias de leitura são utilizadas pelo leitor e qual a relação de tais estratégias com o desempenho diante das retomadas anafóricas e do texto? Ainda, é possível, no que diz respeito às estratégias empregadas pelos participantes, observar semelhanças frente às diferentes anáforas lidas?**

A nossa quinta e última pergunta de pesquisa está subdividida em duas. Então, primeiro, falaremos das estratégias que os leitores participantes da pesquisa utilizaram na sessão de leitura do texto durante os protocolos verbais. A partir disso, faremos as correlações com o desempenho nas questões do teste de compreensão leitora relativas às anáforas e, também, com o desempenho geral. Feito isso, olharemos para as estratégias empregadas diante das anáforas

realizadas com o sintagma “o mesmo” e com o pronome “ele”. Esses resultados fortalecerão nossas bases acerca do processamento de “o mesmo” como anáfora correferencial e do produto da compreensão de textos que contêm tais anáforas.

Para responder a essa pergunta, iniciamos com a catalogação das estratégias utilizadas pelos participantes da pesquisa durante as seções de Protocolo Verbal. Chegamos a um total de oito estratégias diferentes, a saber: Releitura, Monitoramento da Compreensão, Conexões Anafóricas Intersentenciais, Relato Episódico, Leitura em voz alta, Avaliação do texto, Elaboração e Paráfrase. Como explicamos na seção 3.7, as categorias de estratégias surgiram a partir das verbalizações dos próprios participantes, mas considerando, também, o que a literatura apresenta acerca de estratégias (AFFLERBACH; PEARSON; PARIS, 2008; KENDEOU; BROEK; 2007; LESLIE; CALDWELL, 2011; SOLÉ, 1998;). Para retomar, sabemos que um leitor que “é estratégico, utiliza das estratégias para chegar a um objetivo” (AFFLERBACH; PEARSON; PARIS, 2008, p.366). No caso desta pesquisa, o objetivo-alvo durante os protocolos era: os acadêmicos deveriam ler com o objetivo de compreender o texto como um todo. A partir disso, cada participante se valeu de ações específicas para construir o caminho na direção da compreensão do texto.

Existem diferentes estratégias que podem ser usadas no processo de leitura. Os leitores proficientes sabem selecioná-las e usá-las para alcançar aquilo que desejam. Não se pode dizer que existem, de modo intrínseco, estratégias melhores ou piores, mas há estratégias que podem funcionar ou não em um determinado contexto, com determinado objetivo e para um determinado leitor. Cabe ao leitor saber qual delas é a mais adequada. Para entendermos melhor cada uma das categorias de estratégias que identificamos, vale retomar as explicações e aprofundá-las de modo a deixar mais claro o que entendemos com cada uma dessas estratégias adotadas.

A releitura é uma das estratégias utilizadas no momento em que o leitor encontra algum entrave na leitura, quando percebe que o que leu não faz sentido ou, então, quando quer se certificar daquilo que compreendeu. Essa estratégia costuma auxiliar, ainda, em problemas nos níveis sintáticos e prosódicos (ALLINGTON; MCGILL-FRANZEN, 2011). Essa foi uma estratégia utilizada por 16 participantes, ou seja, 61,5% dos participantes precisaram reler algum trecho do texto em algum momento da sessão de protocolo verbal.

O monitoramento da compreensão é, por alguns autores (DOLE; NOKES; DRITS, 2011), também chamado de estratégia metacognitiva. Esse tipo de estratégia é particularmente relevante para a compreensão, porque o leitor acompanha a sua leitura, buscando autoafirmar ou se autoquestionar acerca do que está lendo e do que está compreendendo. Nesse processo,

os leitores costumam formular hipóteses e ir fazendo a checagem durante o progresso no texto. Quando há o monitoramento da compreensão, existem mais chances de o leitor perceber quando uma falha ocorre e, então, acionar alguma outra estratégia para compensar essa falha. Dos 26 participantes da pesquisa, 18 se valeram dessas estratégias, ou seja, 69%. Isso não significa que os outros participantes não tenham se valido do monitoramento da compreensão, mas não demonstraram isso em seus relatos verbais. Essa estratégia tende a se mostrar um importante preditor da compreensão (BAKER; BEALL, 2011), porque revela um leitor que está atento àquilo que compreende ou não.

As conexões anafóricas intersentenciais dizem respeito a verbalizações relativas a diferentes sentenças, ou seja, o participante, para compreender determinados trechos, precisava fazer as conexões entre sentenças. 18 participantes expõem verbalmente essas conexões, ou seja, 69,23% dos participantes verbalizam, de algum modo, acerca desse processamento, o que revela o processo de construção de um modelo mental, que é formado conforme o texto avança (GARNHAM, 2001). Então, ao se valerem dessa estratégia, os participantes estão integrando as informações provindas de diferentes sentenças. Como a compreensão ocorre de modo processual, o uso dessa estratégia revela esse processo por parte dos participantes.

O relato episódico é uma estratégia que pode ser considerada como fuga do texto a depender de como o leitor a utiliza. No caso desta pesquisa, alguma informação contida no texto leva os participantes a lembrarem de algum professor ou de alguma situação específica que viveram. Essa estratégia pode ser útil se o leitor consegue retomar o texto e, a partir da lembrança, formular uma representação coerente do que está lendo, fazendo conexões entre seu conhecimento prévio e as informações apreendidas durante a leitura. Do contrário, pode ser uma estratégia que acabe interferindo nos processos cognitivos de compreensão, por mudar o foco de atenção (STERNBERG; STERNBERG, 2017) do texto. 6 participantes utilizaram dessa estratégia, ou seja, 23,07%.

A leitura em voz alta, por sua vez, quando feita por leitores proficientes é uma estratégia usada para obter um retorno auditivo daquilo que se está lendo e pode ser utilizada quando algo está interferindo no foco de atenção (ALLINGTON; MCGILL-FRANZEN, 2011). Apenas 7 participantes se valeram dessa estratégia (26,92% dos participantes).

Ao avaliarem o texto, os leitores estão sendo ativos e questionando de alguma forma aquilo que leem. Conforme Duke e Pearson (2002, p. 1, tradução nossa), bons leitores “avaliam a qualidade e o valor do texto e reagem a ele de várias maneiras, tanto intelectual quanto

emocionalmente<sup>57</sup>”. Assim, 42,30% dos participantes avaliam o texto em algum momento da sua leitura.

Na elaboração, consideramos os movimentos inferenciais que os participantes realizaram, conectando informações de diferentes partes do texto e desenvolvendo uma reflexão a partir disso ou, então, trazendo para cena o conhecimento prévio construído e relacionando com as informações do texto. Metade dos participantes se vale dessa estratégia em algum momento da leitura.

Já a paráfrase foi a estratégia mais utilizada. Em algum momento, todos os participantes parafrasearam, resumiram ou explicaram alguma parte do texto. De acordo com a literatura (LESLIE; CALDWELL, 2011; JENNINGS; CALDWELL; LERNER, 2014), essa é uma estratégia que se mostra frequente entre leitores proficientes e é um dos processos envolvidos na compreensão (BLOCK; LACINA, 2011). Então, é uma estratégia bastante útil quando o objetivo está em compreender o texto lido. Cabe, agora, refletirmos sobre os resultados obtidos acerca das estratégias utilizadas e do desempenho no teste de compreensão em leitura.

Na sequência, apresentamos a estatística descritiva do Protocolo Verbal:

Tabela 36: Estatística Descritiva – Protocolo Verbal

Medida	Mon. Releitura	Con. Comp.	Con. Anaf.	Relato	Leit. Voz Alta	Av. Texto	Elab.	Paráf.
<b>Média</b>	1,69	2,77	1,54	1,12	0,54	1,27	2,35	9,12
<b>Mínimo</b>	0	0	0	0	0	0	0	0
<b>Máximo</b>	7	12	4	14	5	7	12	15
<b>Desvio Padrão</b>	0,38	0,65	0,24	0,61	0,23	0,38	0,63	0,86

Fonte: a autora.

Pela estatística descritiva já conseguimos visualizar o que apresentamos anteriormente: a estratégia mais utilizada pelos participantes foi a paráfrase. Então, explicações, resumo, repetição com outras palavras de partes consideradas importantes são ações que compreendemos como movimentos que geram uma reflexão a partir do texto. E são estas as estratégias mais recorrentes para estabelecer uma compreensão diante da leitura que lhes foi apresentada. A seguir, reportamos os resultados dos testes de normalidade para, na sequência,

<sup>57</sup> They evaluate the text's quality and value, and react to the text in a range of ways, both intellectually and emotionally.

refletirmos sobre os significados das estratégias utilizadas pelos participantes e suas relações com a anáfora estudada.

Tabela 37: Normalidade – Protocolo Verbal

	Shapiro-Wilk		
	Estatística	df	Sig.
Releitura	0,83	26	0,00
Monitoramento da Compreensão	0,80	26	0,00
Conexões Anafóricas Intersentenciais	0,87	26	0,00
Relato Episódico	0,42	26	0,00
Leitura em voz alta	0,53	26	0,00
Avaliação do Texto	0,70	26	0,00
Elaboração	0,75	26	0,00
Paráfrase	0,93	26	0,11

Fonte: a autora.

Os resultados obtidos por meio dos Testes de Shapiro-Wilk demonstram que os dados não seguem uma distribuição aproximadamente normal para as seguintes variáveis do Protocolo Verbal (PV): releitura, monitoramento da compreensão, conexões anafóricas intersentenciais, relato episódico, leitura em voz alta, avaliação do texto e elaboração. Os resultados obtidos com os Testes de Shapiro-Wilk demonstram, ainda, que os dados seguem uma distribuição aproximadamente normal para a variável paráfrase. A normalidade da paráfrase se dá exatamente por ter sido a estratégia mais usada pelos participantes, todos se valerem dela em algum momento da leitura.

A partir do que o teste de normalidade apresentou, utilizamos as correlações de postos de Spearman para correlacionar as estratégias utilizadas e o desempenho nas questões locais. Obtivemos o seguinte:



Tabela 38: Correlações de Postos de Spearman dos Protocolos Verbais e Compreensão Leitora Local

		<b>Questão 2</b>	<b>Questão 4</b>	<b>Questão 6</b>	<b>Questão 8</b>	<b>Questão 9</b>
<b>Releitura</b>	Cor.	-0,30	-0,26	0,00	-0,02	0,09
	Sig.	0,13	0,18	0,97	0,90	0,63
	N	26	26	26	26	26
<b>Monitoramento Compreensão</b>	Cor.	-0,12	-0,14	-0,14	0,00	-0,10
	Sig.	0,53	0,46	0,47	0,97	0,61
	N	26	26	26	26	26
<b>Conexões anafóricas intersentenciais</b>	Cor.	0,10	0,13	-0,01	-0,07	0,13
	Sig.	0,53	0,49	0,93	0,70	0,50
	N	26	26	26	26	26
<b>Relato Episódico</b>	Cor.	-0,19	-0,33	0,06	0,01	-0,32
	Sig.	0,34	0,09	0,75	0,93	0,11
	N	26	26	26	26	26
<b>Leitura em Voz Alta</b>	Cor.	-0,08	0,06	-0,16	0,18	0,31
	Sig.	0,69	0,73	0,43	0,37	0,11
	N	26	26	26	26	26
<b>Avaliação do texto</b>	Cor.	-0,14	-0,38	-0,14	-0,22	0,27
	Sig.	0,49	0,05	0,46	0,26	0,18
	N	26	26	26	26	26
<b>Elaboração</b>	Cor.	0,41	-0,17	0,10	0,02	0,19
	Sig.	0,03	0,4	0,59	0,89	0,34
	N	26	26	26	26	26
<b>Paráfrase</b>	Cor.	0,08	0,16	0,12	0,20	0,10
	Sig.	0,67	0,42	0,55	0,30	0,60
	N	26	26	26	26	26

Fonte: a autora.

Com relação à questão 2, foram encontradas correlações aproximadamente moderadas, positivas e estatisticamente significativas somente com a variável Elaboração (N = 26,  $p < 0,05$ ), obtidas por meio das correlações de Postos de Spearman, com um tamanho de efeito de 0,18 para explicar a covariância dos dados, ou seja, apenas 18% da covariância dos dados pode ser explicada por meio da associação entre essas variáveis (Elaboração X Questão 2). Dessa forma, temos que quanto mais a estratégia de elaboração é utilizada durante o protocolo verbal, mais os participantes acertam a questão 2. Esse resultado pode ser explicado pelo que solicitava a questão 2, como já apresentado, os participantes precisavam integrar diferentes partes do texto, estabelecendo a correferência entre anáfora e antecedente de modo adequado para poder responder ao solicitado. Já na estratégia de elaboração, foram consideradas as verbalizações

que revelavam sinalizações inferenciais, integrando diferentes partes do texto. Ainda que a questão 2 exigisse uma integração mais local, processos inferenciais podem ser estratégias úteis.

No que diz respeito às outras questões de integração local, não foram encontradas correlações estatisticamente significativas para as variáveis correlacionadas, ou seja, houve uma correlação significativa apenas com a estratégia de elaboração.

Correlacionamos, ainda, as estratégias obtidas com o resultado que os participantes obtiveram no teste de compreensão em leitura como um todo. Para fazer a correlação com a “paráfrase”, utilizamos coeficientes de correlação de postos de Pearson:

Tabela 39: Correlações de Pearson dos Protocolos Verbais e Compreensão Leitora Geral

		<b>Paráfrase</b>
<b>Total TCL</b>	Cor.	0,35
	Sig.	0,07
	N	26

Fonte: a autora.

Ainda que o resultado tenha sido próximo ao nível de significância, não foram encontradas correlações estatisticamente significativas com relação às variáveis Paráfrase e Compreensão Leitora Geral. O que os estudos têm mostrado (LESLIE; CALDWELL, 2011; JENNINGS; CALDWELL; LERNER, 2014; BLOCK; LACINA, 2011) é que a paráfrase, se bem empregada, é um recurso bastante produtivo para a compreensão. Esperava-se, contudo, uma correlação positiva e significativa apenas se essa paráfrase fosse bem construída e se o desempenho em leitura fosse alto. Como vimos na seção 4.1.4, a média obtida pelos participantes no teste de compreensão ficou em 6,66, quando a pontuação máxima era 10. Então, temos um índice de compreensão próximo de 70%, o que pode ser considerado mediano. Além disso, há que se refletir sobre o fato de termos trabalhado com a leitura de texto único em situação única, não sendo possível fazer a comparação de diferentes textos e trazer explicações robustas sobre a falta de significância na correlação entre paráfrase e desempenho geral no teste.

Para correlacionar as demais estratégias à compreensão leitora geral, foram utilizados os testes de correlações de postos de Spearman:

Tabela 40: Correlações de Postos de Spearman dos Protocolos Verbais e Compreensão Leitora Geral

		<b>Total TCL</b>
<b>Releitura</b>	Cor.	-0,33
	Sig.	0,09
	N	26
<b>Monitoramento Compreensão</b>	Cor.	-0,27
	Sig.	0,18
	N	26
<b>Conexões anafóricas intersentenciais</b>	Cor.	0,40
	Sig.	0,04
	N	26
<b>Relato Episódico</b>	Cor.	-0,39
	Sig.	0,04
	N	26
<b>Leitura em Voz Alta</b>	Cor.	0,03
	Sig.	0,86
	N	26
<b>Avaliação do texto</b>	Cor.	-0,39
	Sig.	0,04
	N	26
<b>Elaboração</b>	Cor.	0,15
	Sig.	0,43
	N	26

Fonte: a autora.

Os Testes de Correlações de Postos de Spearman demonstraram haver uma correlação aproximadamente moderada, positiva e estatisticamente significativa para a associação entre as variáveis Conexões Anafóricas Intersenciais e Total TCL ( $N = 26$ ,  $p = 0,04$ ), com tamanho de efeito de 0,22. O fato de existir uma correlação positiva, moderada e estatisticamente significativa sugere que, quanto mais os leitores empregaram estratégias que fizessem as conexões entre as sentenças, maior foi o desempenho no teste de compreensão. Ainda que haja muitos fatores intervenientes, é mais um indício de que: 1. As anáforas, ainda que atuem em um nível de processamento local, têm participação importante nos níveis superiores associados à compreensão; e 2. As anáforas realizadas com o sintagma “o mesmo” não parecem ter afetado negativamente a compreensão geral do texto. Assim, se geraram algum tipo de impasse, este foi possível de ser resolvido com estratégias específicas e não se mostrou empecilho para alcançar escores mais elevados no teste de compreensão.

Além disso, foi obtida correlação negativa e aproximadamente moderada para Relato Episódico e Total TCL ( $N = 26$ ,  $p = 0,04$ ), com tamanho de efeito de 0,16; e negativa e

aproximadamente moderada para Avaliação do Texto e Total TCL ( $N = 26$ ,  $p = 0,04$ ), com tamanho de efeito de 0,20. Nesses casos, o relato episódico e a avaliação do texto não se mostraram estratégias interessantes para aumentar os escores do teste de compreensão, pois, quanto mais tais estratégias foram usadas, menores foram os resultados obtidos no teste, esse resultado pode ser explicado, de modo geral, por deslocamento de atenção. De modo mais específico, no que tange aos relatos episódicos, esse resultado é esperado, já que eles podem ser considerados como fuga do texto, pois os participantes se valem apenas de lembranças relacionadas a disciplinas ou experiências práticas, sem, necessariamente, relacionarem tais lembranças de forma direta com o que o texto apresenta. As avaliações do texto feitas pelos participantes dizem respeito a perspectivas individuais, baseadas em opiniões, as quais, nem sempre, estão na direção do que o texto apresenta ou, então, podem representar uma expectativa de compreensão sem que ela de fato tenha ocorrido, por exemplo, quando os participantes se expressam com: [...] *é interessante abordar os três termos nesse texto* (P11) ou *sim, são muito importantes realmente* (P7). Isto é, são avaliações que não necessariamente demonstram compreensão do todo do texto ou revelam algum tipo de reflexão ou inferência acerca do que está sendo lido.

