

**PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS -
MESTRADO E DOUTORADO
ÁREA DE CONCENTRAÇÃO EM LEITURA: ESTUDOS LINGUÍSTICOS,
LITERÁRIOS E MIDIÁTICOS**

Jordan Junges

**MULTIMODALIDADE SENSORIAL COMO AUXÍLIO OU OBSTÁCULO À
COMPREENSÃO LINGUÍSTICA**

Santa Cruz do Sul

2026

Jordan Junges

**MULTIMODALIDADE SENSORIAL COMO AUXÍLIO OU OBSTÁCULO À
COMPREENSÃO LINGUÍSTICA**

Dissertação apresentada ao Programa de Pós-Graduação em Letras – Mestrado e Doutorado, Área de Concentração em Leitura: Estudos Linguísticos, Literários e Midiáticos, Linha de Pesquisa em Estudos Linguísticos e Cognição, Universidade de Santa Cruz do Sul – UNISC, como requisito parcial para obtenção do título de Mestre em Letras.

Orientadora: Profa. Dra. Rosângela Gabriel

Santa Cruz do Sul

2026

CIP - Catalogação na Publicação

Junges, Jordan

Multimodalidade sensorial como auxílio ou obstáculo à
compreensão linguística / Jordan Junges. – 2026.

278f. : il. ; 29 cm.

Dissertação (Mestrado em Letras) – Universidade de Santa Cruz
do Sul, 2026.

Orientação: Profa. Dra. Rosângela Gabriel.

1. Compreensão leitora. 2. Multimodalidade. 3. Atenção. 4.
Memória de trabalho. 5. Eye tracking. I. Gabriel, Rosângela. II.
Título.

Jordan Junges

**MULTIMODALIDADE SENSORIAL COMO AUXÍLIO OU OBSTÁCULO À
COMPREENSÃO LINGUÍSTICA**

Esta dissertação foi submetida ao Programa de Pós-Graduação em Letras – Mestrado e Doutorado; Área de Concentração em Leitura: Estudos Linguísticos, Literários e Midiáticos; Linha de Pesquisa em Estudos Linguísticos e Cognição, Universidade de Santa Cruz do Sul – UNISC, como requisito parcial para obtenção do título de Mestre em Letras.

Dra. Rosângela Gabriel

Professora orientadora - UNISC

Dra. Rejane Frozza

Professora examinadora - UNISC

Dr. José Ferrari Neto

Professor examinador - UFPB

Santa Cruz do Sul

2026

AGRADECIMENTOS

À minha companheira Ana, que tornou possível este projeto de vida quando uma série de circunstâncias o tornaram extremamente necessário. Muito obrigado por todo o amor, a compreensão e o suporte durante esta desafiadora jornada.

Na intersecção entre a família e a instituição de ensino, agradeço a duas pessoas em especial, que me abasteceram com todos os materiais didáticos imagináveis e que só não ficaram mais surpresas do que eu mesmo quando decidi iniciar este mestrado. À minha cunhada Rosi, uma veterana do PPGL que compartilhou comigo conselhos fundamentais. E à minha sogra Dercy, uma das mães fundadoras do Curso de Letras da Unisc e professora de muitos dos meus professores, que sempre me ajudou com tudo que esteve ao seu alcance. Infelizmente, não deu tempo de ela acompanhar a conclusão deste percurso, mas o seu entusiasmo com a minha formação me acompanhou e me acompanhará em todas as horas.

Aproveito para lembrar as demais pessoas que já partiram e que vivem para sempre, com toda certeza, nas memórias que compartilhamos delas e nos impactos das ações que realizaram. Faço um destaque ao meu padrinho Adair, que faleceu na época mais crítica das enchentes no RS, quando a condição das estradas não permitiu que viajássemos para nos despedir dele. E à minha mãe Marlene, que desde 2013 faz uma falta tremenda, não podendo nos dar novos conselhos e nem celebrar as vitórias conosco. Ela certamente estaria orgulhosa, porque sempre tratou nossa educação e nosso bem-estar como prioridades.

Aos homens de três gerações da minha família: meu pai Irineu Donato, meu irmão Jessé e meu sobrinho Davi, de quem pude estar geograficamente mais perto nesse período. Obrigado pela motivação, pelo reconhecimento e por me garantirem momentos alegres de distração quando era preciso fazer uma pausa.

Aos meus amigos de músicas e risadas, sem os quais a coisa toda perde o sentido. Valeu por (quase) sempre entenderem quando eu estava afundado demais em leituras ou tentando escrever algo que fizesse sentido. Em breve, voltamos à programação normal.

Faço um agradecimento abrangente a todos os servidores públicos federais que lutaram para instituir e manter os direitos de afastamento para participação em programa de pós-graduação *stricto sensu* e de incentivo à qualificação. Da

instituição que me emprega, a Universidade Federal de Santa Maria, agradeço a toda a cadeia de setores e pessoas que aprovaram o meu licenciamento, sobretudo ao meu colega e chefe imediato Jonathan, que não só não encontrou qualquer empecilho como garantiu que o processo fosse encaminhado de forma rápida e segura. Meu reconhecimento, também, aos colegas da Coordenadoria de Comunicação Social e, mais diretamente, do Núcleo Rádio Universidade que adaptaram e absorveram as tarefas que eu realizaria durante esse tempo.

Ao curso de extensão on-line “Introdução à Neurociência da Música”, da Universidade Federal do ABC, no qual tive o prazer de estudar e que foi crucial para eu conseguir visualizar um caminho para mim na pesquisa acadêmica.