Além de considerar as estratégias utilizadas durante a leitura de todo o texto, correlacionamos as estratégias usadas diante das anáforas construídas com o pronome “ele” e com o sintagma “o mesmo”. Nesse caso, temos duas condições e um mesmo participante; por conta disso, utilizamos o teste de Wilcoxon, método não-paramétrico, que compara duas amostras pareadas. A seguir, apresentamos a estatística descritiva do teste:

Tabela 41: Estatística Descritiva – Protocolo Verbal – Comparação “ele” e “o mesmo”

Medida	Média	Mínimo	Máximo	Desvio Padrão
<b>Releitura O mesmo</b>	0,42	0	3	0,85
<b>Mon. Comp. O mesmo</b>	0,46	0	3	0,76
<b>Con. Anaf. O mesmo</b>	0,85	0	2	0,78
<b>Relato O mesmo</b>	0,19	0	3	0,69
<b>Leit. Voz Alta O mesmo</b>	0,08	0	1	0,27
<b>Av. Texto O mesmo</b>	0,23	0	2	0,58
<b>Elaboração</b>	0,19	0	2	0,56

<b>O mesmo</b>				
<b>Paráfrase</b>	1,38	0	2	0,69
<b>O mesmo</b>				
<b>Releitura</b>	0,08	0	1	0,27
<b>Ele</b>				
<b>Mon. Comp.</b>	0,31	0	3	0,73
<b>Ele</b>				
<b>Con. Anaf.</b>	0,54	0	2	0,58
<b>Ele</b>				
<b>Relato</b>	0,19	0	2	0,56
<b>Ele</b>				
<b>Leit. Voz Alta</b>	0,08	0	2	0,39
<b>Ele</b>				
<b>Av. Texto</b>	0,15	0	1	0,36
<b>Ele</b>				
<b>Elaboração</b>	0,62	0	3	1,02
<b>Ele</b>				
<b>Paráfrase</b>	1,58	0	4	1,13
<b>Ele</b>				

Fonte: a autora.

A estatística descritiva mostra que a média obtida para todas as estratégias é baixa. Isso ocorre porque a variabilidade dos dados, diferença entre o mínimo e o máximo, é também baixa. Nesse sentido, os participantes se valeram de poucas estratégias diante das anáforas que são nosso alvo. Esse resultado pode demonstrar que tais anáforas foram processadas automaticamente e não exigiram a mobilização de estratégias para resolver possíveis impasses na leitura, o que sugere que os participantes conseguiram estabelecer a coesão a partir da relação entre as sentenças, integrando a informação prévia com as novas sentenças. Então, ao que parece, os leitores conseguiram usar a informação referencial para buscar em sua memória o antecedente, formando uma representação local coerente (JUST; CARPENTER, 2013), sem que fosse necessário o uso de estratégias específicas nesse momento. Isso se dá para as duas formas anafóricas (o pronome “ele” e o sintagma “o mesmo”).

Para o teste de Wilcoxon, foram comparadas as médias das estratégias de “ele” com as médias das estratégias de “o mesmo”:

Tabela 42: Teste de Wilcoxon - Protocolo Verbal – Comparação “ele” e “o mesmo”

Medida	Z	Sig
<b>Releitura Ele – O mesmo</b>	-2,12	0,03
<b>Mon. Comp. Ele – O mesmo</b>	-0,71	0,47
<b>Con. Anaf. Ele - O mesmo</b>	-1,99	0,04
<b>Relato Ele - O mesmo</b>	0,00	1
<b>Leit. Voz Alta Ele - O mesmo</b>	0,00	1
<b>Av. Texto Ele - O mesmo</b>	-0,64	0,51
<b>Elaboração Ele - O mesmo</b>	-2,41	0,01
<b>Paráfrase Ele - O mesmo</b>	-0,73	0,46

Fonte: a autora.

Os resultados obtidos com os Testes T de Wilcoxon apontam para diferenças nas medianas<sup>58</sup> entre as seguintes variáveis: Releitura Ele – Releitura O Mesmo (N = 26, Z = -2,12, p = 0,03), com tamanho de efeito de 12,3; Conexões Anafóricas Intersentenciais Ele – Conexões Anafóricas Intersentenciais O Mesmo (N = 26, Z = -1,99, p = 0,04), com tamanho de efeito de 12,1; e Elaboração Ele – Elaboração O Mesmo (N = 26, Z = -2,41, p = 0,01), com tamanho de efeito de 12,3. O que esses resultados mostram é que: 1. houve mais releitura de “o mesmo” do que releitura de “ele”; 2. Foram feitas menos conexões anafóricas intersentenciais para “ele” do que para “o mesmo”; 3. Houve menos elaborações para “ele” do que para “o mesmo”. Então, ao comparar as duas formas anafóricas, os resultados apontam para a necessidade de mais estratégias diante do sintagma “o mesmo”. Há que se considerar, contudo, que essa diferença pode ser por conta de a variabilidade dos dados não ser grande, já que o tamanho de efeito obtido mostra que os resultados podem ter sido ao acaso e terem sofrido efeito de variáveis intervenientes.

Nas demais estratégias, não há diferença estatisticamente significativa, ou seja, o tipo de anáfora (“ele” ou “o mesmo”) não parece afetar a estratégia utilizada. Então, a partir dos resultados, as estratégias utilizadas pelos participantes não se mostraram diferentes para cada uma das anáforas. A releitura como estratégia usada durante o protocolo verbal tende a demonstrar uma falta de compreensão por parte do leitor, ou seja, ele relê porque algo naquele trecho não ficou claro ou porque quer confirmar se a compreensão formulada foi adequada. Nesse caso, mais releituras em trechos que continham a anáfora “o mesmo” poderiam, ainda que não necessariamente, ser preditoras de um possível impasse. No entanto, não parece ter havido qualquer empecilho para o estabelecimento das integrações locais possibilitadas pelo uso das anáforas realizadas com o sintagma e com o pronome, como foi possível inferir nos

<sup>58</sup> Como se trata de um teste não-paramétrico de ranqueamento, trabalhamos com as medianas obtidas.

resultados obtidos nas questões que exigiam integração local para serem respondidas no teste de compreensão leitora (seção 4.1.4).

Os participantes também se valeram de mais estratégias de conexões anafóricas intersentenciais diante da anáfora construída com o sintagma “o mesmo”. Então, é possível que, para que a integração ocorresse, tenha sido necessário mobilizar uma estratégia específica, o que não ocorre diante do pronome.

Por sua vez, o fato de os participantes terem feito mais elaborações diante da anáfora realizada com “o mesmo” pode demonstrar mais reflexões e integrações ao conhecimento prévio. Temos indícios disso nos seguintes protocolos:

Quadro 7: Elaboração diante do sintagma “o mesmo”

<b>Participante</b>	<b>Texto</b>	<b>Verbalização</b>
P11	O serviço de detecção de defeitos, nesse sentido, pode não ser confiável em alguns ambientes. O mesmo pode acabar fornecendo informações incorretas ou apresentando altos tempos de respostas na detecção de defeitos de componentes do sistema.	É... por causa de um problema muitas vezes com ruído de... de leitura... é... muitas vezes a informação que a gente recebe... não é a informação real ou o tempo/ a demora para chegar a uma informação é muito grande e... o sistema, por ter uma resposta muito rápida, precisa ser atuado sobre ele rapidamente. Nessas horas, tanto o controle desse sistema como a detecção de falhas é... se torna um desafio grande.
P16	Por melhores que sejam as técnicas usadas na construção desses sistemas, diversas questões podem levá-los a apresentar defeitos: desgastes físicos de componentes, falhas humanas, fenômenos e/ou desastres naturais etc. Portanto, os mesmos devem possuir mecanismos adequados de tolerância a falhas.	áí... já não sei dizer o que é melhor, se uma fiscalização privada ou pública, enfim, mas que tem que ter uma fiscalização rigorosa, isso tem. Ahm... e também uma série de... inúmeros, vários testes antes de... de ser... efetivamente aplicado o sistema.
P18	O comportamento dos componentes em caso de falhas guia a forma de implementação dos detectores de defeitos. O mesmo recebe a denominação de semântica de falhas, que determina a técnica de verificação (ou monitoramento) a ser utilizada pelo serviço de detecção de defeitos.	por exemplo, na minha cabeça, uma coisa que tenha a propensão de pegar fogo vai receber um detector de incêndio... por perto... detector de fumaça...

Fonte: a autora.

Então, os participantes parecem processar a anáfora e, para além disso, fazem elaborações a partir do que leram. De todo modo, apenas três participantes se valeram da estratégia de elaboração diante da anáfora realizada com “o mesmo”.

Além da elaboração, uma estratégia bastante recorrente diante do sintagma “o mesmo” foi a paráfrase, seguindo a lógica encontrada no texto como um todo. Para refletir sobre isso, apresentamos a seguir algumas das paráfrases construídas:

Quadro 8: Paráfrase diante do sintagma “o mesmo”

Trecho	Paráfrase
Por melhores que sejam as técnicas usadas na construção desses sistemas, diversas questões podem levá-los a apresentar defeitos: desgastes físicos de componentes, falhas humanas, fenômenos e/ou desastres naturais etc. Portanto, os mesmos devem possuir mecanismos adequados de tolerância a falhas.	É... Independente das técnicas de construção, nenhum... nenhum sistema é def/ é à prova de falha... Então, a gente sempre tem que ter... <b>desenvolver mecanismos para... para evitar ou pelo menos para... diminuir as perdas se alguma falha ocorrer. (P2)</b>
Por melhores que sejam as técnicas usadas na construção desses sistemas, diversas questões podem levá-los a apresentar defeitos: desgastes físicos de componentes, falhas humanas, fenômenos e/ou desastres naturais etc. Portanto, os mesmos devem possuir mecanismos adequados de tolerância a falhas.	Aí a justificativa de mesmo colocando o sistema certo <b>ele vai precisar de algum jeito conter falhas fora do sistema</b> , desgaste físico [01:27 – 01:29 relê o trecho final] <b>Ou seja, tolerâncias às falhas. (P3)</b>
O serviço de detecção de defeitos, nesse sentido, pode não ser confiável em alguns ambientes. O mesmo pode acabar fornecendo informações incorretas ou apresentando altos tempos de respostas na detecção de defeitos de componentes do sistema.	O serviço de detecção pode não ser confiável... <b>(P9)</b>

Fonte: a autora.

Na verbalização construída por P2, há uma paráfrase que demonstra integração entre as diferentes sentenças e, dentro desse contexto, o estabelecimento adequado do antecedente (sistemas) à anáfora (os mesmos). É possível perceber, então, que o participante começa a formular proposições, as quais, mais adiante na leitura, devem ser integradas entre si.

No mesmo trecho de P2, P3 também realiza uma paráfrase. Vale ressaltar que esse trecho está no início do texto. Então, foi a primeira vez que o sintagma “o mesmo” apareceu aos participantes e não há indícios de rupturas ou de problemas na integração. Já P9 parafraseia, de modo resumido, um dos trechos do penúltimo parágrafo do texto. Esses são exemplos das paráfrases construídas por nossos participantes nos trechos que continham a nossa anáfora-alvo. Embora não reportemos todas as verbalizações, vale ressaltar que, para ser considerada paráfrase, consideramos apenas aquelas que foram construídas de modo coerente e adequado



ao texto, ou seja, não foram tentativas falhas de parafrasear. De modo geral, contudo, como vimos, a quantidade de estratégias utilizadas diante das anáforas não foi em grande número, o que pode ser explicado pelo fato de elas serem processadas automaticamente e não terem exigido qualquer tipo de desautomatização por parte do participante.

Tendo apresentado e refletido sobre os resultados obtidos para responder a cada uma das perguntas de pesquisa, na sequência, fazemos as análises gerais, entrecruzando os resultados obtidos e buscando responder ao grande problema desta investigação.

#### 4.2 DISCUSSÃO DOS RESULTADOS

A grande pergunta que guiou a nossa investigação foi: de que forma e em que medida o sintagma “o mesmo”, como anáfora correferencial, atua na integração local, na formação da base textual e na compreensão em leitura de texto acadêmico-científico? Diante do que apresentamos na seção anterior, podemos levantar algumas respostas, que partem da intersecção dos resultados obtidos por meio dos diferentes instrumentos que usamos, ainda que – sabemos – existam nuances e especificidades envolvidas em cada um desses diferentes níveis de processamento.

Para que a compreensão de um texto ocorra, como já esboçado na fundamentação deste trabalho, é necessário que o leitor desenvolva um modelo mental – uma representação do texto lido. Nesse sentido, o leitor não se lembrará da estrutura exata de um texto ou até das palavras usadas, mas lembrará da significação contida nele. Ocorre que, para se chegar a essa significação, ou seja, a um modelo mental coerente, é necessário considerar os diversos processos que acontecem a partir do texto (OAKHILL; CAIN; ELBRO, 2017; PERFETTI, LANDI; OAKHILL, 2013). Nesta pesquisa, detivemo-nos a um dos elementos responsáveis pela vinculação de frases. De modo mais específico, olhamos para anáforas correferenciais construídas com o sintagma “o mesmo”.

A fim de iniciar a resposta à nossa grande questão, começamos refletindo acerca da integração local, que foi observada a partir dos resultados obtidos por meio do teste de leitura automonitorada, das questões locais do teste de compreensão em leitura, do comportamento diante da anáfora realizada com “o mesmo” e das estratégias utilizadas diante das anáforas que foram foco desta pesquisa.

Para entender o funcionamento desse “o mesmo” como anafórico na língua portuguesa, precisávamos, em primeiro lugar, entender se o seu processamento ocorria de um modo que não atrapalhasse a vinculação entre as frases e, também, que garantisse o estabelecimento

adequado da correferência. Para tanto, comparamos o tempo de processamento desse sintagma com o processamento do pronome “ele”, anafórico que tende a ser utilizado quando não há uma mudança de tópico (GARNHAM, 2001). Como apresentado na seção anterior, o teste de leitura automonitorada mostrou não haver diferenças significativas no tempo de leitura do sintagma “o mesmo” e do pronome “ele”. Isso significa que o processamento de ambos os anafóricos ocorre de modo semelhante. Além disso, a acurácia foi exatamente a mesma, independentemente do anafórico utilizado, então a relação antecedente-anáfora ocorreu tanto em frases construídas com o pronome quanto em frases construídas com o sintagma. Os estudos de processamento anafórico (LEITÃO, 2005; QUEIROZ, 2009; GADELHA, 2012; NÓBREGA-LIMA, 2014; SIMÕES, 2014; GONDIM, 2017) que vêm sendo realizados mostram o pronome como um elemento linguístico de processamento mais rápido se comparado à repetição do antecedente que está sendo retomado, isso porque o pronome possui menos informações semânticas, o que evita uma sobrecarga da memória de trabalho. O sintagma “o mesmo”, por sua vez, parece seguir essa mesma lógica. Ainda que seja um sintagma e não tenha em sua essência a função de anáfora correferencial, possui apenas as informações semânticas necessárias para que se estabeleça a correferência sem sobrecargas no processamento. Nesse sentido, a partir da leitura automonitorada, podemos dizer que a integração local aconteceu e que ambos os anafóricos parecem ter comportamento semelhante na língua portuguesa.

Feitas essas elocubrações iniciais, para além dos resultados estatísticos, mobilizamos algumas reflexões com os resultados obtidos sobre o fato de a média de tempo de leitura do sintagma “o mesmo” ter sido menor que a do pronome “ele” nas duas medidas de tempos com as quais trabalhamos (tempo total e tempo do segmento pós-anafórico). A média de tempo do sintagma, na leitura do segmento pós-anafórico, foi de 460,33 milissegundos, enquanto a do pronome foi de 486,37 milissegundos. Já na leitura do tempo total, obtivemos uma média de 5152,10 milissegundos para o sintagma “o mesmo” e 5417,23 para o pronome. Além disso, ao verificarmos a média<sup>59</sup> de tempo de leitura de cada participante, também obtivemos uma média maior de tempo de leitura do pronome. No que diz respeito ao tempo total, obtivemos um resultado de 5138,12 milissegundos para “o mesmo” e 5378,16 milissegundos para ele. Já no que tange ao tempo de leitura do segmento pós-anafórico, obtivemos 460,38 milissegundos para “o mesmo” e 489,43 milissegundos para ele. A fim de tornar a visualização mais fácil, montamos a tabela a seguir com todos os tempos obtidos a partir da leitura automonitorada:

---

<sup>59</sup> Para responder à pergunta 2, utilizamos a média dos tempos de leitura de cada participante e, a partir disso, calculamos a média geral obtida, ou seja, a média da média.

Tabela 43: Comparação de médias do tempo em milissegundos – Leitura Automonitorada

	<b>“O mesmo”</b>	<b>Ele</b>
<b>Tempo total de leitura – média total</b>	5152,10	5417,23
<b>Tempo de leitura do segmento pós-anafórico – média total</b>	460,33	486,37
<b>Tempo total de leitura – média por participante</b>	5138,12	5378,16
<b>Tempo de leitura do segmento pós-anafórico – média por participante</b>	460,38	489,43

Fonte: a autora.

Assim, é possível que haja uma tendência para a leitura mais rápida do sintagma. No caso desta pesquisa, esse resultado pode ter sido devido ao N amostral e ao fato de terem sido aplicados testes não paramétricos, que são mais conservadores, mas vale mobilizar tal reflexão, já que se mostrou verdadeira nas medidas de tempo analisadas. O sintagma “o mesmo” tem sido usado como anafórico em contextos escritos mais formais, não aparecendo com tanta frequência em contextos orais (MOREIRA, 2007; PEREIRA, 2013). Desse modo, uma das possibilidades a serem levantadas para um processamento mais rápido do sintagma está, exatamente, no contexto com o qual trabalhamos: estudantes universitários lendo frases construídas com uma estrutura próxima daquelas que formam um texto acadêmico-científico. Diante disso, nos experimentos realizados, o anafórico retoma antecedentes que não possuem características de [+animados]. Então, o sintagma pode ter assumido, nesse contexto, um papel de neutralidade, que seria mais adequado para a retomada de antecedentes [-animados]. Em se tratando, portanto, de um processamento automático, é possível que o sintagma “o mesmo” esteja assumindo o papel de uma forma mais eficiente para a retomada de um antecedente [-animado] em contextos mais formais de escrita. De todo modo, como não encontramos diferenças significativas, seria interessante, testar esse tempo com um N amostral maior para saber se o processamento mais rápido do sintagma se mantém e se o fato de não haver uma diferença significativa entre esses anafóricos se confirma. De toda forma, diante dos resultados obtidos no teste de leitura automonitorada, em um processamento automático, o sintagma não parece gerar custos mais elevados ou problemas em retomar o antecedente ao qual se vincula.

Ainda no que diz respeito à atuação local do sintagma, trabalhamos, no teste de compreensão leitora, com questões que tinham por objetivo exatamente o processamento local. A média obtida nessas questões foi de 3,7 pontos no índice de acertos, de um total de 5 pontos, ou seja, houve um índice médio de acerto de 74%. A resolução adequada de tais questões dependia do processamento correto das anáforas construídas com o sintagma, com o pronome e, também, do estabelecimento de outros elementos de coesão local. Assim, o fato de a maioria dos participantes ter ficado com um desempenho mediano pode não ter sido apenas por questões

relativas às anáforas investigadas. Para refletir de modo mais específico acerca da resolução do sintagma, vale analisar duas questões mais detidamente, a saber: questão 2 e a questão 9 (Apêndice F).

No que diz respeito à questão 2, para além de realizar a integração possibilitada pelo sintagma “o mesmo”, o participante deveria compreender a explicação sobre “semântica de falhas”, ou seja, era necessário que ocorresse integração local e, feito isso, deveria ser construída uma explicação adequada para o termo e, além disso, os participantes deveriam indicar as linhas do texto para as quais recorreram a fim de desenvolver a resposta. Para entender a pontuação dos participantes nesta questão, trazemos algumas respostas:

*P1: Semântica de falhas é o comportamento dos componentes de um sistema em caso de falhas. Linhas 13 e 14.*

*P2: Semântica de falhas é o modo como os dispositivos se comportam quando estão apresentando falhas. 13-14*

*P3: Semântica de falhas constitui nos métodos utilizados para identificar falhas em componentes defeituosos. 13, 14, 15, 16, 17, 18 e 19.*

*P4: A semântica de falhas é como os componentes vão se comportar se caso ocorrer alguma falha no sistema em algum momento da operação. Linhas 13 até 16.*

*P5: Semântica de falhas é o conjunto de comportamentos de um componente em caso de falha. A partir de conhecimento prévio do comportamento do componente sabemos qual método de detecção de falha, por exemplo variável a ser medida ou timeouts para uma resposta, deve ser implementado. linhas 13 a 19.*

*P11: É o comportamento dos componentes do sistema quando ocorrem falhas. Linhas 13 e 14.*

*P15: Semântica de falhas corresponde ao conjunto de comportamento dos dispositivos quando estes encontram-se em modo de falha em sua operação, o mesmo contém todos os comportamentos que podem ocorrer com o dispositivo e irá guiar na adoção da melhor técnica para verificação de defeitos. Linhas 13 e 14.*

*P16: É a forma de implementação do sistema de detecção de falhas conforme o comportamento dos componentes e dispositivos quando falham. Linhas 13 e 14.*

*P18: Semântica de falhas é o comportamento dos componentes em caso de falhas guiando a instalação de detectores de defeitos. Como na língua portuguesa onde a semântica determina a estrutura de uma oração onde um componente requer outro (um sujeito requer um verbo por exemplo) a semântica das falhas orienta a estrutura dos possíveis defeitos e suas prevenções. Linhas 9 a 19.*

*P19: É o comportamento do componente em caso de falha. Ela guia o tipo de técnica utilizada na detecção dos defeitos. 13 a 16.*

P1, P2, P3, P4 e P5 obtiveram 0,5 pontos, as outras cinco respostas apresentadas alcançaram 1 ponto. De modo geral, apenas 7 participantes obtiveram pontuação máxima, 1 participante zerou a questão e os outros obtiveram 0,5 pontos. Ocorre, contudo, que a não obtenção da pontuação máxima parece não ter se relacionado diretamente com o estabelecimento e a compreensão do sintagma “o mesmo” como anafórico. Isso é possível verificar, uma vez que as respostas dadas pelos participantes mostram que eles conseguiram integrar o antecedente “o comportamento dos componentes diante de falhas” com o sintagma, entendendo, assim, que “semântica de falhas” é o nome dado a esse comportamento. Essa compreensão era possível somente com a integração adequada das duas sentenças, possibilitada pela resolução do anafórico. No entanto, tiveram dificuldades de explicar a semântica de falhas em si, ponto solicitado pela questão. Então, a integração parece ter ocorrido, embora a construção de uma explicação (que exigia inferências mais elaboradas) tenha ficado prejudicada. Essa dificuldade pode se relacionar com questões superiores de compreensão (KINTSCH; RAWSON, 2013; KINTSCH, 1998), mas pode, também, estar relacionada a problemas em expressar por escrito aquilo que foi solicitado.

Já a questão 9 solicitava que os participantes localizassem os outros dois sintagmas “o mesmo” que estavam no texto, explicitassem seu referente e explicassem a informação trazida nesses trechos. Nesse caso, 16 participantes obtiveram pontuação máxima, ou seja, 61,53% dos participantes do estudo conseguiram resolver adequadamente a questão. Os outros 10 demonstraram dificuldades com a explicação solicitada, mas não em integrar antecedente e anáfora construída com o sintagma. Então, mais uma vez a construção anafórica não parece ter sido um problema, isto é, ela cumpre o seu papel de integração local, possibilitando a realização de inferências locais e possibilitando a construção de uma microestrutura coerente.

Para ter mais indícios acerca da medida de integração local possibilitada pelo sintagma, analisamos, também, as estratégias empregadas diante das duas anáforas utilizadas. O que os resultados explicitados na seção 4.1.5 demonstram é que os participantes se valeram de poucas estratégias diante das anáforas. As estratégias, portanto, foram mais utilizadas como meio para auxiliar nas inferências elaborativas e na formação de uma representação mental do texto. Nessa reformulação, não se usa a estrutura exata do texto lido. Para que isso seja feito, são necessárias habilidades de elaboração e inferência, fazendo com que o leitor comece a construir os seus próprios sentidos a partir do lido e da intersecção com seus conhecimentos prévios.

No estudo que realizamos, a estratégia “paráfrase” foi a mais utilizada pelos participantes, inclusive diante da anáfora “o mesmo”. Essa estratégia demonstra uma tentativa

de formular uma representação mental acerca do texto lido, já que exige que o leitor sintetize, selecione e faça uma reformulação das ideias principais do texto, movimento necessário, de acordo com Kintsch (1998) e Kintsch e van Dijk (1978), para que proposições sejam formuladas. Esse movimento parafrásico, portanto, é possível porque a análise semântica já foi construída e as proposições já foram correlacionadas em um nível micro (KINTSCH; RAWSON, 2013).

Então, neste ponto, a partir da estratégia mais mobilizada durante as sessões de protocolos verbais, começamos a responder acerca da formação da base textual por parte dos participantes deste estudo, já que a paráfrase é uma estratégia que auxilia nessa formação “do significado do texto tal qual expresso verdadeiramente pelo texto” (KINTSCH; RAWSON, 2013, p.229). Dessa forma, à medida que os participantes foram integrando as diferentes partes do texto, foram procurando construir a sua representação acerca do significado do texto. No caso do protocolo verbal, foram fazendo essa construção por etapas já que a solicitação era que verbalizassem seus pensamentos durante a atividade de leitura.

O que essas utilizações de paráfrases podem, então, nos indicar é que, de acordo McNamara et al (2007, p. 401, tradução nossa), em primeiro lugar, o leitor está transformando

o material em uma representação mais familiar e, conseqüentemente, mais fácil de memorizar. Em segundo lugar, a capacidade de parafrasear representa o nível mais básico de compreensão porque, para parafrasear com sucesso, deve-se ser capaz de processar a estrutura básica e as relações da frase para transformar o texto literal em palavras mais familiares.<sup>60</sup>

Assim, os participantes da pesquisa, ao utilizarem da estratégia de paráfrase, estão demonstrando um movimento em direção à construção da base textual, ou seja, do significado do texto, o que serve de fortalecimento para a construção de um modelo situacional coerente (MCNAMARA et al, 2007). Com isso em mãos, conseguimos, aos poucos construir um caminho em direção à relação entre o processamento da anáfora, mais especificamente do sintagma “o mesmo”, e a formulação de um modelo situacional do texto lido.

Além da paráfrase, outra estratégia mais recorrente diante da anáfora “o mesmo” foi a conexão interssentencial, que acontece, como apresentado na seção 3.7, quando o participante relaciona alguma informação contida em uma sentença com a informação contida em outra

---

<sup>60</sup> Describing the text in one’s own words serves two functions. First, it allows the reader to transform the material into a representation that is more familiar and consequently more memorable. Second, the ability to paraphrase roughly translates into the most basic level of comprehension because, to paraphrase successfully, one must be able to process the basic structure and relations of the sentence to transform the verbatim text into more familiar words. (MCNAMARA et al, 2007, p.401).

sentença (em outra parte do texto) na qual existe uma conexão anafórica. Esse movimento também indica que o participante integrou antecedente e anafórico e conseguiu relacionar as informações contidas nas distintas partes lidas. Assim, o anafórico “o mesmo” cumpre com a sua função de, do mesmo modo que o pronome, referir “conceitos que foram mencionados ou abordados recentemente, que foram introduzidos explicitamente no texto e que estão ativados na memória de trabalho” (KINTSCH; RAWSON, 2013, p. 233). O sintagma, portanto, parece indicar de forma adequada qual o antecedente a que se refere, adquirindo seu significado com base em seus antecedentes (OAKHILL; CAIN; ELBRO, 2017).

Ainda no que tange à formação da base textual, podemos levantar o resultado de uma das questões do teste de compreensão leitora. A questão 10 solicitava que os participantes fizessem, por escrito, uma síntese do que haviam lido. Dos 26 participantes, 12 gabaritaram a questão, os demais obtiveram 0,5 pontos. A média de acertos ficou em 0,6. Isso significa uma média próxima da pontuação máxima (1,0). Nesse sentido, é possível aventar que, em sua maioria, eles conseguiram construir uma significação do texto, formulando uma base textual adequada. Esse é o primeiro passo para a construção de uma representação mental adequada do texto. Além disso, o fato de as estratégias utilizadas pelos leitores durante as sessões de protocolo verbal não terem revelado qualquer impasse no que diz respeito ao sintagma em si, sugere que os participantes não tiveram seu processamento interrompido ou prejudicado por essa construção.

Então, até este ponto, já temos indícios de que o leitor consegue integrar as frases em que o sintagma aparece, formando a base textual do texto. No entanto, sabemos, como discutido ao longo do capítulo de fundamentos teóricos, que compreensão acontece quando o leitor forma uma representação mental adequada do texto lido, transpondo, assim, a base textual e relacionando os conhecimentos provindos do texto com conhecimentos prévios relevantes. A última parte da nossa pergunta se direciona a esse aspecto “compreensão”.

Ocorre que a compreensão não é uma medida exata a qual simplesmente se chega. Compreensão é construção e se dá a partir dos diversos fatores já citados ao longo do nosso trabalho: texto, leitor, situação e contexto da leitura (KINTSCH, 1998; KINTSCH, 2013; KINTSCH; RAWSON, 2013; OAKHILL; CAIN; ELBRO, 2017; SOUZA; SEIMETZ-RODRIGUES; WEIRICH, 2019). Dessa forma, para que conseguíssemos avaliar a compreensão dos participantes desta pesquisa, a ferramenta utilizada foi o produto obtido a partir do teste de compreensão leitora. Com o resultado geral alcançado nele, obtivemos informações acerca da compreensão dos participantes, que, em sua maioria, ficou em um nível mediano e alto. De modo geral, temos que 25 participantes obtiveram escores acima de 50%.

Sabemos que a compreensão não é um processo unitário, mas processos de integração entre distintos componentes (KINTSCH, 1998; KINTSCH; KINTSCH, 2005; KINTSCH, 2013; KINTSCH; RAWSON, 2013; OAKHILL; CAIN; ELBRO, 2017), que envolvem os diferentes fatores os quais temos citado ao longo do nosso estudo, dentre eles estão as informações contidas no texto e o conhecimento prévio relevante do leitor. Para este estudo, não checamos o conhecimento prévio específico dos participantes no que diz respeito ao texto lido em si, apenas controlamos o fato de todos provirem de um mesmo curso e de já terem alguma experiência em sua área de estudo, não sendo calouros.

Os escores gerais do teste mostram, então, um nível mediano de compreensão. No entanto, esse resultado pode ter inúmeros elementos por trás, tendo em vista todos os pilares necessários para se construir uma compreensão coerente e relevante e, ainda, considerando que trabalhamos com apenas um texto. O que podemos fazer é considerar que: 1. O anafórico “o mesmo” parece ter cumprido sua função e não ter gerado entraves para a formulação da base textual. 2. O resultado final para a compreensão geral pode ter sido devido a questões que envolviam níveis mais elevados de compreensão, como as questões 1, 3 e 7, as quais requeriam análise e reflexão. Analisando mais detidamente os resultados dessas questões do teste, temos o seguinte:

Questão 1: apenas um participante atingiu pontuação máxima e um participante zerou a questão, ou seja, 24 participantes alcançaram meio ponto aqui. Isso porque não conseguiram elaborar explicações suficientemente claras acerca dos sistemas distribuídos, o que requeria um conhecimento prévio sobre o tema, pois a explicação não se apresentava no texto.

Questão 3: 15 participantes atingiram pontuação máxima nessa questão, 2 zeraram e 9 obtiveram 0,5 pontos. Para responder a essa questão, os participantes deveriam ter compreendido o texto, mas, para além disso, deveriam conseguir fazer relações com conhecimentos de sua área. A maioria conseguiu responder de forma suficientemente adequada.

Questão 7: 4 participantes gabaritaram, 3 zeraram e 19 obtiveram 0,5 pontos nessa questão. Aqui, também era necessário ir para além do texto a fim de que se construísse uma resposta adequada ao solicitado. O que se percebe, então, é que as dificuldades começam a aparecer em níveis superiores de processamento, que requerem vinculação a outros conhecimentos e a formulações mais elaboradas, sobretudo no que diz respeito ao componente inferencial (PERFETTI; LANDI; OAKHILL, 2013), que se apoia no texto, mas que o transcende. O que temos, então, é o fato de que o primeiro passo para a formação de uma base textual adequada e de um modelo mental coerente parece ter se mostrado efetivo com o uso do sintagma “o mesmo” como anafórico.



Portanto, considerando todos os aspectos avaliados, temos indícios que levam à sustentação da hipótese levantada para a presente investigação. Isto é, “o mesmo” segue um padrão de processamento semelhante ao do pronome “ele”. Sendo assim, os resultados em termos de processamento e de produto indicam que “o mesmo” pode exercer a função de anáfora correferencial sem que isso gere entraves para a compreensão em leitura. Então, ainda que esse sintagma não se caracterize como pronome, ele pode assumir tal função e atuar como anáfora correferencial, sem que isso gere complicações ou atrasos no processamento e na compreensão.

#### 4.2.1 *Outliers* e o N amostral: o que revelam sobre os resultados obtidos?

Antes de encerrarmos este texto, é relevante tratarmos dos casos em que os participantes, em algumas variáveis de alguns testes, se distanciaram dos demais, tornando-se *outliers*. Como dissemos no início das análises, optamos por não os retirar da rodada estatística, isso porque entendemos que eles também representam a amostra e fazem parte do grupo com o qual trabalhamos, ainda que possam apresentar particularidades. Além disso, ao retirar um *outlier*, outros surgem, então o N amostral diminuiria demasiadamente se optássemos por removê-los.

Os testes que tiveram participantes *outliers* foram o de leitura automonitorada e o protocolo verbal. No que diz respeito aos protocolos, os *outliers* aparecem em estratégias outras que não as de paráfrase. Isso porque todo e qualquer participante que usou qualquer uma das outras estratégias diante dos anafóricos acaba se caracterizando como outlier, já que elas não foram muito recorrentes e o N amostral é pequeno. Por conta disso, optamos por não analisar todos os *outliers*, mas focar em um, P22, que se caracteriza como *outlier* tanto na leitura automonitorada quanto nas seguintes estratégias utilizadas nos protocolos verbais: avaliação do texto diante da anáfora “o mesmo”, releitura e monitoramento da compreensão diante da anáfora “ele”.