Ao Grupo de Pesquisa Linguagem e Cognição, que me proporcionou conviver com pessoas tão brilhantes quanto amáveis, que cooperaram comigo de inúmeras formas. Aos bolsistas de Iniciação Científica Rodrigo e Dani, que foram os primeiros a participar dos nossos testes-piloto. À Mari e à Vitória, que só posso chamar de calouras porque o acaso me fez iniciar o curso antes delas, já que estão quilômetros à frente em todo o restante. Ao Brendon, por ter a melhor biblioteca e por fazer questão de indicar e emprestar os livros mais pertinentes, além de ter colaborado com dedicação com um segundo teste-piloto que fizemos e de ter deixado valiosos apontamentos sobre ele. À professora, pós-doutora e cantora Aline, que acompanhou com muito carinho a minha turma e, ainda mais agora, vai poder fazer isso por muitos outros alunos. À Monique, colega com quem foi uma satisfação produzir e apresentar tantos trabalhos de disciplinas, que nem considerou tomar um fôlego antes de avançar para o doutorado. E à Kadine, que me ajudou em basicamente tudo: me acolheu desde o meu primeiríssimo contato com o *eye tracker*, foi sujeito de teste em um dos pilotos, recrutou alunos para participar da pesquisa e compartilhou, sem reservas, o seu extenso conhecimento para melhorar meus capítulos de teoria, metodologia e análise.

À Profa. Dra. Régine Kolinsky e à Université Libre de Bruxelles, que, por uma felicidade, possibilitaram que usássemos, aqui, um dispositivo de rastreamento ocular, algo que contribuiu muito com a nossa pesquisa e atendeu a uma antiga curiosidade pessoal minha. Obrigado, também, à Ma. Julia Justino, que gentilmente nos repassou importantes dicas para operarmos esse equipamento.

Aos professores que me honraram e alegraram aceitando participar como avaliadores da minha banca, até mesmo porque já tinham tido impacto fundamental

no percurso desta dissertação. A Profa. Rejane Frozza já havia deixado valiosas contribuições por ocasião da banca de qualificação e, sem o empenho dela, provavelmente não teríamos conseguido tantos participantes para a pesquisa, sobretudo no recorte das Ciências Exatas. E o Prof. José Ferrari Neto, que conheci em um desses grandes golpes de sorte, quando ele ministrava um curso de análise de dados na pesquisa em leitura em um congresso na cidade de Lajeado, justamente quando eu começava a escrever o capítulo de resultados e precisava de várias diretrizes. Além de ter apresentado o software estatístico Jamovi e de ter ensinado o seu funcionamento naquele evento, o professor ainda voltou a nos aconselhar de forma mais personalizada em uma reunião do grupo de pesquisa.

À minha estimada professora orientadora Rosângela, por ter aceitado este enferrujado aluno, vindo da afastada área da Comunicação Social e com uma intenção de pesquisa um tanto destoante. Muito obrigado por ser essa referência para os âmbitos acadêmico e pessoal, capaz de lidar com os problemas mais graves sempre com uma atitude de adaptação e superação, sem jamais perder a elegância e a leveza. Nem me iludo que seja possível chegar nesse mesmo nível, mas é uma boa meta distante a se perseguir.

Aos participantes da pesquisa, é claro, pois, sem eles, não haveria dados novos para discutir. Agradeço por terem nos doado seu valioso tempo, já que o mais simples a fazer era terem seguido com suas rotinas de atividades ou de descanso.

Ao Programa de Pós-Graduação em Letras, uma nova família que ganhei. Aos excelentes professores, em todas as dimensões, que tive o prazer de conhecer. Vou sentir muita falta, acima de tudo, dos debates em sala de aula. À Luana, secretária do Programa, que é um verdadeiro ponto de referência para todos nós alunos, resolvendo todo tipo de questão burocrática enquanto mantém um ambiente aconchegante para a convivência de todos. Por fim, deixo um agradecimento especial para todas as minhas colegas (a maioria esmagadora de mulheres me autoriza e me obriga a falar no feminino), que fazem o mesmo que eu, mas cuidando de cônjuge, filhos, pais, departamentos escolares, turmas de aula e outros empregos, produzindo resultados acadêmicos iguais ou melhores do que os meus. É um mistério como vocês conseguem isso. Se a amostra que eu vi representa, razoavelmente, o todo, podemos ter muita fé no futuro da educação.

*Escrever sobre música
é como dançar sobre arquitetura
(Autor desconhecido).*

RESUMO

A compreensão leitora, capacidade que une a compreensão linguística geral com o reconhecimento e a decodificação visual das palavras (Hoover; Gough, 1990; Scarborough, 2001), é uma habilidade fundamental para se viver plenamente em nossa sociedade letrada. Apesar desse status, as atividades de leitura, que demandam grande quantidade de processamento cognitivo, constantemente se desenvolvem em situações muito abaixo das ideais, quase sempre em disputa com a infinidade de distrações do mundo contemporâneo. A atenção é a porta de entrada que define quais estímulos – entre eles, a leitura – recebem tratamento cognitivo privilegiado e podem chegar à consciência e também às memórias. Quando se trata de materiais linguísticos, ou, mais amplamente, semióticos, esses estímulos costumam vir misturados em múltiplas modalidades sensoriais simultâneas. Tendo isso em vista, este estudo tem como objetivo analisar possíveis interferências, desde a promoção de auxílio até a geração de obstáculo, provocadas a partir da associação, em um único texto, das modalidades sensoriais visual e auditiva à compreensão linguística de estudantes universitários. A fim de conduzir a investigação em uma população universitária ($n = 25$), desenhou-se um teste composto por nove narrativas curtas (com cerca de 150 palavras cada), apresentadas em nove formatos diferentes (escrito, falado, escrito e falado, em vídeo legendado, em vídeo narrado, cantado, com música de fundo instrumental, com distrações sonoras, e com distrações visuais) para medir a compreensão linguística enquanto processo, a partir de rastreamento ocular, e enquanto produto, a partir de questionários com perguntas abertas curtas. Juntamente, foi aplicado um questionário de perfil, direcionado para conhecer hábitos de leitura e de consumo de mídias pelos participantes, e um questionário de percepção sobre preferências percebidas durante o teste de compreensão leitora. As métricas de tempo de reação para resposta aos questionários de compreensão e de acurácia não constataram quaisquer diferenças estatisticamente significativas entre as condições experimentais; porém, esta última medida revelou uma tendência de a leitura silenciosa ser mais eficiente do que os formatos de texto falado, cantado e acompanhado por música instrumental. As informações de *eye tracking* apontaram nivelamento entre todas as condições experimentais baseadas em texto estático, ocultando possíveis variações de processamento da memória de trabalho nesse