No que diz respeito ao experimento de leitura automonitorada, P22 se distancia dos demais na leitura de quatro frases, quais sejam:

*A prova solicita os dados do estudo. A mesma exige exemplos práticos e argumentos embasados.*

*O livro foca em concepções econômicas. O mesmo debate criticamente os efeitos do capitalismo.*

*A aula expõe os programas de incentivo. Ela aborda especialmente a formação inicial.*

*O tema permite inúmeras abordagens. Ele provoca reações diversas no campo de estudo.*

O tempo de leitura desse participante se eleva bastante diante desses estímulos. Não se pode dizer que o tipo de anáfora seja responsável pelo tempo maior de leitura nesse caso, pois não houve um padrão, apenas frases com “ele” ou apenas frases com “o mesmo”. Além disso, ao analisarmos o julgamento de aceitabilidade que P22 fez desses conjuntos de sentenças, temos, respectivamente, as seguintes pontuações: 7, 5, 7 e 6, ou seja, em uma avaliação deliberada, ele não considera tais conjuntos problemáticos. O que pode ter ocorrido, durante a leitura automonitorada, foi uma queda atencional no momento da leitura dessas sentenças, fazendo com que o tempo de leitura ampliasse.

No que diz respeito ao protocolo verbal, temos uma avaliação do texto diante da anáfora “o mesmo”:

Quadro 9: Avaliação do texto diante da anáfora "o mesmo"

Trecho	Avaliação do texto
O serviço de detecção de defeitos, nesse sentido, pode não ser confiável em alguns ambientes. O mesmo pode acabar fornecendo informações incorretas ou apresentando altos tempos de respostas na detecção de defeitos de componentes do sistema.	É, de fato, dependendo de onde tu está trabalhando, a detecção de defeito pode ser complicada em locais com temperatura muito alta, pode interferir na... no sistema...

Fonte: a autora.

Este participante foi, portanto, um dos poucos a utilizar a estratégia de avaliação de texto diante dessa anáfora. A avaliação feita, contudo, não mostra qualquer tipo de entrave ou de problema diante da construção anafórica; muito pelo contrário, o participante parece processar o sintagma e relacioná-lo de modo adequado ao seu antecedente, fazendo uma avaliação que vai ao encontro do que o texto afirma sobre o serviço de detecção de defeitos em alguns ambientes.

Além disso, ele se caracteriza como *outlier* também no uso de duas estratégias diante da anáfora realizada com o pronome:

Quadro 10: Monitoramento da Compreensão e Releitura diante da anáfora "ele"

Trecho	Estratégias
O desempenho de um detector de defeitos é definido em termos de sua acurácia e rapidez. Ele é adequado se os seus parâmetros operacionais forem configurados de forma a atender as demandas das aplicações, mediante as condições de carga e disponibilidade dos recursos computacionais do ambiente distribuído.	Certo, detector de defeitos, ele tem que... é... acurácia, seria precisão, né, é... e velocidade, tá! ele tem que atender as aplicações e tem que suportar condições de carga variável... disponibilidade de recursos (monitoramento da compreensão)
[...] por isso, eles podem degradar o desempenho da aplicação ou do próprio serviço de detecção, fornecendo informações incorretas ou tempos de detecção inadequados, se os intervalos entre verificações dos defeitos não forem definidos de forma apropriada.	é... [10:40 começa a reler] o tempo de detecção for inadequado [começa em voz alta e depois silencia até 10:54] uhum (releitura)

Fonte: a autora.

Dessas duas estratégias utilizadas diante do pronome, a releitura poderia indicar algum tipo de entrave no trecho lido, ou seja, algo não ficou claro e precisou ser feita a releitura. Já o monitoramento da compreensão indica que o participante está buscando confirmar a sua compreensão acerca do que leu, mostrando um movimento importante em direção à compreensão. Como destacamos anteriormente, tal estratégia tende a ser um preditor da compreensão (BAKER; BEALL, 2011), porque revela um leitor que está atento àquilo que compreende ou não. No caso desta pesquisa, contudo, não foi uma estratégia recorrente, embora tenha se repetido nas verbalizações deste participante. Por conta disso, vale refletir sobre o resultado de P22 no teste de compreensão em leitura. Ele é o único que fica abaixo da média no teste, obtendo um escore de 4,4 pontos, ou seja, é possível que haja alguma dificuldade inerente ao próprio participante. Contudo, ao que parece, mesmo esse participante que se mostra mais frequente como *outlier* não apresenta indícios que poderiam levar a diferentes resultados no que diz respeito ao processamento da anáfora “o mesmo” e da sua influência da formação da base-textual e da compreensão em leitura.

## 5 CONSIDERAÇÕES FINAIS

No decorrer de todo este texto, procuramos tecer um diálogo acerca da leitura e dos níveis de processamento que atuam para a concretização dela. Para tanto, focalizamos nossa atenção no processamento anafórico e na compreensão em leitura, olhando, mais especificamente, para anáforas construídas com o sintagma “o mesmo”, o qual tem assumido função de anáfora correferencial nos mesmos termos que o pronome “ele”. Para fechar essa discussão, organizamos estas conclusões da seguinte maneira: i. retomada do grande objetivo estabelecido e dos instrumentos usados para atingi-lo; ii. reflexão sobre os achados da pesquisa; iii. interlocução de diferentes áreas para que a pesquisa ocorresse; iv. as limitações do estudo; v. as direções a serem tomadas em pesquisas futuras e, por fim, vi. as contribuições da pesquisa realizada.

O objetivo da presente investigação estava em compreender de que forma e em que medida “o mesmo” usado como anáfora correferencial influencia no processamento e na compreensão em leitura de texto acadêmico-científico. Para que fosse possível atingir tal objetivo, trabalhamos com acadêmicos de um curso específico, qual seja: Engenharia de Controle e Automação, e nos valem de diferentes instrumentos, os quais nos auxiliaram na análise da compreensão em leitura, levando em conta seu produto e, também, diferentes níveis de processamento, dentre os quais aquele em que ocorre o processamento anafórico.

A leitura automonitorada teve relevante papel no que diz respeito ao estabelecimento da relação entre anáfora e antecedente, além de, por meio do tempo de leitura, trazer indícios importantes acerca do processamento em nível local da anáfora realizada com “o mesmo”, comparando-a com a anáfora realizada com “ele”. Os protocolos verbais auxiliaram na análise de níveis superiores de compreensão, trazendo indícios de informações disponíveis na memória de trabalho do leitor e possibilitando a descrição dos processos cognitivos envolvidos na leitura. Diante disso, foi possível ter acesso aos movimentos feitos pelos participantes na construção da compreensão do texto. Além disso, o teste de compreensão leitora também nos auxiliou no que diz respeito aos movimentos em direção à compreensão a partir da gravação das ações dos participantes na tela do computador quando do momento de responder às questões. Ainda, trouxe resultados acerca da compreensão como produto, a partir do olhar lançado a diferentes aspectos cognitivos (localizar, avaliar, analisar, inferir), que evidenciam os distintos níveis envolvidos no processo da compreensão. Já o teste de aceitabilidade mostrou a avaliação metacognitiva feita pelos participantes acerca de conjuntos de frases construídas com o pronome e com o sintagma.

Cada uma desses instrumentos, portanto, nos auxiliou na resposta a questões específicas que conduziram esta investigação, mas foi a partir da sua interrelação que pudemos atingir o grande objetivo que guiou o estudo. Obtivemos alguns resultados bastante interessantes e, em sua maioria, que foram ao encontro das hipóteses formuladas. Em primeiro lugar, o processamento do sintagma se mostrou semelhante ao do pronome nos contextos investigados – posição de sujeito, retomando antecedentes [-animados]. Nesse sentido, ambas as retomadas mostraram cumprir sua função de anáfora, retomando o antecedente de modo acurado e rápido. Em contrapartida, o julgamento de aceitabilidade evidenciou que os participantes do estudo avaliaram como mais aceitáveis as frases construídas com o pronome, ainda que não tenham considerado inaceitáveis as frases com o sintagma. O que temos, então, é que, em um nível cognitivo, o sintagma é processado sem maiores custos se comparado ao pronome; e, num nível metacognitivo, o pronome é a forma preferencial. Isso, como vimos, pode ser devido a alguns fatores, como a própria frequência de uso do pronome, que tende a ser o recurso anafórico mais usado quando não há mudança de tópico. Já “o mesmo” tem um uso mais restrito, menos frequente, como anáfora correferencial (AMORIM, 2009; PEREIRA, 2013).

No que diz respeito ao teste de compreensão leitora, utilizamos de dados provenientes de duas formas: produto e processo. O produto se trata do resultado do teste em si, enquanto o processo diz respeito ao comportamento dos participantes diante do texto no momento de responder às questões. Os resultados obtidos mostraram não ter havido qualquer impasse diante do estabelecimento do antecedente das anáforas construídas com “o mesmo”, nem diante da construção de uma base textual coerente do texto lido. Já o processo, ou seja, o comportamento dos participantes diante do texto, mostrou que eles retomam o texto com bastante frequência para responder às questões. No que diz respeito, mais especificamente, às questões relativas às anáforas, era esperada tal ação, já que se tratava de questões de localização. Além disso, ao correlacionarmos o comportamento diante da anáfora “o mesmo” com o comportamento diante do texto como um todo, obtivemos padrões semelhantes, o que indica que não há comportamentos específicos diante do sintagma, mas comportamentos generalizados, mostrando que “o mesmo” atua como anáfora correferencial sem que isso demande de mais ou diferentes estratégias por parte dos participantes. Ainda, no que tange à gravação de tela, vale destacar, conforme Schlichting (2021, p.223), que “[t]er acesso aos movimentos realizados durante a tarefa de teste de compreensão em leitura possibilita acompanhar o processo – e defrontá-lo com o produto – da compreensão”, o que foi essencial para poder entender como os participantes agiam diante das anáforas construídas com o sintagma.

Resultados semelhantes foram obtidos a partir das verbalizações feitas pelos participantes nas sessões de protocolos verbais. Isto é, as estratégias utilizadas diante do sintagma seguiram o mesmo padrão das estratégias utilizadas no texto como um todo, sendo a paráfrase a estratégia mais utilizada. A utilização dessa estratégia mostra que os participantes estão buscando construir a compreensão a partir do processamento da estrutura básica do texto (MCNAMARA et al, 2007), não demonstrando entraves ou problemas diante dessa construção anafórica. Dessa forma, a partir da interlocução feita entre os resultados obtidos por meio dos diferentes testes realizados, podemos dizer que o sintagma “o mesmo” parece atuar de modo adequado como um anafórico correferencial assim como o pronome, não prejudicando a formação de uma base textual coerente pelo leitor e nem interferindo de modo negativo nos processos superiores de compreensão. Dessa forma, ainda que haja insurgências (BAGNO, 2012; SACCONI, 1994; ALMEIDA, 2009) contra o uso do sintagma “o mesmo” como anáfora correferencial, ele atua nessa função de modo proficiente e cumpre com o esperado de uma anáfora correferencial na tecitura textual (KOCH; ELIAS, 2018; KOCH; ELIAS, 2015).

Os resultados mostram, ainda, que os estímulos foram bem construídos. Na leitura automonitorada, por exemplo, a acurácia é exatamente a mesma para as frases realizadas com “o mesmo” e para as frases que contêm o pronome “ele”. Além disso, embora alguns participantes tenham sido considerados *outliers*, não há nenhuma frase que tenha provocado, por ela mesma, tempo de leitura maior.

Para além dos resultados proporcionados pela pesquisa, cabe o destaque para o diálogo efetuado entre diferentes áreas do conhecimento, porque este estudo só foi possível diante da aproximação entre Psicolinguística e Engenharia. Outros estudos, como o de Schlichting (2021), têm ressaltado a importância de relacionar diferentes áreas do saber no sentido de criar movimentos que possibilitem ampliar os horizontes investigativos, saindo de nossa área de especialização. Quando estudamos a compreensão em texto acadêmico-científico, é relevante que texto e participantes provenham da mesma área, exatamente para que seja possível um maior controle experimental. Por conta disso, seria contraproducente selecionar estudantes de cursos superiores aleatórios, já que a leitura é um fenômeno complexo e que envolve, como discutimos ao longo deste texto, diferentes aspectos e processos. Então, além das contribuições que esta investigação acrescenta nos estudos sobre leitura, sobre processamento anafórico e sobre o uso do sintagma “o mesmo” como anáfora correferencial, há que se considerar, também, o fato de investigar uma área popularmente conhecida pelo envolvimento em cálculos com ênfase nas demandas tecnológicas, mas que também apresenta necessidades voltadas para a leitura e a escrita.

Além da intersecção entre Psicolinguística e Engenharia, podemos ressaltar também os diálogos entre as Ciências Cognitivas, o Processamento em Leitura, a Linguística Textual, a Sociolinguística e o Processamento Anafórico. Todas essas áreas trouxeram conhecimentos relevantes para construir a base que sustenta a pesquisa realizada. A Psicolinguística já nasce, por si só, da intersecção entre áreas, então realizar um estudo que traga conhecimentos de diferentes áreas de especialidade da Linguística e da Psicologia também se mostra como contribuição no processo de construção do conhecimento. Isto é, buscamos ir a fundo para entender o funcionamento do sintagma “o mesmo” dentro do texto e no processamento dele de modo a compreender se, em uma área de especialidade, ele cumpre com a função que lhe tem sido atribuída: de anáfora correferencial, sem gerar problemas para a compreensão em leitura.

Todo o percurso desta pesquisa foi desenhado de forma a respeitar o método e a considerar os objetivos traçados. No entanto, vale ressaltar algumas limitações inerentes ao processo investigativo. A primeira delas é o número amostral obtido. Para pesquisas estatísticas, é interessante que o N amostral seja robusto, a fim de que seja possível fazer generalizações acerca de uma população. No entanto, quando trabalhamos com a compreensão leitora e com área específica de especialidade, nem sempre é possível ampliar demasiadamente esse número. Sem contar que a pesquisa que realizamos foi desenvolvida em etapas distintas com instrumentos distintos e era prioridade manter os mesmos participantes nos diferentes momentos. Isso também limita o número de participantes, já que nem sempre há acadêmicos com disponibilidade para participar de diferentes momentos da pesquisa. Outro ponto está no fato de termos trabalhado com a leitura de apenas um texto. Aqui, mais uma vez, entra o fato de estarmos em uma área de especialidade e já termos demandado dos participantes dois momentos distintos para a participação da pesquisa. A leitura de mais um texto exigiria que os participantes precisassem ceder mais de seu tempo à pesquisa, o que poderia limitar ainda mais o N amostral. Desse modo, optamos por equilibrar os fatores, organizando a investigação em apenas dois momentos com a leitura de um texto e contando com a participação do maior número possível de participantes.

Para pesquisas futuras, contudo, esses aspectos podem ser considerados. Assim, é interessante que outros textos sejam lidos, que o número amostral seja aumentado, sobretudo no experimento de leitura automonitorada, e que outros cursos também sejam testados no que diz respeito, mais especificamente, ao processamento do sintagma. Além disso, faz-se relevante pensar e estudar acerca do processo de compreensão leitora em diferentes áreas de especialidade, considerando as características específicas de tais áreas. Ademais, relacionar o processamento anafórico do sintagma “o mesmo” com a capacidade individual da memória de

trabalho também é questão a se considerar em outras investigações, já que a capacidade da memória de trabalho se relaciona com a manutenção e manipulação do antecedente e, por conseguinte, com o processamento da anáfora de modo a estabelecer a correferência adequadamente. A animacidade do antecedente retomado pelo anafórico “o mesmo” também é ponto a ser investigado de modo mais minucioso. Os nossos resultados mostram que esse sintagma é processado de modo semelhante ao pronome em contextos nos quais o antecedente é [-animado], mas cabe ponderar se essa lógica se materializa em contextos com antecedentes [+animados], isto é, faz-se relevante que outros estudos sejam desenvolvidos a fim de entender as restrições impostas ao uso e processamento do sintagma “o mesmo”.

Assim, entendemos que esta pesquisa acrescenta uma peça para o entendimento do quebra-cabeça da leitura, trazendo discussões acerca da relação entre o processamento de anáforas correferenciais e a compreensão. De modo mais específico, uma anáfora correferencial por muitos estigmatizada mostra cumprir a sua função, qual seja: estabelecer a coesão entre partes de um texto possibilitando a formação de uma base textual coerente. Além disso, este é mais um estudo que traz contribuições que podem subsidiar a prática pedagógica no que tange ao ensino da leitura. De modo mais específico, no que diz respeito à constituição linguística do texto e a aspecto cognitivo de processamento, fornecendo fundamentos para o trabalho com a leitura em sala de aula – seja ela do ensino superior ou não. Nesse sentido, este trabalho contribui com a sistematização de aspectos de retomada, “considerando diferentes construções anafóricas a serem realizadas” (FRANZEN; SOUZA, 2020, p.23).

Com isso, chegamos ao final deste texto e desta pesquisa, findamos uma etapa, mas não o diálogo que abrimos com este estudo. Há, ainda, muito a se discutir sobre leitura nas mais distintas áreas do saber. Há muito a se pensar sobre o processamento anafórico, sua relação com a compreensão em leitura e o papel dos diferentes anafóricos utilizados na tecitura textual. Portanto, encerramos um ciclo, mas o diálogo permanece.



## REFERÊNCIAS

ABBUHL, Rebekha.; GASS, Susan.; MACKEY, Alison. Experimental research design. In: PODESVA, Robert J.; SHARMA, Devyani (edit.). **Research Methods in Linguistics**. New York: Cambridge University Press, 2013, p. 116-134.

ADAMS, Marilyn Jager. Modeling the connections between word recognition and reading. In: ALVERMANN, Donna E.; UNRAU, Norman J.; RUDDELL, Robert B. (edit). **Theoretical Models and Processes of Reading**. 6 ed. Newark, DE: International Reading Association, 2013, p. 783-806.

AFFLERBACH, Peter; JOHNSTON, Peter. Research methodology: on the use of verbal reports in reading research. **Journal of Reading Behavior**, 1984, v.16, n.4, p.307-322.

AFFLERBACH; Peter; PEARSON; P. David; PARIS, Scott G. Clarifying differences between reading skills and reading strategies. **The Reading Teacher**, 61(5), 2008, p. 364–373.

AFFLERBACH, Peter; CHO, Byeong-Young. Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In: ISRAEL, Susan E., DUFFY, GERALD G. **Handbook of research on reading comprehension**. New York: Routledge, 2011, p.69-90.

ALLINGTON, Richard L.; MCGILL-FRANZEN, Anne. Comprehension Difficulties among Struggling Readers. In: ISRAEL, Susan E., DUFFY, GERALD G. **Handbook of research on reading comprehension**. New York: Routledge, 2011, p.551-568.

ALMEIDA, Napoleão Mendes de. **Gramática metódica da língua portuguesa**. 46ed. São Paulo: Saraiva, 2009.

ALMOR, Amit; KEMPLER, Daniel; MacDONALD; Maryellen C.; ANDERSEN, Elaine S.; TYLER, LORRAINE K. Why do Alzheimer patients have difficulty with pronouns? **Brain and Language**, 67, 1999, p. 202–227.

ALMOR, Amit. Noun-phrase anaphora and focus: The Informational Load Hypothesis. **Psychological Review**. 1999, v.106, n. 4, p. 748-765.

ALMOR, Amit. Constraints and mechanisms in theories of anaphor processing. In: PICKERING, M.; CLIFTON, C.; CROCKER, M. (Ed.). **Architectures and Mechanisms for Language Processing**. Cambridge University Press: England, 2000.

ALONSO, Jesús; MATEOS, María del Mar. Comprensión lectora: modelos, entrenamiento y evaluación. **Infancia y Aprendizaje**. 1985, 31-31, p.5-19.

ALVES, Giorvan Ânderson dos Santos. **Processamento correferencial em idosos com e sem a doença de Alzheimer**. 2012. 144 f. Tese (Doutorado) - Curso de Linguística, Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa, 2012.

AMORIM, Neide Correia S. **O item linguístico mesmo**: confrontando usos e funções no português do Brasil. 2009. 228p. Tese (Doutorado) - Curso de Linguística, Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa, 2009.

APOTHÉLOZ, Denis. Papel e funcionamento da anáfora na dinâmica textual. In: CAVALCANTE, Mônica Magalhães; RODRIGUES, Bernadete Biasi; CIULLA, Alena. **Referenciação**. São Paulo: Contexto, 2019.

APOTHÉLOZ, Denis; CHANET, Catherine. Definido e demonstrativo nas nomeações. In: CAVALCANTE, Mônica Magalhães; RODRIGUES, Bernadete Biasi; CIULLA, Alena. **Referenciação**. São Paulo: Contexto, 2019.

ARROYO, Adolfo Alonso; FUJITA, Mariângela Spotti Lopes; GIL-LEIVA, Iisidoro; PANDIELLA, Andrés. Protocolo verbal: análisis de la producción científica, 1941-2013. **Inf. e Soc.**: João Pessoa, v. 25, n. 2, p. 61-75, maio/ago., 2016.

BADDELEY, Alan. Memória de Trabalho. In: BADDELEY, A; ANDERSON, M.C.; EYSENCK, M.W. **Memória**. Porto Alegre: Artmed, 2011.

BAGNO, Marcos. **Gramática Pedagógica do Português Brasileiro**. São Paulo: Parábola Editorial, 2012.

BAKER, Linda; BEALL, Lisa Carter. Metacognitive Processes and Reading Comprehension. In: ISRAEL, Susan E., DUFFY, GERALD G. **Handbook of research on reading comprehension**. New York: Routledge, 2011, p.373-388.

BART, Daniel; DAUNAY, Bertrand. **Pode-se levar a sério o PISA?** O tratamento do texto literário em uma avaliação internacional. Campinas, SP: Mercado das Letras, 2018. Trad. Adilson Ribeiro de Oliveira, Daniella Lopes Dias Ignácio Rodrigues, Juliana Alves Assis e Maria Angela Paulino Teixeira Lopes.

BECHARA, Evanildo. **Moderna gramática do português**. Rio de Janeiro: Nova Fronteira, 2015.

BLOCK, Cathy Collins; LACINA, Jan. Comprehension Instruction in Kindergarten through Grade Three. In: ISRAEL, Susan E., DUFFY, GERALD G. **Handbook of research on reading comprehension**. New York: Routledge, 2011, p.494-509.

BONNET, Gérard. Tener presentes las singularidades lingüísticas y culturales en las evaluaciones internacionales de las competencias de los alumnos: ¿una nueva dimensión para PISA? **Revista de Educación**, “**PISA. Programa para la Evaluación Internacional de Alumnos**”. n° extraordinario, 2006, p.91-109.

BRASIL. Ministério da Educação. **Pisa 2000**: relatório nacional. Brasília, 2001.

BRASIL. Ministério da Educação. **Brasil no PISA 2015**: análises e reflexões sobre o desempenho dos estudantes brasileiros. OCDE - Organização para a Cooperação e Desenvolvimento Econômico. São Paulo: Fundação Santillana, 2016.

CARVALHO, Margarete Gonçalves Macedo de. **Avaliação da Compreensão Leitora: parâmetros e critérios essenciais para a construção de instrumentos.** Orientadora: Ana Cláudia de Souza. 2022. Tese (Doutorado) - Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina. Florianópolis, 2022.

CAVALCANTE, Liliane Carvalho Félix. **Compreensão leitora de elementos anafóricos e de humor em narrativas: um design experimental.** 2017. 212p. Tese (Doutorado) - Curso de Linguística, Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa, 2017.

CAVALCANTI, Marilda C. (1989). **Interação leitor-texto: Aspectos de interpretação pragmática.** Campinas: Editora da UNICAMP, 1989.

CHOMSKY, Noam. **Lectures on government and binding.** Dordrecht: Foris, 1981.

CHOMSKY, Noam. **Knowledge of language, its nature, origin and use.** New York: Praeger, 1986.

COHEN, Andrew D. Recent uses of mentalistic data in reading strategy research. **D.E.L.T.A.**, 1987, v.3, n.1, p. 57-84.

CONTE, Maria-Elisabeth. Encapsulamento anafórico. In: CAVALCANTE, Mônica Magalhães; RODRIGUES, Bernadete Biasi; CIULLA, Alena. **Referenciação.** São Paulo: Contexto, 2019.

CORREIA, Débora Vasconcelos. **Relações entre memória procedimental e linguagem em pessoas que gaguejam: um estudo com base no processamento da correferência anafórica em português brasileiro.** 2014. 80 f. Dissertação (Mestrado) - Curso de Linguística, Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa, 2014.

COSTA-FERREIRA, FERREIRA, Jullyane Glaicy da; FERRARI-NETO, José. Processing of coreferential relations and their relationship with the working memory: the influence of syntactic distance on anaphora processing in Brazilian Portuguese. **Journal of Language and Linguistic Studies**, 17 (Special Issue 2), 2021, p. 1012–1033.

CRESWELL, John W.; CLARK, Vicki L. Plano. **Pesquisa de métodos mistos.** Tradução: Magda França Lopes. 2 ed. Porto Alegre: Penso, 2013.

CUNHA-LIMA, Maria Luiza. **Indefinido, anáfora e a construção textual da referência.** Tese (Doutorado) - Curso de Linguística, Universidade Estadual de Campinas, Campinas, 2004, 231p.

CUNHA, Celso; CINTRA, Lindley. **Nova Gramática do Português Contemporâneo.** Nova Fronteira: Rio de Janeiro, 2001.

DANEMANN, Meredyth. Individual differences in reading skills. In: BARR, Rebecca; KAMIL, Michael L; MOSENTHAL, Peter; PEARSON, P. David. **Handbook of Reading Research.** v. II. Lawrence Erlbaum Associates: Mahwah, New Jersey, 1996, p.512-538.

DAUNAY, Bertrand; BART, Daniela. Entrevista com Bertrand Daunay e Daniel Bart: Pode-se levar o PISA a sério? Adilson Ribeiro de Oliveira. **Scripta**, Belo Horizonte, v. 23, n. 48, p. 201-215, 2º quadrimestre de 2019.

DOLE, Janice A.; NOKES, Jeffery D.; DRITS, Dina. Cognitive Strategy Instruction. In: ISRAEL, Susan E., DUFFY, GERALD G. **Handbook of research on reading comprehension**. New York: Routledge, 2011, p.347-372.

DUKE, Nell K.; PEARSON, David. Effective Practices for Developing Reading Comprehension. **Scholastic Red**. 2002, p.1-27. Disponível em: <http://faculty.washington.edu/smithant/DukeandPearson.pdf>. Acesso em: 13 mar. 2022.

ERICSSON, K. Anders; SIMON, Hebert A. Verbal Reports as Data. **Psychological Review**, 1980, v.87, n.3, p. 215-251.

ERICSSON, K. Anders; SIMON, Hebert A. **Protocols analysis: verbal reports as data**. Revised edition. Cambridge: MIT Press, 1993. 435p.

EYSENCK, Michael W. Memória semântica e conhecimento armazenado. In: BADDELEY, A; ANDERSON, M.C.; EYSENCK, M.W. **Memória**. Porto Alegre: Artmed, 2011.

FINGER-KRATOCHVIL, Claudia. **Estratégias para o desenvolvimento da consciência lexical: relações com a compreensão em leitura**. Tese (Doutorado). Florianópolis: Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, 2010.

FISCHER, Adriana.; HEINIG, Otilia Lizete de Oliveira Martins (org.). **Linguagens em uso nas engenharias**. Blumenau: Edifurb, 2014.

FLICK, Uwe. **Introdução à pesquisa qualitativa**. Tradução de Joice Elias Costa. 3 ed. Porto Alegre: Artmed, 2009. 405 p.

FRANZEN, Bruna Alexandra. **Letramentos: o dizer de engenheiros relativo ao seu campo de trabalho**. Dissertação (Mestrado). Universidade Regional de Blumenau. Blumenau: Programa de Pós-Graduação em Educação, 2012.

FRANZEN, Bruna Alexandra; SOUZA, Ana Cláudia. Processamento anafórico e leitura: revisão sistemática e ponderações para ensinar e aprender a ler. **Revista Signum: Estudos da Linguagem**, Londrina, v. 23, n. 1, 2020, p.9-30.

FOX, Emily; ALEXANDER, Patricia A. Text comprehension: a retrospective, perspective, and prospective. In: ISRAEL, Susan E.; DUFFY, Gerald G. (Eds.). **Handbook of research on reading comprehension**. Routledge, New York, 2009, p. 227-239.

GADELHA, Luísa de Araújo Pereira. **Processamento da correferência anafórica de pronomes e nomes repetidos em brasileiros aprendizes de francês como L2**. 2012. 94 f. Dissertação (Mestrado) - Curso de Linguística, Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa, 2012.

GAGNÉ, Ellen; YEKOVICH, Carol Walker; YEKOVICH, Frank R. **The cognitive psychology of school learning**. New York, USA: Harper Collins, 1993.

GARNHAM, Alan. **Mental models and the interpretation of anaphora**. London: Psychology Press Ltd, 2001.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 5 ed. São Paulo: Atlas, 2007.

GONDIM, Eva Vilma Aires Cabral. **Investigação teórico metodológica sobre a penalidade do nome repetido em português brasileiro**. Dissertação (Mestrado) – Curso de Linguística, Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa, 2017, 100p.

GOODMAN, Kenneth S. Reading: a psycholinguistic guessing game. In: SINGER, Harry; RUDDER, Robert B. (Eds.). **Theoretical Models and Processes of Reading**. Newark, Del.: International Reading Association, 1976. p. 477-508. Disponível em: [https://uascentral.uas.alaska.edu/onlinelib/Fall-2007/ED674-JD1/Goodman\\_article.pdf](https://uascentral.uas.alaska.edu/onlinelib/Fall-2007/ED674-JD1/Goodman_article.pdf). Acesso em: 23 mai. 2019.

GOODMAN, Kenneth S. Unidade na leitura – um modelo psicolinguístico transacional. **Letras de Hoje**. Porto Alegre, v. 26, n. 4, dezembro 1991. Disponível em: <http://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/16125/10586>. Acesso em: 15 jun. 2020.

GORDON, Peter C.; GROSZ, Barbara J.; GILLIOM, Laura A. Pronouns, names, and the centering of attention. **Cognitive Science**, v. 17, n. 3, p. 311-347, 1993.

GRAY, David E. **Pesquisa no mundo real**. Tradução de Roberto Cataldo Costa. 2 ed. Porto Alegre: Penso, 2012.

GOUGH, Philip B. One second of reading. In: SINGER, H.; RUDDER, R. B. (Ed.). **Theoretical models and processes of reading**. 2.ed. Newark: International Reading Association, 1976. p.509-535.

HALLIDAY, Michael A. R.; HASAN, Ruqaiya. **Cohesion in English**. London: Longman, 1976.

HANKAMER, Jorge; SAG, Ivan A. Deep and Surface Anaphora. **Linguistic Inquiry**, v. 7, n.3, p. 391-428, 1976.

HEINIG, Otilia Lizete de Oliveira Martins. O papel do professor no processo da construção de sentido na leitura. In: SOUZA, Ana Cláudia de; SEIMETZ-RODRIGUES, Cristiane; FINGER-KRATOCHVIL, Claudia; BARETTA, Luciane; BACK, Angela Cristina Di Palma (org.). **Diálogos linguísticos para a leitura e a escrita**. Florianópolis: Insular, 2019. p. 107-129.

ILARI, Rodolfo; GERALDI, João Wanderley. **Semântica**. São Paulo: Ática, 1990.

ILARI, Rodolfo. Anáfora e correferência: por que as duas noções não se identificam? **Cadernos de Estudos Linguísticos**, Campinas, (41):91-109, Jul./Dez. 2001.

JENNINGS, Joyce Holt; CALDWELL, Joanne Schudt; LERNER, Janet W. **Reading Problems: assessment and teaching strategies**. 7 ed. Boston: Pearson, 2014.

JUST, Marcel Adam; CARPENTER, Patricia A. A Theory of Reading: From Eye Fixation to Comprehension. In: ALVERMANN, Donna E.; UNRAU, Norman J.; RUDDEL, Robert B. **Theoretical Models and Processes of Reading**. 6 ed. Newark, DE: International Reading Association, 2013.

KAISER, Elsi. Experimental Paradigms in Psycholinguistics. In: PODESVA, Robert J.; SHARMA, Devyani. (Ed.). **Research Methods in Linguistics**. Cambridge: Cambridge University Press, 2013. p. 135-168.

KENDEOU, Panayiota; BROEK, Paul van Den. The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. **Memory & Cognition**. 35 (7), 2007, 1567-1577.

KENEDY, Eduardo; OTHERO, Gabriel de Ávila. **Para conhecer: Sintaxe**. São Paulo: Contexto, 2018.

KINTSCH, Walter; RAWSON, Katherine A. Compreensão. In: SNOWLING, Margaret J.; HULME, Charles. (Org.). **A ciência da leitura**. Tradução de Ronaldo Cataldo Costa. Porto Alegre: Penso, 2013. p. 227-244.

KINTSCH, Walter. Revisiting the Construction–Integration Model of Text Comprehension and Its Implications for Instruction. In: ALVERMANN, Donna E.; UNRAU, Norman J.; RUDDEL, Robert B. **Theoretical Models and Processes of Reading**. 6 ed. Newark, DE: International Reading Association, 2013, p.807-839.

KINTSCH, Walter. **Comprehension: a paradigm for cognition**. Cambridge: Cambridge University Press, 1998. 461p.

KINTSCH, Walter. The role of knowledge in discourse comprehension: A construction-integration model. **Psychological Review**, v. 95, n. 2, 1988, p.163-182.

KINTSCH, Walter; VAN DIJK, Teun. Toward a model of text comprehension and production. **Psychological Review**. v. 85, n. 5, 1978, p.363-394.

KLEIMAN, Angela. **Texto e leitor: aspectos cognitivos da leitura**. 12 ed. Campinas, SP: Pontes, 2009.

KLEIMAN, Angela. **Oficina de leitura: teoria e prática**. 16 ed. Campinas, SP: Pontes, 2016.

KOCH, Ingedore Villaça; ELIAS, Vanda Maria. **Ler e escrever: estratégias de produção textual**. 2ed. São Paulo: Contexto, 2018.

KOCH, Ingedore Villaça; ELIAS, Vanda Maria. **Ler e compreender: os sentidos do texto**. 3ed. São Paulo: Contexto, 2015.

KOCH, Ingedore Villaça; TRAVAGLIA, Luiz Carlos. **A coerência textual**. 18ed. São Paulo: Contexto, 2015.

KUMARAVADIVELU, B. Intake Facton and Intake Processes in Adult Language Learning. **Applied Language Learning**. 1994, vol. 5, nº 1, pp. 33-71. Disponível em: <https://core.ac.uk/download/pdf/145720569.pdf>. Acesso em: 26 jul. 2021.

LABERGE, David; SAMUELS, S. Jay Toward a theory of automatic information processing in reading. **Cognitive Psychology**, 1974, 6, 293-323.

LEFFA, Vilson J. Perspectivas no estudo da leitura: texto, leitor e interação social. In: LEFFA, Vilson J.; PEREIRA, Aracy, E. (orgs.). **O ensino da leitura e produção textual: alternativas de renovação**. Pelotas: Educat, 1999. p. 13-37.

LEITÃO, Márcio Martins. **O processamento do objeto direto anafórico no português brasileiro**. Tese (Doutorado) - Curso de Linguística, Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005, 160p.

LEITÃO, Márcio Martins.; LIMA, Juciane Nóbrega de; CALAÇA, Flávia G. Coreference processing: sentential links and structural parallelism. In: FRANÇA, A.I.; MAIA, M. (orgs.). **Papers in Psycholinguistics: proceedings of the first International Psycholinguistics Congress**. Rio de Janeiro: Imprinta, 2010, p. 311-316.