recorte, e foram percebidas diferenças em formatos de áudio e de vídeo, com destaque à maior dilatação das pupilas, cuja interpretação não é conclusiva. Considerando a totalidade dos dados, inclusive as preferências dos participantes, o formato de texto com melhor e mais consistente desempenho foi o unimodal visual. Em geral, os resultados indicam que estudantes universitários processam a compreensão linguística com grande flexibilidade para lidar com multimodalidade e distrações. A quantidade de modos sensoriais ocupados pela informação comunicada parece ter impacto menor do que parâmetros como concisão, estruturação e eficácia nativa da codificação aplicada. Estudos futuros com textos informacionalmente mais densos, instrumentos regulados para a medição de processos mais complexos de compreensão leitora, ou considerando públicos com leitura menos proficiente e menor adaptação a distrações provavelmente apresentarão conclusões distintas que ajudarão a colocar os achados desta pesquisa em perspectiva.

Palavras-chave: Compreensão linguística. Compreensão leitora. Multimodalidade. Atenção. Memória de trabalho. Eye tracking.

ABSTRACT

Reading comprehension, the ability that combines general language comprehension with the visual recognition and decoding of words (Hoover; Gough, 1990; Scarborough, 2001), is a fundamental skill for living fully in our literate society. Despite this status, reading activities, which demand a large amount of cognitive processing, are constantly carried out in suboptimal scenarios, almost always competing against the myriad of distractions of the contemporary world. Attention is the gateway that defines which stimuli – among them, reading – receive privileged cognitive treatment and can reach consciousness and also memory. When it comes to linguistic, or more broadly, semiotic materials, these stimuli are usually mixed in multiple simultaneous sensory modalities. With this in mind, this study aims to analyze possible interferences, ranging from providing assistance to generating obstacles, caused by the association, in a single text, of visual and auditory sensory modalities over the language comprehension of college students. In order to conduct an investigation on an undergraduate population ($n = 25$), a test was designed consisting of nine short narratives (with approximately 150 words each), presented in nine different formats (written, spoken, written and spoken, in subtitled video, in narrated video, sung, with instrumental background music, with sound distractions, and with visual distractions) to measure language comprehension as a process, applying eye tracking, and as a product, using questionnaires with short open-ended questions. Additionally, a profile questionnaire was applied, aimed at understanding the participants' reading and media consumption habits, and a perception questionnaire about perceived preferences during the reading comprehension test. The reaction time metrics for answering the comprehension questionnaires and the accuracy scores did not show any statistically significant differences among the experimental conditions; however, the latter measure revealed a tendency for silent reading to be more efficient than the spoken, sung, and instrumental music-accompanied text formats. Eye tracking data showed leveled effects across all experimental conditions based on static text, masking possible variations in working memory processing in this segment. Differences were observed in audio and video formats, particularly larger pupil dilation, whose interpretation is not conclusive. Considering all data, including participant preferences, the unimodal visual text format performed best and most consistently. In general, the results indicate that

undergraduate students process language comprehension with great flexibility in dealing with multimodality and distractions. The number of sensory modes occupied by the communicated information seems to have less impact than parameters such as conciseness, structuring, and native effectiveness of the applied encoding. Future studies with more informationally dense texts, instruments calibrated to measure more complex reading comprehension processes, or considering audiences with less proficient reading and less adaptability to distractions will likely present different conclusions that will help put the findings of this research into perspective.

Keywords: Language comprehension. Reading comprehension. Sensory multimodality. Attention. Working memory. Eye tracking.

LISTA DE ILUSTRAÇÕES

Figura 1 – Área de trabalho global	37
Figura 2 – Esquema revisado da memória de trabalho	52
Figura 3 – Esquema do simbolismo linguístico	62
Figura 4 – Modelo de cordas de Scarborough (2001)	75
Figura 5 – Redes corticais da leitura	77
Figura 6 – Passagens narrativas e questionários de Daneman e Carpenter (1980)	98
Figura 7 – Formatos de exibição de texto por condição experimental	102
Figura 8 – Esquema da pesquisa experimental	112
Figura 9 – Resultado da calibração do rastreador ocular efetuada em um participante	114
Quadro 1 – Sequenciamento dos estímulos para cada participante	115
Quadro 2 – Esquema de avaliação das respostas	117
Figura 10 – AOIs delimitadas no estímulo 1A	121
Figura 11 – AOIs delimitadas em trecho do estímulo 1D	122
Figura 12 – Idades dos participantes (n = 25)	125
Figura 13 – Semestre cursado pelos participantes	126
Figura 14 – Número de indivíduos pela frequência de leitura por tipo de material	127
Figura 15 – Número de indivíduos pela frequência de consumo de mídias ao dia	128
Figura 16 – Número de indivíduos pela maneira preferida para compreender conteúdos	129
Figura 17 – Número de indivíduos pela ambientação preferida para atividade	130
Figura 18 – Tempo de resposta aos questionários por participante	135
Figura 19 – Tempo total (leituras de textos e resposta aos questionários) por participante	135
Figura 20 – Resultado do Teste de Friedman para Tempo (TQ) por participante	137
Figura 21 – Tempos por condição experimental	139
Figura 22 – Classificação das condições experimentais pelo tempo (Q) para resposta aos questionários de compreensão leitora	140