LEITÃO, Márcio Martins; RIBEIRO, Antonio João Carvalho; MAIA, Marcus. Penalidade do nome repetido e rastreamento ocular em português brasileiro. **Revista Linguística**. Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro. v.8, n.2, dezembro de 2012, p. 35-55.

LEITÃO, Márcio Martins. Processamento Anafórico. In: MAIA, Marcos (org.). **Psicolinguística, psicolinguísticas: uma introdução**. São Paulo: Contexto, 2015. p. 45-58.

LESLIE, Lauren; CALDWELL, Joanne. Formal and Informal Measures of Reading Comprehension. In: ISRAEL, Susan E., DUFFY, GERALD G. **Handbook of research on reading comprehension**. New York: Routledge, 2011, p.403-427.

LEZAMA, Carlos Gelormini. **Processing repeated names, overt pronouns and null reference in Spanish**. Dissertação (Mestrado em Linguística). University of South Carolina, Columbia, 2008.

MAIA, Marcus A. R. **The comprehension of object anaphora in brazilian portuguese**. Unpublished Phd dissertation, USC, 1994.

MAIA, Jefferson de Carvalho. **O processamento de expressões correferenciais em português**. Dissertação (Mestrado) - Curso de Estudos Linguísticos, Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2013, 120p.

MAIA, Jefferson de Carvalho; CUNHA-LIMA, Maria Luiza. Referenciação e técnicas experimentais: aspectos metodológicos na investigação do processamento correferencial em português brasileiro. **Revista Estudos dos Linguagem**. Belo Horizonte, v. 22, n. 1, p. 67-93, jan./jun. 2014.

MANOLI, Polyxeni; PAPADOPOULOU, Maria. Reading strategies versus reading skills: two faces of the same coin. **Procedia - Social and Behavioral Sciences**. n.46, 2012, p. 817-821.

MARCUSCHI, Luiz Antônio. Anáfora indireta: o barco textual e suas âncoras. **Revista Letras**, Curitiba, n. 56, p. 217-258. jul./dez. 2001.

MCNAMARA, Danielle; O'REILLY, Tenaha; ROWE, Michael; BOONTHUM, Chutima; LEVINSTEIN, Irwin. iSTART: A Web-Based Tutor That Teaches Self-Explanation and Metacognitive Reading Strategies. In: MCNAMARA, Danielle (org.). **Reading Comprehension Strategies: Theories, Interventions, and Technologies**. Lawrence Erlbaum Associates, Inc., Publishers: New Jersey, 2007. p. 397-420.

MELO, Maria de Fátima B. de. **O processamento da co-referência do sujeito pronominal em sentenças formadas por verbos de comunicação linguística no português do Brasil**. Tese (Doutorado). Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2003.

MITCHELL, Don C. On-line methods in language processing: introduction and historical review. In: CARREIRAS, Manuel; CLIFTON, Charles E. (ed.). **The on-line study of sentence comprehension: eyetracking, ERPs and beyond**. Nova Iorque: Psychology Press, 2004. p. 15-32.

MONDADA, Lorenza; DUBOIS, Danièle. Construção dos objetos de discurso e categorização: uma abordagem dos processos de referenciação. In: CAVALCANTE, Mônica Magalhães; RODRIGUES, Bernadete Biasi; CIULLA, Alena. **Referenciação**. São Paulo: Contexto, 2019.

MOREIRA, Emília Laudicéia. **O uso de o(s) mesmo(s) como elemento anafórico numa modalidade escrita do português do Brasil**. Dissertação (Mestrado) – Curso de Pós-Graduação em Letras, Ciências Humanas Letras e Artes, Universidade Federal do Paraná, Curitiba, 2007, 114p.

MOSS, Jarrod; SCHUNN, Christian; SCHNEIDER, Walter; McNAMARA, Danielle S.; VANLEHAN, Kurt. The neural correlates of strategic reading comprehension: Cognitive control and discourse comprehension. **NeuroImage**. 58, 2011, p. 675-686.

NÓBREGA-LIMA, Juciane. **Paralelismo e foco estrutural no processamento da correferência de pronomes e de nomes repetidos**. Dissertação (Mestrado) - Curso de Linguística, Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa, 2014, 77p.

OAKHILL, Jane; CAIN, Kate; ELBRO, Carsten. **Compreensão de leitura: teoria e prática**. Tradução e Adaptação de Adair Sobral. São Paulo: Hogrefe, 2017.

PARIS, Scott G.; LIPSON, Marjorie Y.; WIXSON, Karen. K. Becoming a strategic reader. In: RUDDELL, R. B.; RUDDELL, M. R.; SINGER, H. (Ed.). **Theoretical models and processes of reading**. 4. ed. Newark, Delaware: International Reading Association, 1994. p. 788-810.



PEREIRA, Ivelã. **Mesmo**: a funcionalidade de um item linguístico camaleônico. 2013. 313 f. Dissertação (Mestrado) – Curso de Linguística, Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2013.

PERFETTI, Charles A; LANDI, Nicole; OAKHILL, Jane. A aquisição da habilidade de compreensão da leitura. In: SNOWLING, Margaret J.; HULME, Charles. (Org.). **A ciência da leitura**. Tradução de Ronaldo Cataldo Costa. Porto Alegre: Penso, 2013. p. 227-244.

PINTO, Maria da Graça Lisboa Castro. Uma breve abordagem à leitura e à escrita na perspectiva da Psicolinguística. In: SOUZA, Ana Cláudia de; SEIMETZ-RODRIGUES, Cristiane; FINGER-KRATOCHVIL, Claudia; BARETTA, Luciane; BACK, Angela Cristina Di Palma (org.). **Diálogos linguísticos para a leitura e a escrita**. Florianópolis: Insular, 2019. p. 85-106.

PINTO, Maria da Graça Lisboa Castro. Os meandros da escrita acadêmica. Alguns recados aos estudantes universitários. **Linha D'Água (Online)**, São Paulo, v. 31, n. 1, p. 9-27, jan.-abril 2018.

PRESSLEY, Michael; AFFLERBACH, Peter. **Verbal protocols of reading**: the nature of constructively responsive reading. Hillsdale, Hove: Lawrence Erlbaum Associates, 1995. 157p.

PROCAILO, Leonilda. **Reading digital texts in L2**: working memory capacity, text mode, and reading condition accounting for differences in processes and products of reading. Tese (Doutorado). Programa de Pós-Graduação em Inglês: estudos linguísticos e literários. Universidade Federal de Santa Catarina, Florianópolis, 2017.

QUEIROZ, Karla Lima de. **Processamento da correferência**: pronomes lexicais, nomes repetidos, hiperônimos e hipônimos como formas de retomadas anafórica inter-sentencial do sujeito em Português Brasileiro. Dissertação (Mestrado) - Curso de Linguística, Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa, 2009, 61f.

RASINSKI, Timothy V.; REUTZEL, D. Ray; CHARD, David, LINAN-THOMPSON, Sylvia. Reading Fluency. In: KAMIL, Michael L; PEARSON, P. David; MOJE, Elizabeth Birr; AFFLERBACH, Peter P. **Handbook of Reading Research**. v. IV. Routledge: New York and London, 2011, p. 286-319.

RUDELL, Robert B.; UNRAU, Norman J. Reading as a motivated meaning-construction process: the reader, the text, and the teacher. In: ALVERMANN, Donna E.; UNRAU, Norman J.; RUDELL, Robert B. **Theoretical Models and Processes of Reading**. 6 ed. Newark, DE: International Reading Association, 2013, p.1015-1068.

RUMELHART, David E.; MCCLELLAND, James L. Interactive Processing Through Spreading Activation. In: LESGOLD Alan M.; PERFETTI, Charles A. (Eds.). **Interactive Processes in Reading**. Hillsdale, NJ: Lawrence Erlbaum Associates Inc., 1981. p. 37-60.

RUMELHART, David E. Toward and Interactive Model of Reading. In: ALVERMANN, Donna E.; UNRAU, Norman J.; RUDELL, Robert B. **Theoretical Models and Processes of Reading**. 6 ed. Newark, DE: International Reading Association, 2013.

SACCONI, Luiz Antônio. **Nossa gramática: teoria e prática**. 18.ed São Paulo: Atual, 1994.

SAG, Ivan A.; HANKAMER, Jorge. Toward a theory of anaphoric processing. **Linguistics and Philosophy**, 7, 1984, p. 325-345.

SAMUELS, S. Jay; KAMIL, Michael L. Models of reading process. In: PEARSON, P. D.; MOSETHAL, P.; KAMIL, M.; BARR, R. (Eds.). **Handbook of reading research**. New York: Longman, Inc. 1984.

SCHLICHTING, Thais de Souza. **Letramentos em contexto de aprendizagem ativa nas engenharias: “construindo o edifício de palavras para nele ser inquilino”**. Dissertação (Mestrado). Blumenau: Programa de Pós-Graduação em Educação, 2016.

SCHLICHTING, Thais de Souza. **Compreensão leitora e linguagem técnica no Ensino Superior: um diálogo entre Psicolinguística e Engenharia**. Orientadora: Ana Cláudia de Souza. 2021. Tese (Doutorado) - Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina. Florianópolis, 2021.

SCHÜTZE, Carson; SPROUSE, Jon. Judgment data. In: PODESVA, Robert J.; SHARMA, Devyani. (Ed.). **Research Methods in Linguistics**. Cambridge: Cambridge University Press, 2013. p. 27-50.

SCLIAR-CABRAL, Leonor. Inter-relação entre o biológico e o cultural: psicolinguística e educação. In: MAIA, Marcus (org.). **Psicolinguística e educação**. Campinas, SP: Mercado de Letras, 2018. P. 25-55.

SCOVEL, THOMAS. **Psycholinguistics**. New York: Oxford University Press, 1998.

SICURO-CORRÊA, Leticia Maria. **Aquisição da linguagem e problemas do desenvolvimento linguístico**. Rio de Janeiro: Ed. PUC-Rio; São Paulo: Loyola, 2006.

SILVA, Celia S. A natureza e a função do input na interação em sala de aula de língua estrangeira. In: MELO, Lélia E. **Tópicos de psicolinguística aplicada**. 2. ed. – São Paulo: Humanitas/FFLCH/USP, 1999, pp. 121-132.

SIMMS, Leonard J.; ZELAZNY, Kerry; WILLIAMS, Trevor F.; BERNSTEIN, Lee. Does the number of response options matter? Psychometric perspectives using personality questionnaire data. **Psychological Assessment**, 31(4), 557-566, 2019. <https://doi.org/10.1037/pas0000648>

SIMÕES, Antonia Barros Gibson. **A influência da coesão e da coerência no processamento correferencial de pronomes e nomes repetidos em português brasileiro**. 2014. 104 f. Dissertação (Mestrado) - Curso de Linguística, Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa, 2014.

SOLÉ, Isabel. **Estratégias de leitura**. 6.ed. Tradução de Cláudia Schilling. Porto Alegre: ARTMED, 1998.

SOUZA, Ana Cláudia. **Leitura, metáfora e memória de trabalho**: três eixos imbricados. Tese (Doutorado). Florianópolis: Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, 2004.

SOUZA, Ana Cláudia; RODRIGUES, Cassio. Protocolos verbais: uma metodologia na investigação de processos de leitura. In: TOMITCH, Lêda Maria Braga (org.). **Aspectos cognitivos e instrucionais da leitura**. Bauru, SP: EDUFSC, 2008.

SOUZA, Ana Cláudia. Leitura emergente: a alfabetização como chave à produção de sentidos a partir do escrito. In: SOUZA, Ana Cláudia; GARCIA, Wladimir Antônio da Costa. **A produção de sentidos e o leitor**: os caminhos da memória. Florianópolis: NUP/CED/UFSC, 2012.

SOUZA, Ana Cláudia de. Leitura, produção de sentidos e ensino. **Lendu – Linguagem, Ensino e Educação**. v. 2, n. 1, 2011. Edição Especial II SELEP. Disponível em: <http://periodicos.unesc.net/lendu/issue/view/122/showToc>. Acesso em: 25 jun. 2019.

SOUZA, Ana Cláudia de; SEIMETZ-RODRIGUES, Cristiane; WEIRICH, Helena Cristina. Ensinar a estudar ensinando a ler: potências dos roteiros de leitura. In: SOUZA, Ana Cláudia de; SEIMETZ-RODRIGUES, Cristiane; FINGER-KRATOCHVIL, Claudia; BARETTA, Luciane; BACK, Angela Cristina Di Palma (org.). **Diálogos linguísticos para a leitura e a escrita**. Florianópolis: Insular, 2019. p. 164-200.

SOUZA, Ana Cláudia de; FRANZEN, Bruna Alexandra; SCHLICHTING, Thais de Souza. Método na pesquisa psicolinguística sobre leitura: técnicas de coleta de dados. **Fórum Linguístico**, Florianópolis, v.16, n.2, p. 3849-3860, abr./jun.2019.

STERNBERG, Robert. J.; STERNBERG, Karin. **Psicologia cognitiva**. Tradução Noveritis do Brasil. 2 ed. São Paulo: Cengage Learning, 2017.

TEIXEIRA, Elisângela. Nogueira. **Preferências sintáticas e semânticas no processamento da correferência anafórica**: evidências de movimentação ocular. Tese (Doutorado) - Curso de Linguística, Departamento de Letras Vernáculas, Universidade Federal do Ceará, Fortaleza, 2013, 185p.

TOMITCH, Lêda Maria Braga. **Reading**: text organization perception and working memory capacity (Tese). Universidade Federal de Santa Catarina: Florianópolis, 1995. 366p.

TOMITCH, Lêda Maria Braga. Desvelando o processo de compreensão leitora: protocolos verbais na pesquisa em leitura. **Signo**. 2007, 32 (53), p. 42-53.

TOMITCH, Lêda Maria Braga. A metodologia da pesquisa em leitura: das perguntas de compreensão à ressonância magnética funcional. In: TOMITCH, Lêda Maria Braga (org.). **Aspectos cognitivos e instrucionais da leitura**. Bauru, SP: EDUFSC, 2008.

TOMITCH, Lêda. **Produção textual acadêmica**. Florianópolis: UFSC/CCE/LLE, 2012.

van DIJK, Teun A.; KINTSCH, Walter. **Strategies of discourse comprehension**. San Diego, California, Academic Press, 1983.

VASCONCELOS, Manuela Leitão de. **Processamento da correferência em pacientes com afasia de expressão**. 2012. 132 f. Dissertação (Mestrado) - Curso de Linguística, Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa, 2012.

## APÊNDICE A – Instruções para a realização da técnica de Protocolo Verbal

### INSTRUÇÕES ACERCA DA SUA PARTICIPAÇÃO NA PESQUISA

Você está participando de uma pesquisa sobre leitura. O estudo é constituído de duas etapas. Na primeira, realizada hoje, você fará: a) uma atividade de leitura e verbalização de pensamentos; e b) uma atividade de leitura, seguida de uma pergunta de compreensão, de uma lista de frases. A segunda etapa acontecerá no fim de novembro e consistirá na realização de duas atividades: a) leitura e interpretação de um texto; e b) leitura e avaliação de uma lista de frases.

Se desejar, você obterá um certificado de participação nesta atividade, gerado pelo sistema de registros da UFSC (Universidade Federal de Santa Catarina). Esse certificado pode ser contabilizado como ACC (Atividades Acadêmico-Científico-Culturais).

**Obrigada pela sua participação!  
Ela é fundamental ao desenvolvimento da pesquisa!**

### INSTRUÇÕES PARA A TAREFA DO PROTOCOLO VERBAL

Esta é uma atividade de leitura. Você lerá um texto, exposto parágrafo a parágrafo na tela do computador, e verbalizará seus pensamentos, de acordo com os seguintes procedimentos:

1. Leia o texto silenciosamente, com o objetivo de compreendê-lo.
2. Embora a leitura seja silenciosa, você deve interrompê-la para relatar os seus pensamentos. Sinta-se à vontade para falar sobre o que está pensando durante a leitura: você pode verbalizar dúvidas, questionamentos etc.
3. Você também deve interromper a leitura silenciosa para verbalizar sempre que encontrar um círculo vermelho, como lembrete: ●. Ele aparecerá no título e, em cada parágrafo do texto, haverá três círculos vermelhos: dois pelo meio e um posicionado ao final. Lembre-se de sempre relatar seus pensamentos nesses pontos, podendo também verbalizar quando sentir desejo ou necessidade.
4. Leia o texto e fale sobre ele e sobre o que se passar na sua mente até o final.
5. É importante que a sua verbalização seja audível e clara.
6. Leve o tempo que julgar necessário para executar a tarefa de leitura.
7. Durante a leitura do texto, para passar ao parágrafo seguinte, utilize a tecla ► e, para voltar, utilize a tecla ◀.
8. Avance e volte na leitura sempre que achar necessário.
9. Tente ler como se você estivesse sozinho. A pesquisadora não irá interferir, a menos que se passe um longo período sem que você verbalize seus pensamentos.
10. Antes de iniciar a tarefa, será feito um treinamento com a leitura de um outro texto, seguindo os procedimentos descritos acima. Durante o treinamento, você poderá conversar com a pesquisadora, que estará a sua disposição para esclarecer as dúvidas.
11. As indicações de nota que aparecerem ao longo do texto são todas relativas a referências bibliográficas.

Com a sua permissão, a sessão principal de leitura será gravada em áudio. O treino não será gravado.