Figura 23 – Resultado do Teste de Friedman para tempo (T) por condição experimental	142
Figura 24 – Resultado do Teste de Friedman para tempo (Q) por condição experimental	143
Figura 25 – Resultado do Teste de Friedman para Tempo (TQ) por condição experimental	144
Figura 26 – Médias de pontuação por texto de cada participante	148
Figura 27 – Médias marginais estimadas para pontuação total por participante	149
Figura 28 – Classificação das condições experimentais pela acurácia nas respostas aos questionários de compreensão leitora	151
Figura 29 – Pontuações (total e estratificada por natureza da questão) por condição experimental	152
Figura 30 – Médias marginais estimadas para pontuação total por condição	153
Figura 31 – Porcentagens de fixações, sacadas e “outros” durante leitura por participante	163
Figura 32 – Porcentagens de fixações, sacadas e “outros” durante leitura por condição	164
Figura 33 – Número de fixações e sacadas por segundo por participante	165
Figura 34 – Número de fixações e sacadas por segundo por condição	165
Figura 35 – Diâmetro médio da pupila por participante	166
Figura 36 – Diâmetro médio da pupila por condição	167
Figura 37 – Duração média das fixações por participante	168
Figura 38 – Duração média das fixações por condição	168
Figura 39 – Amplitude média das sacadas por participante	169
Figura 40 – Amplitude média das sacadas por condição	170
Figura 41 – Velocidade de pico média de sacadas por participante	171
Figura 42 – Velocidade de pico média de sacadas por condição	171
Figura 43 – Porcentagens de tempo de fixação em AOI de resposta (referente ao tempo total de leitura) e de número de fixações em AOI de	

resposta (referente ao número total de fixações) por participante	173
Figura 44 – Porcentagens de tempo de fixação em AOI de resposta (referente ao tempo total de leitura) e de número de fixações em AOI de resposta (referente ao número total de fixações) por condição	173
Figura 45 – Número de votos por participante para melhor condição, a partir da narrativa escolhida (n = 27)	177
Figura 46 – Número de votos por participante para pior condição, a partir da narrativa escolhida (n = 26)	178
Figura 47 – Número de votos por participante para pior ordem do estímulo, a partir da narrativa escolhida (n = 26)	179
Figura 48 – Número de votos totais para condição mais fácil (n = 50)	181
Figura 49 – Número de votos totais para condição mais difícil (n = 50)	182
Figura 50 – Médias marginais estimadas entre classes de modalidade para tempo de questionário	187
Figura 51 – Tempo de questionário entre classes unimodal e bimodal	188
Figura 52 – Médias marginais estimadas entre classes de modalidade para pontuação total em cada texto	189
Figura 53 – Pontuação total por texto entre classes unimodal e bimodal	190
Figura 54 – Associação entre tempo de resposta aos questionários e pontuação no teste de compreensão leitora	198
Figura 55 – Tempo de questionário entre classes unimodal, submodal e bimodal	204
Figura 56 – Pontuação total por texto entre classes unimodal, submodal e bimodal	204
Figura 57 – Mapas de calor com a soma das fixações para todas as condições D e E para o texto 1	206
Figura 58 – Tempo de questionário entre classes unimodal, bimodal(-) e submodal(-)	211
Figura 59 – Pontuação total por texto entre classes unimodal, bimodal(-) e submodal(-)	211

Figura 60 – Pontuação total por texto considerando presença de perturbação	212
Figura 61 – Mapa de calor com a soma de todas as fixações no estímulo 6l	216
Figura 62 – Pontuação total por texto considerando presença de silêncio	221
Figura 63 – Pontuação total por texto considerando a modalidade de sentido que contém o conteúdo verbal	222
Figura 64 – Tempo de resposta aos questionários por ordem do estímulo	224
Figura 65 – Pontuação total por ordem do estímulo	225
Figura 66 – Porcentagem de número de fixações sobre AOI de resposta, com relação ao tempo de leitura, por ordem do estímulo	227
Figura 67 – Pontuação total por narrativa	228
Figura 68 – Pontuação em questões curtas (R1, R2 e R3) por narrativa	229
Figura 69 – Pontuação total geral (máx = 270) por gênero	232
Figura 70 – Médias marginais estimadas para pontuação total geral (máx = 270) por área de estudo	233
Figura 71 – Médias marginais estimadas para tempo de resposta aos questionários por ambientação preferida para atividade de trabalho	236

LISTA DE TABELAS

Tabela 1 – Idade dos participantes por área de estudo	125
Tabela 2 – Descrição dos tempos por participante	132
Tabela 3 – Descrição dos tempos por condição	137
Tabela 4 – Comparação do tempo (TQ) do participante com a média por condição experimental	145
Tabela 5 – Descrição da pontuação total por participante	147
Tabela 6 – Descrição das pontuações por condição	150
Tabela 7 – Médias marginais estimadas para pontuação total por condição	154
Tabela 8 – Comparação da pontuação total do participante com a média por condição experimental	155
Tabela 9 – Medidas de tendência central para as métricas de rastreamento ocular	158
Tabela 10 – Matriz de Correlações entre as medidas de rastreamento ocular (apenas condições A, G, H e I)	160
Tabela 11 – Matriz de Correlações entre as medidas de rastreamento ocular (incluindo todas as condições experimentais)	161
Tabela 12 – Frequências de escolhas para melhor texto	175
Tabela 13 – Frequências de escolhas para melhor texto	175
Tabela 14 – Proporções de votos para condições fáceis	183
Tabela 15 – Proporções de votos para condições difíceis	183
Tabela 16 – Médias marginais estimadas para pontuação total por texto considerando presença de perturbação	213