## APÊNDICE B – Instruções para a realização da leitura automonitorada

### INSTRUÇÕES PARA A TAREFA DE LEITURA AUTOMONITORADA

Esta atividade consiste na leitura de frases na tela do computador. Essas frases estarão divididas em segmentos e quem controla o tempo de leitura de cada segmento é você. Cada vez que você lê um segmento e dá o comando para o próximo, o anterior desaparece. Os procedimentos serão os seguintes:

- 1- Você lerá as instruções que aparecerão na tela do computador, no momento em que a atividade iniciar.
- 2- Após as instruções, aparecerão, na tela do computador, duas frases que estarão encobertas, como no seguinte exemplo: \_\_\_\_\_ . \_\_\_\_\_
- 3- Cada parte da frase será revelada de modo progressivo, ou seja, você lerá uma parte por vez, de modo atento e na sua velocidade normal de leitura com a finalidade de compreensão.
- 4- Para revelar o primeiro segmento da primeira frase, você deve apertar a tecla verde.
- 5- Após ler o primeiro segmento, em uma velocidade natural de leitura, você, novamente, clica no botão verde, para enxergar o próximo segmento. Esse procedimento acontecerá sucessivamente, até que você tenha lido as duas frases inteiras.
- 6- Mantenha seu indicador direito em cima da tecla verde.
- 7- As frases finalizam quando o segundo ponto final aparecer, ou seja, quando terminar a segunda frase do conjunto.
- 8- Após a leitura das duas frases ter sido efetuada, você clica, novamente, na tecla verde e, então, surgirá uma pergunta de compreensão das frases.
- 9- Caso a resposta seja “sim”, você clicará na tecla verde, caso a resposta seja “não”, você clicará na tecla vermelha.
- 10- Em seguida, aparecerão as próximas duas frases.
- 11- Lembre-se de ler cada segmento em sua velocidade normal de leitura.
- 12- Durante a pergunta de compreensão, mantenha o indicador direito posicionado na tecla verde e o indicador esquerdo posicionado na tecla vermelha.
- 13- Você será avisado quando a atividade finalizar.
- 14- Antes de iniciar a tarefa, será feito um treinamento para que você se habitue com o teste. Caso apareçam dúvidas durante esse treinamento, você pode questionar a pesquisadora. Busque sanar todos os problemas que, porventura, possam aparecer. Durante o teste, não haverá interação com a pesquisadora.

## APÊNDICE C – Instruções para a tarefa de Julgamento de Aceitabilidade

### INSTRUÇÕES SOBRE A PARTICIPAÇÃO NA PESQUISA

Dando continuidade à sua participação no estudo sobre leitura na engenharia, neste segundo encontro vamos solicitar que: a) leia uma lista de frases e as avalie; b) leia um texto e responda a questões relativas a ele.





Agradecemos pela sua participação em mais este momento.

### INSTRUÇÕES PARA A REALIZAÇÃO DO TESTE DE ACEITABILIDADE

Em seu computador há um arquivo em Excel com uma lista de frases. Você deve avaliar essas frases a partir dos seguintes procedimentos:

- 1 Leia cada frase silenciosamente e avalie o quão aceitável ou adequada você a considera.
- 2 A avaliação será feita em uma escala de 1 a 7, **sendo 1 não aceitável e 7 totalmente aceitável.**
- 3 Faça um X no número da escala que você acredita representar melhor o quão aceitável ou adequada é a frase lida.
- 4 Avalie cada frase conforme o seu conhecimento e a sua percepção de falante da língua portuguesa.
- 5 Não importa se você consegue compreender a frase ou se ela faz sentido em relação ao seu conhecimento de mundo. Importa saber como você avalia cada frase em relação à língua portuguesa ou ao seu conhecimento de língua.
- 6 Não há pegadinhas nem alterações que possam tornar as frases artificiais.
- 7 Não há respostas corretas ou incorretas. O que queremos é saber a sua avaliação das frases; por isso, é essencial que você faça uma leitura atenta e que respeite suas próprias percepções.
- 8 Por gentileza, salve o arquivo conforme você for avaliando cada frase.

As frases que você lerá serão semelhantes às seguintes:

1	O edital trouxe muitas especificações para a pesquisa. Foi desenvolvido da melhor forma.	 1 2 3 4 5 6 7 
2	Os fundamentos da análise estavam comprometidos. Os dados foram copiados da internet.	 1 2 3 4 5 6 7 

## APÊNDICE D – Estímulos utilizados no julgamento de aceitabilidade e na leitura automonitorada

O edital trouxe especificações para a pesquisa. Apresenta as normas detalhadamente.  
 A apresentação não focou nos problemas educacionais. Além disso, os dados eram antigos.  
 Uma empresa estabelece objetivos constantemente. Essa prática movimenta a equipe.  
 O projeto tratava sobre controladores robustos. Tal projeto, contudo, não obteve sucesso.  
 A pesquisa foi cuidadosamente desenhada. A apresentação do método comprova o rigor.  
 O sistema de análise estava comprometido. Por isso, os resultados não eram confiáveis.  
 O texto não cumpriu com o esperado. Além disso, a temática abordada não se mostra relevante.  
 O trabalho não apresentou problemas teóricos. Sua abordagem, contudo, era polêmica.  
 A pesquisa estuda sistemas de produção. Contudo, ao ler o resumo não se compreende a abordagem.

O trabalho final deve estar claro e coerente. Para isso, o texto precisa ser bem escrito.  
 A introdução apresentou adequadamente o trabalho. Contudo, não trouxe o objetivo do estudo.  
 O laudo técnico não traz explicações. Por conta disso, o problema não pode ser solucionado.  
 Os fundamentos da análise estavam comprometidos. Os dados foram copiados descaradamente.  
 A investigação foi feita em laboratório. Não refletiu a realidade nesse contexto.  
 Os dados passaram por análise estatística. Mas os resultados não foram significativos.  
 A pesquisa sobre tecnologia tem pontos delicados. Tais pontos, portanto, devem ser reavaliados.

O projeto de pesquisa não foi bem desenvolvido. Isso aconteceu também com o trabalho final.  
 As informações não foram suficientes para o trabalho. É necessário detalhar alguns pontos.  
 A investigação trouxe resultados práticos. Havia, portanto, rigor metodológico no estudo.  
 Os processos industriais foram apresentados. Mas esse não foi o assunto tratado no trabalho.  
 A empresa, aparentemente, teve problemas com o laudo. Este não estava devidamente claro.  
 O projeto obteve financiamento. Seu foco contempla, apropriadamente, o ensino da leitura.  
 A prática requer conhecimento da teoria. A prática, além disso, orienta a teoria.  
 A introdução está bem formulada. Apresenta, de modo claro, o trabalho e o assunto mobilizado.  
 O relatório trata da situação ambiental. Mas, o texto dele não está bem estruturado.  
 O conceito está um pouco confuso. A sua teoria de apoio conflita com os estudos da área.  
 A lista mostra as sugestões apresentadas. Ela divulga com estrito rigor as próximas etapas.  
 A lista mostra as sugestões apresentadas. A mesma divulga com estrito rigor as próximas etapas.

A obra trata sobre empreendedorismo. A mesma discute com pertinência novas abordagens.  
 A obra trata sobre empreendedorismo. Ela discute com pertinência novas abordagens.  
 A aula expõe os programas de incentivo. Ela aborda especialmente a formação inicial.  
 A aula expõe os programas de incentivo. A mesma aborda especialmente a formação inicial.  
 A prova solicita os dados do estudo. A mesma exige exemplos práticos e argumentos embasados.

A prova solicita os dados do estudo. Ela exige exemplos práticos e argumentos embasados.  
 O texto define as fases da inovação. Ele exhibe, em detalhes, o papel da tecnologia.  
 O texto define as fases da inovação. O mesmo exhibe, em detalhes, o papel da tecnologia.  
 O livro foca em concepções econômicas. O mesmo debate criticamente os efeitos do capitalismo.  
 O livro foca em concepções econômicas. Ele debate criticamente os efeitos do capitalismo.  
 A ação busca novidades estratégicas. Ela sugere análises profundas e reflexões precisas.  
 A ação busca novidades estratégicas. A mesma sugere análises profundas e reflexões precisas.



A tese propõe novos métodos científicos. A mesma discorre sobre a linguagem e o ensino superior.

A tese propõe novos métodos científicos. Ela discorre sobre a linguagem e o ensino superior.

O tema permite inúmeras abordagens. Ele provoca reações diversas no campo de estudo.

O tema permite inúmeras abordagens. O mesmo provoca reações diversas no campo de estudo.

O filme mostra uma investigação real. O mesmo resume de forma coerente o trabalho feito.

O filme mostra uma investigação real. Ele resume de forma coerente o trabalho feito.

O caso explica as dúvidas apontadas. Ele levanta, também, uma discussão bem embasada.

O caso explica as dúvidas apontadas. O mesmo levanta, também, uma discussão bem embasada.

O fato ressalta a crise na ciência. O mesmo revela, com clareza, as fragilidades encontradas.

O fato ressalta a crise na ciência. Ele revela, com clareza, as fragilidades encontradas.

## APÊNDICE E – Instruções para o teste de Compreensão em Leitura

### INSTRUÇÕES PARA A REALIZAÇÃO DA ATIVIDADE DE LEITURA

Neste momento, você vai participar de uma avaliação de leitura, constituída por um texto e 10 questões a ele relacionadas. Enquanto você realiza a atividade, o *software* VLC deve ser mantido ativado. Não o desative ou feche. Suas atividades serão gravadas pelo *software*.

Quanto à avaliação, mobilize seu conhecimento prévio, mas não desconsidere o texto, salvo na primeira questão, que visa conhecer o quanto você já sabe sobre o que leu. Para que a atividade seja realizada com sucesso, observe as seguintes recomendações:

1. Leia todo o texto antes de ler e responder as questões.
2. Você pode fazer anotações e destaques no texto, utilizando os recursos do PDF, sempre que achar necessário. Pode, ainda, fazer anotações utilizando o bloco de notas, que se encontra aberto em seu computador.
3. Leia atentamente cada questão, preste atenção ao enunciado e responda tudo o que é solicitado.
4. Respeite a sequência em que as questões estão apresentadas.
5. Releia o texto ou partes dele sempre que necessário.
6. Empenhe-se na elaboração ou seleção das respostas, pois os resultados da avaliação somente corresponderão à expectativa próxima da verdadeira, se você se dedicar à resolução das questões propostas.
7. Construa as respostas às questões discursivas. Você não deve transcrever ou copiar partes do texto.
8. Certifique-se de que o cursor está sobre o campo de resposta a cada questão. Caso contrário, o que for digitado não será registrado.
9. Você dispõe do tempo máximo de 60 minutos para realizar a avaliação.
10. Não escreva seu nome nos arquivos de avaliação. Você tem um código na pesquisa, a pesquisadora irá fornecê-lo, coloque esse código no início do arquivo com as questões.
11. Não use a internet e não converse com os/as colegas nem os/as consulte durante a realização da atividade.
12. Caso ocorra algum problema, chame a pesquisadora.
13. Esclareça todas as suas dúvidas, antes de iniciar a avaliação.

**Obrigada pela sua participação. Ela é fundamental ao andamento da pesquisa.**

## APÊNDICE F – Teste de Compreensão em Leitura gabaritado

### Teste de leitura

#### Texto: “Fundamentos de detecção de defeitos em sistemas distribuídos”

Com base no texto que você acabou de ler, responda às seguintes questões:

**01.** Explique o que são os sistemas distribuídos e para que servem na área da Engenharia de Controle e Automação.

#### **Análise da Questão:**

**Aspecto:** Refletir e Analisar

**Nível:** 6 (solicita que o leitor levante hipóteses ou avalie criticamente um texto complexo sobre um assunto desconhecido, levando em consideração critérios ou perspectivas múltiplas e aplicando interpretações sofisticadas externas ao texto)

**Tipo:** aberta

#### **Respostas corretas:**

Sistemas distribuídos são sistemas compostos por dispositivos autônomos que trocam informação entre si para a realização de tarefas de forma colaborativa. A relação com a área de controle e automação se dá pelo fato de algumas tarefas não poderem ser resolvidas de maneira centralizada, por exigirem dispositivos muito complexos e caros para isso. Por outro lado, com sistemas distribuídos essas tarefas podem ser resolvidas de maneira eficiente, exigindo custos menores, porém sujeitas a maior complexidade do ponto de vista algorítmico. (1,0)

Ou:

Sistemas distribuídos são aqueles cujos componentes estão localizados em computadores interligados em rede, que se comunicam e coordenam suas ações apenas através de mensagens e que, no “ponto final”, agem como se fossem apenas um sistema. Tais sistemas se relacionam com a área de engenharia de controle e automação dada à necessidade de interligar vários equipamentos de forma a minimizar custos e aumentar a operacionalidade de uma aplicação. O profissional da área, por sua vez, precisa saber trabalhar com e *produzir* esses sistemas. (1,0)

#### **Respostas Parciais:**

Sistemas distribuídos são uma coleção de computadores (máquinas autônomas) independentes que, para o usuário, aparenta ser um sistema único, coerente e articulado. (0,5)

Sistemas distribuídos são sistemas compostos por dispositivos autônomos que trocam informação entre si para a realização de tarefas de forma colaborativa. (0,5)

Sistemas distribuídos são aqueles cujos componentes estão localizados em computadores interligados em rede, que se comunicam e coordenam suas ações apenas através de mensagens e que, no “ponto final”, agem como se fossem apenas um sistema. (0,5)

A relação com a área de controle e automação se dá pelo fato de algumas tarefas não poderem ser resolvidas de maneira centralizada, por exigirem dispositivos muito complexos e caros para isso. Por outro lado, com sistemas distribuídos essas tarefas podem ser resolvidas de maneira eficiente, exigindo custos menores, porém sujeitas a maior complexidade do ponto de vista algorítmico. (0,5)

Tais sistemas se relacionam com a área de engenharia de controle e automação dada à necessidade de interligar vários equipamentos de forma a minimizar custos e aumentar a

operacionalidade de uma aplicação. O profissional da área, por sua vez, precisa saber trabalhar com e *produzir* esses sistemas. (0,5)

Outras respostas (0,0)

**Pontuação:** 1 ponto

*Obs.: essa questão foi pensada para controlar o conhecimento prévio do leitor. O conceito de sistemas distribuídos é fundamental para a compreensão do texto (estabelecimento da coerência), mas não está apresentado de modo explícito nele.*

**02.** Com base nas informações do texto lido, faça o que se pede: a) explique o que é semântica de falhas; b) indique a quais linhas do texto você recorreu para elaborar sua resposta.

**Análise da Questão:**

**Aspecto:** Recuperar e Localizar

**Nível:** 5 (“requerem que o leitor localize e organize informações profundamente integradas, inferindo sobre quais informações no texto são relevantes”)

**Tipo:** aberta

**Resposta correta:**

É a forma como os componentes de um sistema se comportam durante uma falha, esse comportamento determina a técnica a ser utilizada na detecção de defeito de um determinado componente. Linhas 13 a 16 (1,0)

**Resposta Parcial.**

É o comportamento dos componentes de um sistema em caso de falhas. (0,5)

É a técnica a ser utilizada na detecção de defeito de um determinado componente. (0,5)

Outras respostas (0,0)

**Pontuação:** 1 ponto

*Obs.: a resposta correta à questão requer compreensão da anáfora da linha 14. Isso significa que parte da resposta requerida para esta questão será considerada para o cômputo do desempenho em compreensão leitora local.*

**03.** De acordo com o texto, qual serviço pode ser usado quando se trata da tolerância de falhas? Partindo das informações trazidas pelo texto, explique o processo de implementação desse serviço.

**Análise da Questão:**

**Aspecto:** Refletir e Analisar

**Nível:** 3 (exige que o leitor demonstre uma compreensão refinada do texto em relação a conhecimentos do cotidiano)

**Tipo:** aberta

**Resposta correta:**

Serviço de tolerância de falhas: detector de defeito. O serviço de detecção de defeitos é implantado a partir do comportamento dos componentes em caso de falhas (semântica de

falhas). Ou seja, é necessário, primeiro, compreender como as falhas afetam o comportamento dos componentes do sistema. (1,0)

Ou:

Serviço de tolerância de falhas: detector de defeito. O detector de defeitos pode ser implantado através da predição comportamental dos componentes do sistema, i.e., pode-se estimar o comportamento esperado de cada dispositivo e, quando esse comportamento é extrapolado, atribui-se ao dispositivo o estado de falha. (1,0)

*(requer uma compreensão do assunto tratado no texto e uma relação com o conhecimento de mundo dos participantes).*

**Resposta Parcial:**

A implantação do detector de defeitos deve ser feita a partir do comportamento dos componentes. (0,5)

O detector de defeito pode ser implantado em sistemas distribuídos como, por exemplo, na aplicação de compartilhamento de arquivo. (0,5)

Serviço de tolerância de falhas: detector de defeito. (0,5)

Resposta com base apenas no conhecimento prévio (0,5)

Outras respostas (0,0)

**Pontuação:** 1 ponto

**04.** Assinale apenas a alternativa correta. De acordo com o texto lido, para um bom funcionamento, é necessário que os sistemas de automação:

- a) verifiquem o funcionamento dos dispositivos ou componentes.
- b) apresentem altos tempos de respostas na detecção de defeitos.
- c) possuam mecanismos que auxiliem na tolerância de falhas.
- d) possuam temporizadores para a detecção de falhas.

**Análise da Questão:**

**Aspecto:** recuperar e localizar

**Nível:** 3 (exige que o leitor localize uma informação, integrando fragmentos do texto)

**Tipo:** fechada - múltipla escolha simples

**Resposta correta:** letra “c”, pois são esses mecanismos que proporcionam aos sistemas de automação a capacidade de tolerar falhas (característica fundamental para o bom funcionamento desses sistemas). (1,0)

Outras respostas (0,0)

**Pontuação:** 1 ponto

*Obs: para responder corretamente à questão, é necessário resolver a anáfora construída com “os mesmos” na linha 5. Isso significa que esta questão será considerada para o cômputo do desempenho em compreensão leitora local.*

**05.** Cite quais são as Hipóteses do Modelo de Sistemas Distribuídos apresentadas no texto e explique como elas podem auxiliar na construção de sistemas distribuídos tolerantes a falhas.

**Análise da Questão:**

**Aspecto:** Integrar e Interpretar

**Nível:** 6 (requer “precisão da análise e a atenção refinada a detalhes pouco perceptíveis nos textos.”)

**Tipo:** aberta

**Resposta correta:**

Comportamento temporal; comportamento de falhas dos componentes de software, dos dispositivos e dos canais de comunicação. Elas auxiliam na construção dos sistemas distribuídos de modo a determinar as propriedades do serviço que detecta os defeitos e esse processo é fundamental para situar os protocolos a serem usados na construção de sistemas distribuídos robustos (que consigam detectar defeitos e, assim, evitar falhas no sistema). (1,0)  
Ou

Comportamento temporal; comportamento de falhas dos componentes de software, dos dispositivos e dos canais de comunicação. Elas auxiliam na construção dos sistemas distribuídos de modo a determinar as propriedades do serviço. (1,0)

**Resposta Parcial:**

Comportamento temporal; comportamento de falhas dos componentes de software, dos dispositivos e dos canais de comunicação. Auxiliam na detecção dos defeitos. (0,5)

Comportamento temporal; comportamento de falhas dos componentes de software, dos dispositivos e dos canais de comunicação. (0,5)

As Hipóteses do Modelo de Sistema Distribuído auxiliam construção dos sistemas distribuídos de modo a determinar as propriedades do serviço que detecta os defeitos e esse processo é fundamental para situar os protocolos a serem usados na construção de sistemas distribuídos robustos (que consigam detectar defeitos e, assim, evitar falhas no sistema). (0,5)

Faltando alguma das hipóteses e resposta sobre o auxílio incompleta (0,5)

Outras respostas (0,0)

**Pontuação:** 1 ponto

**06.** Sobre a detecção de defeitos, assinale, de acordo com o texto lido, V para as informações verdadeiras e F para as falsas:

- ( ) A acurácia e a rapidez são requisitos que definem o desempenho do detector de defeitos.
- ( ) Os detectores de defeitos atuam monitorando o funcionamento dos componentes do sistema.
- ( ) A detecção de defeito é um mecanismo de tolerância a falhas seguro para os sistemas distribuídos porque não consome recursos computacionais.
- ( ) A implementação dos detectores de defeitos não depende da semântica das falhas.
- ( ) A detecção de defeitos é uma das formas de prevenir a falha nos sistemas de automação.
- ( ) A detecção de defeitos é uma etapa dos processos de tolerância a falhas.

**Análise da Questão:**

**Aspecto:** Integrar e Interpretar

**Nível:** 3 (requer o entendimento da relação ou a construção do significado de uma palavra ou oração.)

**Tipo:** fechada - múltipla escolha complexa

**Resposta correta:** V V F F V V (1,0)

**Respostas Parciais:**

Cada item: 0,167

Outras respostas (0,0)

**Pontuação:** 1 ponto

*Obs: o item 4 requer a adequada resolução da anáfora da linha 14. Isso significa que esse item será considerado para o cômputo do desempenho em compreensão leitora local.*

**07.** De acordo com o texto lido, por que a detecção de defeitos é um desafio a ser enfrentado quando se trata de sistemas distribuídos?

**Análise da Questão:**

**Aspecto:** Refletir e Analisar

**Nível:** 5 (requer avaliação crítica ou levantamento de hipóteses, com base em conhecimento especializado)

**Tipo:** aberta

**Resposta correta:**

Porque, em sistemas distribuídos, as falhas são parciais. Isso significa que alguns componentes falham enquanto outros continuam funcionando, o que torna a detecção mais complexa. Além disso, as propriedades e o desempenho do serviço de detecção de defeitos dependem das garantias da qualidade do serviço entregues pelo sistema, a detecção pode não ser confiável por acabar fornecendo informações incorretas ou altos tempos de resposta na detecção. (1,0)

Ou:

Porque como, nos ambientes de automação distribuídos, as propriedades e o desempenho do serviço de detecção de defeitos dependem das garantias da qualidade do serviço entregues pelo sistema, a detecção pode não ser confiável por acabar fornecendo informações incorretas ou altos tempos de resposta na detecção. (1,0)

**Resposta Parcial:**

Porque as propriedades e o desempenho do serviço de detecção de defeitos dependem das garantias da qualidade do serviço entregues pelo sistema. (0,5)

Porque, em sistemas distribuídos, as falhas são parciais. Isso significa que alguns componentes falham enquanto outro continuam funcionando, o que torna a detecção mais complexa. Um dos requisitos para um sistema ser considerado distribuído é o fato de possuir independência em relação às falhas, ou seja, cada dispositivo pertencente ao sistema deve funcionar mesmo que algum outro componente apresente falha. O que torna muito mais difícil a implementação de um sistema de detecção de falhas (*resposta apenas com base em conhecimento prévio*) (0,5)

Outras respostas (0,0)

**Pontuação:** 1 ponto

**08. a)** Diga a que se referem, no texto, as seguintes expressões:

*los* (linha 4): \_\_\_\_\_

os mesmos (linha 5): \_\_\_\_\_  
 Ela (linha 10): \_\_\_\_\_  
 desses sistemas (linha 21): \_\_\_\_\_  
 Ele (linha 46): \_\_\_\_\_  
 seus (linha 46): \_\_\_\_\_

**08. b)** É possível afirmar que essas expressões auxiliam na continuidade e na coesão do texto? Por quê?

**Análise da Questão:**

**Aspecto:** Recuperar e Localizar

**Nível:** 4 (“requerem que leitor localize e organize diferentes tipos de informação integrada”)

**Tipo:** aberta

*(requer integração de informação local e o estabelecimento da coesão textual a fim de construir a coerência do assunto apresentado)*

**Resposta correta:**

- a) Sistemas de automação; sistemas de automação; detecção de defeitos; sistemas de automação (modernos); desempenho (de um detector de defeitos); detector de defeitos.
- b) Sim, pois todas elas estão relacionadas diretamente à temática do texto (detecção de defeitos em sistemas de automação (distribuídos)) e todas elas exigem a compreensão de um elemento anteriormente citado, garantindo a progressão textual. (1,0)

**Respostas Parciais:**

Cada item anafórico: 0,1

Segunda pergunta: 0,4

Outras respostas (0,0)

**Pontuação:** 1 ponto

*Obs1: no que diz respeito aos dois primeiros anafóricos, foi considerada como resposta adequada apenas aquelas que apresentaram o sintagma todo “sistemas de automação”. Isso porque, no texto, antes dos dois anafóricos solicitados, há uma retomada feita com “desses sistemas” (que se refere a sistemas de automação). Assim, entende-se que quem respondeu apenas “sistemas” tomou como referente adequado “desses sistemas” e, nesse caso, está equivocado. Em outros contextos, considerou-se parte de sintagma como correto porque, dentro do contexto textual, é possível perceber que o participante de fato compreendeu qual o referente.*

*Obs2: cada anafórico que compõe o item “a” será considerado separadamente para o cômputo do desempenho em compreensão leitora local.*

**09.** As expressões “o mesmo”, que aparecem nas linhas 14 e 33 do texto, referem-se, respectivamente, a quê? Explique, com as suas palavras, a informação trazida em cada um dos trechos em que essas expressões aparecem.

**Análise da Questão:**

**Aspecto:** Recuperar e Localizar

**Nível:** 4 (“requerem que leitor localize e organize diferentes tipos de informação integrada”)

**Tipo:** aberta

**Resposta correta:**



Comportamento (dos componentes em caso de falhas); Serviço (de detecção de defeitos).

Linha 14: O comportamento dos componentes de um sistema é denominado de semântica de falhas é essa semântica que vai determinar a técnica de verificação a ser utilizada pelo serviço de detecção de defeitos. Linha 33: O serviço de detecção de defeitos pode, muitas vezes, não ser confiável por apresentar informações incorretas ou altos tempos de resposta. (1,0)

**Resposta Parcial:**

Comportamento (dos componentes); Serviço (de detecção de defeitos). (0,5)

Linha 14: O comportamento dos componentes de um sistema é denominado de semântica de falhas é essa semântica que vai determinar a técnica de verificação a ser utilizada pelo serviço de detecção de defeitos. Linha 33: O serviço de detecção de defeitos pode, muitas vezes, não ser confiável por apresentar informações incorretas ou altos tempos de resposta. (0,5)

Comportamento (dos componentes). Linha 14: O comportamento dos componentes de um sistema é denominado de semântica de falhas é essa semântica que vai determinar a técnica de verificação a ser utilizada pelo serviço de detecção de defeitos. (0,5)

Serviço (de detecção de defeitos). Linha 33: O serviço de detecção de defeitos pode, muitas vezes, não ser confiável por apresentar informações incorretas ou altos tempos de resposta. (0,5)

Outras respostas (0,0)

**Pontuação:** 1 ponto

*Obs: cada anafórico e cada explicação solicitada serão considerados separadamente para o cômputo do desempenho em compreensão leitora local.*

**10.** Releia o texto atentamente e elabore uma síntese das ideias principais dele, ou seja, faça um resumo. O que se espera de você é que retextualize, com suas próprias palavras, as principais ideias do texto lido.

**Análise da Questão:**

**Aspecto:** Integrar e Interpretar (essa questão tem por objetivo captar se o participante formou um modelo situacional adequado do texto lido)

**Nível:** 6 (requer que o leitor demonstre a compreensão completa do texto, envolvendo a integração das informações)

**Tipo:** aberta

**Resposta correta:**

O acadêmico deve explicar que os sistemas de automação precisam ter a capacidade de tolerar falhas e que a detecção de defeitos é um dos mecanismos utilizados nesse processo (explicar o a implementação desse serviço). No entanto, existem sistemas que possuem uma complexidade maior na implementação de detectores de defeitos: os sistemas distribuídos. Essa implantação é mais difícil porque as falhas são parciais, então é preciso que haja garantias da qualidade do serviço. Uma das formas de contornar essa complexidade é usando as hipóteses do modelo de sistema distribuído. (1,0)

**Resposta Parcial:**

O texto aborda o assunto da detecção de defeitos como um dos mecanismos de tolerância de falhas em sistemas de automação distribuídos. (0,5)

Outras respostas (0,0)

**Pontuação:** 1 ponto

*Obs geral: alguns participantes confundem os termos "falha" e "defeito" (por exemplo, em alguns casos, respondem: serviço de detecção de falhas (em vez de serviço de detecção de defeitos)). Esse tipo de equívoco não foi descontado, porque, embora, na área, "falha" e "defeito" possuam especificações/sentidos diferentes, os participantes (uma parte ao menos) parecem não ter esse conhecimento prévio e isso não está especificado no texto.*

## APÊNDICE G – Texto utilizado nas tarefas de coleta de dados

### FUNDAMENTOS DE DETECÇÃO DE DEFEITOS EM SISTEMAS DISTRIBUÍDOS

Os sistemas de automação requerem a capacidade de tolerar falhas, já que defeitos ou mau funcionamento podem implicar prejuízos financeiros, ambientais ou, até mesmo, perda de vidas humanas. Por melhores que sejam as técnicas usadas na construção desses sistemas, diversas questões podem levá-los a apresentar defeitos: desgastes físicos de componentes, falhas humanas, fenômenos e/ou desastres naturais etc. Portanto, os mesmos devem possuir mecanismos adequados de tolerância a falhas. Os termos “falha”, “erro” e “defeito”, utilizados neste texto, seguem o padrão de terminologia amplamente adotado pela comunidade científica internacional da área de tolerância a falhas <sup>[1]</sup>.

A detecção de defeitos é um serviço utilizado para construir mecanismos de tolerância a falhas. Ela consiste na verificação do funcionamento dos dispositivos ou componentes, de modo que, quando não se comportam de acordo com objetivos especificados, ações possam ser tomadas para substituição de componentes defeituosos, recuperação ou reconfiguração do sistema. O comportamento dos componentes em caso de falhas guia a forma de implementação dos detectores de defeitos. O mesmo recebe a denominação de semântica de falhas<sup>[2]</sup>, que determina a técnica de verificação (ou monitoramento) a ser utilizada pelo serviço de detecção de defeitos. Por exemplo, se um dispositivo do sistema pode falhar produzindo valores incorretos, testes podem ser usados para comparar as saídas do dispositivo com saídas de referência. De outro lado, se um dispositivo falha somente omitindo valores de saída (temporária ou permanentemente), temporizadores podem ser usados para detectar defeitos <sup>[3]</sup>.

A natureza descentralizada de sistemas de automação modernos tem impacto no projeto de detectores de defeitos. Em alguns desses sistemas, os dispositivos de automação estão distribuídos e interconectados a partir de redes de comunicação, por exemplo: Sistemas de Controle Sobre Rede (NCS, *Networked Control Systems*); Sistemas de Controle Distribuídos (DCS, *Distributed Control Systems*)<sup>[4]</sup>, aplicações distribuídas sobre Redes Veiculares Ad Hoc (VANET, *Vehicular Ad hoc Networks*)<sup>[5]</sup>, Sistemas Ciber-Físicos (CPS, *Cyber-Physical Systems*)<sup>[6]</sup>, entre outros. A natureza descentralizada traz uma série de benefícios, tais como a possibilidade da supervisão remota, a interconexão, a adição e a remoção de novos dispositivos, a facilidade de integração entre sistemas etc. Todavia, nesses ambientes de automação distribuídos, as propriedades e o desempenho do serviço de detecção de defeitos dependem das garantias de qualidade de serviço (QoS, *Quality of Service*) entregues pelos sistemas de processamento e de comunicação utilizados.

O serviço de detecção de defeitos, nesse sentido, pode não ser confiável em alguns ambientes. O mesmo pode acabar fornecendo informações incorretas ou apresentando altos tempos de respostas na detecção de defeitos de componentes do sistema. Isso acontece se a infraestrutura de computação subjacente possuir períodos em que não há garantias com relação aos tempos de processamento e transmissão das mensagens, o que gera a possibilidade de comprometimento no funcionamento de certos sistemas (ou subsistemas) de automação com requisitos temporais. Portanto, as hipóteses do modelo de sistema distribuído - isto é, o comportamento temporal e o comportamento de falhas dos componentes de *software*, dos dispositivos e dos canais de comunicação - são usadas para determinar as propriedades do serviço de detecção de defeitos durante sua operação. Essas propriedades ou garantias

influenciam, por sua vez, as características dos protocolos usados na construção de sistemas distribuídos robustos – como, por exemplo, consenso distribuído<sup>[7]</sup> e difusão atômica<sup>[8]</sup>; duas técnicas usadas na construção de estratégias de replicação para tolerância a falhas<sup>[2]</sup>.

O desempenho de um detector de defeitos é definido em termos de sua acurácia e rapidez. Ele é adequado se os seus parâmetros operacionais forem configurados de forma a atender as demandas das aplicações, mediante as condições de carga e disponibilidade dos recursos computacionais do ambiente distribuído. Detectores de defeitos consomem por si recursos computacionais para o processamento e a transmissão das mensagens de monitoramento e requerem parâmetros operacionais adequados para o seu funcionamento; por isso, eles podem degradar o desempenho da aplicação ou do próprio serviço de detecção, fornecendo informações incorretas ou tempos de detecção inadequados, se os intervalos entre verificações dos defeitos não forem definidos de forma apropriada.

(Adaptação de excerto de texto de Macêdo; Sá, 2018<sup>[9]</sup>)

## Referências

- [1] AVIZIENIS, A. et al. Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and Secure Computing*, New York, v. 1, n. 1, p. 11-33, Jan./Mar. 2004.
- [2] CRISTIAN, F. *Understanding fault-tolerant distributed systems*. Communications of the ACM, New York, v. 34, n. 2, p. 56-78, Feb. 1991.
- [3] JALOTE, P. *Fault tolerance in distributed systems*. New Jersey: Prentice Hall, 1994.
- [4] LIAN, F.-L.; MOYNE, J.; TILBURY, D. Network design consideration for distributed control systems. *IEEE Transactions on Control Systems Technology*, New York, v. 10, n. 2, p. 297-307, Mar. 2002.
- [5] WANG, Y.; LI, F. Vehicular ad hoc networks. In: MISRA, S.; WOUNGANG, I.; MISRA, S. C. (Ed.). *Guide to Wireless Ad Hoc Networks*. London: Springer, 2009. p. 503-525. (Computer Communications and Networks).
- [6] SHA, L. et al. Cyber-physical systems: A new frontier. In: TSAI, J. J. P.; YU, P. S. (Ed.). *Machine Learning in Cyber Trust: security, privacy, and reliability*. New York: Springer, 2009. p. 3-13.
- [7] GUERRAOUI, R. et al. Consensus in asynchronous distributed systems: a concise guided tour. In: KRAKOWIAK, S.; SHRIVASTAVA, S. (Ed.). *Advances in distributed systems: from algorithms to systems*. Berlin: Springer, 2000. p. 33-47. (Lecture Notes in Computer Science, v. 1752).
- [8] DÉFAGO, X.; SCHIPER, A.; URBÁN, P. Total order broadcast and multicast algorithms: Taxonomy and survey. *ACM Computing Surveys*, New York, v. 36, p. 372-421, December 2004.
- [9] MACÊDO, R. J. de A.; SÁ, A. S. Fundamentos de Detecção de Defeitos em Sistemas. In: MACÊDO, R. J. de a.; FARINES, J-M. A (org.). *Projeto de Sistemas Distribuídos e de Tempo Real para Automação*. Salvador: EDUFBA, 2018.