LISTA DE ABREVIATURAS E SIGLAS

AC	acurácia
AIC	<i>Akaike Information Criterion</i>
ANOVA	<i>Analysis of Variance</i>
AOI	<i>Area of Interest</i>
BIC	<i>Bayesian Information Criterion</i>
CAAE	Certificado de Apresentação para Apreciação Ética
CI	Construção-Integração
CEP	Comitê de Ética em Pesquisa
DP	desvio-padrão
EEG	eletroencefalograma
ENADE	Exame Nacional de Desempenho de Estudantes
ENEM	Exame Nacional do Ensino Médio
EP	erro-padrão
EVA	Etileno Acetato de Vinila
FBI	<i>Federal Bureau of Investigation</i>
fMRI	<i>functional magnetic resonance imaging</i>
GAMLj3	<i>General analyses for linear models (Version 3.6.4)</i>
GPT-4o	<i>Generative Pre-trained Transformer 4 Omni</i>
IC	intervalo de confiança
IIQ	intervalo interquartil
INAF	Indicador Nacional de Alfabetismo Funcional
PET	<i>positron emission tomography</i>
ppm	palavras por minuto
Q-Q	quantil-quantil
SAS	sistema atencional supervisor
TCLE	Termo de Consentimento Livre e Esclarecido
TDAH	transtorno do déficit de atenção com hiperatividade
TR	tempo de reação
ULB	<i>Université Libre de Bruxelles</i>
UNISC	Universidade de Santa Cruz do Sul
VWFA	<i>Visual Word Form Area</i>
YARC	<i>York Assessment of Reading for Comprehension</i>

SUMÁRIO

1 INTRODUÇÃO.....	20
2 REFERENCIAL TEÓRICO.....	24
2.1 Atenção e consciência.....	24
2.1.1 Sistemas de atenção.....	26
2.1.2 Processamento inconsciente.....	32
2.1.3 Pensamento consciente.....	35
2.2 Memórias.....	39
2.2.1 Classificação das memórias pelo seu conteúdo.....	41
2.2.2 Operações sobre as memórias.....	44
2.2.3 Memórias de longo e curto prazo.....	47
2.2.4 Memória de trabalho.....	50
2.2.5 Chunking.....	53
2.3 Multimodalidade e linguagem.....	57
2.3.1 Signos e linguagens não verbais.....	59
2.3.2 Linguagem verbal.....	66
2.3.3 Redundância, irrelevância e interferência entre modalidades.....	68
2.4 Leitura.....	72
2.4.1 Compreensão leitora.....	79
2.4.2 Instrumentos de avaliação da compreensão leitora.....	85
3 METODOLOGIA.....	92
3.1 Objetivos.....	92
3.1.2 Objetivo geral.....	92
3.1.3 Objetivos específicos.....	92
3.2 Hipóteses.....	93
3.3 Participantes.....	94
3.4 Instrumentos.....	95
3.4.1 Questionário de perfil do participante.....	95
3.4.2 Textos para compreensão leitora.....	96

3.4.2.1 Modalidades e versões de apresentação dos textos.....	100
3.4.3 Questionários de compreensão leitora.....	107
3.4.4 Dispositivo de rastreamento ocular (eye tracker).....	107
3.4.5 Questionário de percepção sobre o teste de compreensão leitora.....	109
3.5 Procedimentos de produção dos dados.....	109
3.6 Computação dos dados coletados.....	117
4 APRESENTAÇÃO E ANÁLISE DOS DADOS.....	124
4.1 Resultados do Questionário de perfil do participante.....	124
4.2 Resultados dos questionários de compreensão leitora.....	131
4.2.1 Tempo de reação nos Questionários de compreensão leitora.....	131
4.2.2 Pontuações nos questionários de compreensão leitora.....	146
4.3 Resultados de rastreamento ocular (eye tracking).....	156
4.4 Resultados do questionário de percepção sobre o teste de compreensão leitora.....	174
4.5 Comparações entre unimodalidade e bimodalidade.....	184
4.6 Multimodalidade e carga sobre a memória de trabalho.....	195
4.7 Comparações com condições submodais.....	203
4.8 Inserção de elementos não redundantes contínuos e intermitentes.....	210
4.9 Possíveis fatores alternativos.....	219
5 CONSIDERAÇÕES FINAIS.....	239
REFERÊNCIAS.....	250
APÊNDICE A – Termo de Consentimento Livre e Esclarecido (TCLE).....	260
APÊNDICE B – Questionário de perfil do participante.....	262
APÊNDICE C – Textos e testes para compreensão leitora.....	268
APÊNDICE D – Questionário de percepção sobre o teste de compreensão leitora.....	274
APÊNDICE E – Instruções dadas ao participante antes do início do teste presencial (conforme apresentadas em um tablet).....	276
APÊNDICE F – Gabarito simplificado para correção dos questionários de compreensão leitora.....	277

APÊNDICE G – Recortes de pontuações por condição experimental e classe de modalidade.....	278
--	------------

1 INTRODUÇÃO

A atenção é um recurso limitado e limitante, porque apenas permite que se pense profundamente a respeito de um recorte mínimo da realidade presente ou do universo imaginado. Ainda que existam pessoas que se consideram plenamente capazes de executar múltiplas tarefas complexas simultaneamente e sem perda de desempenho, essa é uma alegação que não suporta uma análise mais rigorosa: a ciência assegura que a mente só pode lidar com um pensamento consciente por vez (Dehaene, 2014, 2020). E, se essa restrição já representa um desafio até em cenários mais controlados, mesmo sem distrações externas significativas, ainda podemos nos perder nas próprias memórias e reflexões. Apesar dessa limitação, nossa cultura parece empenhada em criar ambientes cada vez mais bem desenhados para competir continuamente pela atenção de cada indivíduo, culminando, até o momento, nas apelativas e onipresentes mídias digitais e móveis.

A despeito desse cenário conturbado, uma tradicional expectativa da sociedade segue inalterada: todo cidadão tem o dever de aprender cada vez mais para que ajude a coletividade sendo o mais produtivo possível. E é a leitura que, há séculos, transmite as informações mais densas e completas, de modo que, como defende Morais (2019), dominar essa capacidade é a única forma de, verdadeiramente, acessar os próprios direitos e de ocupar os espaços de poder nesta cultura letrada. Ler e compreender o que é lido, uma habilidade tipicamente associada ao silêncio e à concentração, é algo que tem conseguido se adaptar às condições contemporâneas, ou tem simplesmente sofrido perdas?

É uma preocupação que inquieta estudiosos há décadas. De acordo com Vasilev, Kirkby e Angele (2018), ela vem desde, pelo menos, os anos 1940, quando aparelhos de rádio portáteis se tornaram populares entre estudantes. Sem jamais desaparecer do radar dos pesquisadores e em constante renovação – incluindo testes dos efeitos sobre a compreensão leitora, provocados por desde ruídos de trânsito até interações com aparelhos de televisão e mídias sociais –, o tema se espalha para as conversas cotidianas e até mesmo para as políticas públicas. Para aproveitar um acontecimento recente e amplamente divulgado, basta recordar que os aparelhos celulares foram proibidos nas salas de aula brasileiras por força de lei federal em 2025 (Brasil, 2025).

De fato, há um bom número de estudos que chegam a uma conclusão que parece consideravelmente intuitiva: estímulos concorrentes são prejudiciais à compreensão leitora (Salamé; Baddeley, 1989; Armstrong; Boiarsky; Mares, 1991; Perham; Currie, 2014; Vasilev; Kirkby; Angele, 2018; Santosh *et al.*, 2024). Entretanto, há que se considerar outros relatos bastante comuns de pessoas que preferem estudar ou trabalhar enquanto escutam música ou deixam algum telejornal sendo exibido na mesma sala. Poderia haver uma explicação razoavelmente simples para isso? Talvez esses ruídos acionados deliberadamente sejam apenas mais previsíveis e, assim, serviriam para encobrir eventuais distrações não controladas e passíveis de causar mais surpresas? Ou, por uma linha de raciocínio consideravelmente diferente, o que porventura possa acontecer é que esses estímulos extras operam no nível emocional, simplesmente aumentando o ânimo do leitor?

Como já existem dúvidas sobre os efeitos da distração nesses cenários mais típicos, quando a tarefa principal (a leitura) e a secundária (ouvir música ou programas de TV, comumente) envolvem duas informações não relacionadas, uma outra situação causa ainda mais incertezas: e quanto aos casos nos quais uma mesma informação chega por dois meios diferentes? Nessas circunstâncias, em que se aproveita a bimodalidade sensorial para, por exemplo, enviar mensagens idênticas ou equivalentes para a visão e a audição, a compreensão linguística talvez possa ganhar um impulso? Mousavi, Low e Sweller (1995) encontraram evidências de que um material educativo era aprendido de maneira superior quando apresentado simultaneamente em modalidades sensoriais distintas, uma conclusão semelhante à de Mayer e Moreno (1998), que realizaram testes com vídeos instrucionais que continham ou legendas ou narração, esta última combinação com melhores resultados. Brasel e Gips (2014), de maneira semelhante, usaram audiovisuais publicitários e perceberam que os espectadores tinham mais facilidade para recordar informações sobre a marca e o produto anunciados quando havia legendas redundantes exibidas na tela.

Resulta daí, então, uma controvérsia: por um lado, o processamento de um texto que ocupa as entradas visual e auditiva pode, muito bem, sobrecarregar o limitado sistema cognitivo do sujeito, prejudicando a compreensão; mas, por outro, é plausível que potencialize o entendimento da mensagem, com ambas as modalidades formando um sistema redundante que confere mais segurança para a

integridade das informações. No centro da questão, estão o funcionamento e os limites da memória de trabalho, conforme o quadro teórico de Baddeley e Hitch (1974), bem como a maneira pela qual a atenção e a consciência processam estímulos complexos, repletos de significados a serem decodificados, guardados e reelaborados.

Considerando as dúvidas que atravessam o tema, chegamos ao problema de pesquisa: há benefícios para a compreensão linguística quando o texto é recebido simultaneamente em duas modalidades, visual e auditiva? A partir de uma reflexão inicial, a hipótese de trabalho que direcionou a investigação foi a de que os textos multimodais, quando apresentam informações redundantes e não contêm elementos demasiadamente distrativos, proporcionam vantagens para a compreensão do leitor/ouvinte, porque fazem proveito de mais canais de recepção disponíveis ao sujeito e, assim, potencializam as chances de que o conteúdo receba atenção, além de favorecer que os dados trafegando em diferentes modos sejam combinados de maneiras mais significativas e memoráveis.

Diante disso, o objetivo geral desta dissertação é analisar possíveis interferências, desde a promoção de auxílio até a geração de obstáculo, provocadas a partir da associação, em um único texto, das modalidades sensoriais visual e auditiva à compreensão linguística de estudantes universitários. Especificamente, serão analisados e comparados os desempenhos de leitores e ouvintes em uma série de apresentações distintas de textos, desde as condições unimodais (leitura ou audição) até o que se entende como os piores cenários possíveis (leitura perturbada por sons ou imagens intermitentes), passando por condições de multimodalidade mais típicas (material audiovisual).

Para realizar essa investigação, utilizamos um método experimental, empregando testes de compreensão leitora – aplicados, de igual forma e sem prejuízo, para aferir a compreensão linguística em sentido amplo, quando o componente do reconhecimento e decodificação visual das palavras (Hoover; Gough, 1990; Scarborough, 2001) está ausente – e relacionando esses dados com medidas de rastreamento ocular coletadas e respostas a questionários adicionais focados em identificar o perfil e os hábitos e preferências dos sujeitos. O público selecionado é composto por alunos universitários da cidade de Santa Cruz do Sul - RS, e os materiais apresentados são curtas passagens narrativas.

Este trabalho se divide em cinco capítulos. Ao concluirmos esta introdução, avançamos para o referencial teórico, que está organizado em quatro seções principais. Na primeira delas, são discutidos os conceitos de atenção e consciência; a segunda parte trata das classificações das memórias e outros temas relacionados; a terceira seção aborda tópicos relativos à multimodalidade sensorial e à linguagem; e a quarta parte encerra aprofundando a temática da leitura, da sua compreensão e de como esta pode ser medida. O capítulo 3 é o momento em que a metodologia é explicada em detalhes, informando objetivos, hipóteses e procedimentos escolhidos para a coleta e análise de dados. O quarto capítulo reúne os resultados e os complementa com a discussão daquilo que foi encontrado. Por último, o capítulo 5 traz a exposição das considerações finais.

2 REFERENCIAL TEÓRICO

Com o propósito de organizar a base conceitual deste trabalho de forma linear e progressiva, optou-se por iniciar pelos aspectos mais amplos da cognição, que são compartilhados até mesmo com outros animais, e avançar até tópicos mais relacionados com criações culturais humanas. Dessa maneira, visando chegar, por fim, às questões específicas da leitura, a seção seguinte discute os conceitos elementares da atenção e da consciência.

2.1 Atenção e consciência

Ainda que sejam tratadas, muitas vezes, como sinônimas, atenção e consciência são funções cognitivas diferentes. É certo que apresentam algumas semelhanças e superposições de tarefas desempenhadas; no entanto, é importante definir as divisões mais evidentes e mais relevantes entre as duas.

Tsuchiya e Koch (2009) indicam a relação íntima entre os dois construtos quando exemplificam esta interação mais típica: ao se prestar atenção a um objeto, os atributos deste se tornam conscientes e, no sentido inverso, esse objeto desaparece da consciência quando a atenção é dirigida para outra coisa. Esses autores trabalham a partir da definição bastante restritiva da atenção como sendo “atenção seletiva”, deixando de lado, expressamente, outras acepções comuns, tais como alerta ou níveis de excitação. A atenção, assim admitida por eles, é uma forma de selecionar a informação que seja relevante para o estado presente do organismo, recortando uma pequena fração de toda a complexa realidade que o cerca no momento. A consciência, segundo Tsuchiya e Koch (2009, p. 64, tradução nossa¹), tem atribuições diferentes:

[...] resumir todas as informações relevantes relativas ao estado atual do organismo e seu ambiente e tornar esse sumário compacto acessível aos estágios de planejamento do cérebro, detectar anomalias e erros, tomada de decisões, linguagem, inferir o estado interno de outros animais,

¹ As traduções presentes nesta dissertação são de responsabilidade do autor.

estabelecer metas de longo prazo, elaborar modelos recursivos, e pensamento racional.²

Convictos de que atenção e consciência apresentam diferentes funções e, portanto, necessariamente devem ser processos diferentes, Tsuchiya e Koch (2009) avançam e identificam situações em que elas são acionadas ou conjuntamente ou de forma dissociada. Se, para a dilatação da pupila em resposta à luz, podem ser dispensadas ambas a consciência e a atenção *top-down* (descendente; controlada voluntariamente, nos centros hierarquicamente superiores do cérebro), e se, no outro extremo, a capacidade de reportar completamente um evento pressupõe que atenção e consciência trabalhem juntas, existem notáveis casos intermediários. Pela interpretação dos autores, a atenção sozinha serve para se efetuar algumas buscas visuais ou ter pensamentos diversos, ao passo que a consciência é condição suficiente para que ocorra a reportabilidade parcial de uma percepção ou para se captar a essência (“*gist*”, no original, p. 66) de uma cena. Tsuchiya e Koch (2009) apresentam, adicionalmente e de forma curiosa, situações em que atenção e consciência inclusive trabalham uma contra a outra. O conhecido fenômeno de “ter algo na ponta da língua” seria um exemplo disso, já que a concentração demasiada na tarefa de achar uma palavra exata, muitas vezes, atrapalha justamente que ela ascenda até o nível da consciência. Não raro, o termo correto é recuperado assim que a pessoa pensa em qualquer outro assunto.

Dehaene (2014) apresenta um entendimento um tanto diferente. Em primeiro lugar, ele identifica pelo menos três conceitos que a ciência contemporânea atribui à consciência, sendo eles: vigília, a condição de se estar acordado; atenção, no sentido de concentração da mente sobre um pedaço da informação disponível; e acesso consciente “[...] o fato de que algumas das informações focadas pela atenção finalmente entram em nosso conhecimento³ e se tornam reportáveis para outros” (Dehaene, 2014, p. 8-9⁴). O autor considera que apenas esta última

² Do original: [...] *summarizing all relevant information pertaining to the current state of the organism and its environment and making this compact summary accessible to the planning stages of the brain, to detecting anomalies and errors, decision-making, language, inferring the internal state of other animals, setting long-term goals, making recursive models, and rational thought* (Tsuchiya; Koch, 2009, p. 64).

³ O termo em inglês utilizado no texto original é “*awareness*”, comumente traduzido para o português como “consciência”.

⁴ Do original: [...] *the fact that some of the attended information eventually enters our awareness and becomes reportable to others* (Dehaene, 2014, p. 8-9).

definição, o acesso consciente, representa a consciência propriamente dita, pois é um estágio mais alto, para o qual estar em vigília e atento pode não bastar.

Há um cuidado extremo por parte de Dehaene (2014) para manter o conceito de consciência dentro de limites bem definidos a fim de que seja um objeto razoavelmente claro para ser tratado pela ciência experimental. Sendo assim, o acesso consciente, também explicado por ele como quando a mente toma posse de algo, é o significado estrito da consciência. A propósito da atenção, o autor, da mesma forma que Tsuchiya e Koch (2009), restringe-a ao conceito mais específico da “atenção seletiva”, ou seja, ela é o filtro complexo que isola um dos possíveis pensamentos e, então, permite que este adentre a consciência.

Considerando as posições teóricas expostas (Tsuchiya; Koch, 2009; Dehaene, 2014), é possível perceber que são bastante próximas nos pontos mais críticos. Entretanto, as principais divergências são a respeito da classificação de algumas operações mentais como totalmente conscientes ou portadoras de apenas algum nível de pré-consciência. De forma mais marcada, Dehaene (2014) coloca a atenção seletiva como condição indispensável para a formação de um pensamento consciente, o que diverge da noção de Tsuchiya e Koch (2009), que concebe alguns estados de consciência ocorrendo livres de atenção.

2.1.1 Sistemas de atenção

Qual o propósito básico da atenção? Segundo Dehaene (2020), é resolver a saturação de informações, um problema muito comum e antigo. A vantagem trazida por esse mecanismo, de acordo com o autor, fez com que o sistema evoluísse em uma grande variedade de espécies animais que, inclusive, fazem uso de circuitos muito aproximados daqueles que há nos humanos. Dehaene (2020) aponta que os sistemas de atenção funcionam em um arranjo piramidal, no qual muitos neurônios distintos lidam, em paralelo, com os mais diferentes estímulos (visão, audição, olfato, tato, entre outros) e vão efetuando sucessivos estágios de triagem até que se filtre e se mantenha a informação considerada a mais essencial. Apenas dessa forma é que se torna possível processar algo em profundidade, porque os recursos do cérebro jamais seriam suficientes para dispensar o mesmo tratamento para os milhões de *bits* chegando a cada segundo (Dehaene, 2020). Tsuchiya e Koch (2009) compartilham dessa noção da existência de sobrecarga informacional e, citando um

exemplo específico do sentido visual, comentam que cada olho carrega cerca de um *megabyte* por segundo de dados brutos. Somente uma pequena fração desse *input* é processada em tempo real; o restante é tratado em uma largura de banda reduzida, ou seja, aquilo que não recebe atenção acaba por ser negligenciado (Tsuchiya; Koch, 2009).

Um aspecto que muitas vezes é esquecido é que a atenção não apenas lida com uma quantidade gigantesca de informações que chegam de fora; ela também sofre competição por aquilo que está dentro do organismo. Um cenário típico que se pode visualizar nesse sentido é quando uma pessoa subitamente se recorda de algo importante, e então uma memória rouba toda a concentração, que passa a não apontar mais para o mundo exterior. Como diz Dehaene (2014), a todo momento os nossos sentidos trabalham com um sem número de potenciais percepções e, paralelamente, em nossa memória está uma abundância de conhecimento que pode, da mesma forma, emergir até o nível da consciência. É o que acontece também com outras operações introspectivas, como o raciocínio sobre um problema abstrato ou na “fala interna” (Dehaene, 2014), aquela voz que se manifesta dentro da nossa mente quando falamos com nós mesmos, em silêncio objetivo, sem produzir vibrações acústicas materiais.

Outra situação que relembra que o universo interior da mente tem muito peso na forma como as informações são processadas é quando a atenção é regida por um controle totalmente consciente. Tsuchiya e Koch (2009) consideram que a atenção seletiva pode se basear por dois tipos de fatores: exógenos, ou *bottom-up* (ascendente; dos sensores externos para os centros superiores da mente), ou endógenos, ou *top-down* (no sentido inverso, iniciando do alto da hierarquia cognitiva). As pistas exógenas que podem capturar a atenção são as características imanentes do evento ou objeto que o tornam salientes ao restante dos elementos. No caso da visão, conforme referido pelos autores, atributos como cor, orientação, textura e movimento tem grande potencial natural para atrair o olhar. Porém, quando há a decisão superior de se concentrar sobre qualquer ponto específico – seja uma região delimitada, uma determinada característica ou um objeto em particular –, a atenção endógena pode suplantar muitas das respostas mais básicas e até instintivas aos estímulos externos (Tsuchiya; Koch, 2009).

A atenção, por mais que possa ser experienciada como algo unificado, está presente em vários circuitos cerebrais (Dehaene, 2020). Três sistemas ou redes

atencionais se destacam. Elas foram apresentadas em Posner e Petersen (1990) – com algumas diferenças nessa esquematização original –, refinadas em Petersen e Posner (2012) e sintetizadas por Dehaene (2020), que, então, as identifica como alerta, orientação e controle executivo.

O primeiro sistema de atenção, de acordo com a explicação de Dehaene (2020), é o de alerta, ocupado com o *quando* estar atento, controlando os níveis de vigília. O autor afirma que os sinais emitidos por essa rede são capazes de mobilizar todo o corpo, quando necessário. O sistema de alerta está muito ligado a eventos extremos, como traumas, situações de risco à vida ou outras fontes de emoções fortes. Filmes ou, especialmente, jogos eletrônicos podem estimular o alerta em condições seguras, fornecendo um contexto favorável para potencializar funções cognitivas e aprendizados diversos (Dehaene, 2020).

Posner e Petersen (1990) indicam que essa rede atencional faz com que o indivíduo responda mais rapidamente, mas a desvantagem é que os erros se tornam mais comuns; todavia, Dehaene (2020, p. 152) aparentemente contradiz que seja sempre esse o caso, pois “Um jogador de videogame consegue tomar decisões ultrarrápidas sem comprometer seu desempenho”.⁵ Dehaene (2020) aponta que o estado de alerta é capaz de modular significativamente o aprendizado de duas formas: primeiro, há uma facilitação para a escrita dos eventos daquele momento na memória, e segundo – comprovado, pelo menos, em testes com camundongos – a liberação de diferentes neuromoduladores é capaz de induzir a recuperação momentânea de alguma plasticidade no córtex cerebral.

A segunda rede atencional é a da orientação, que amplifica o *quê* é objeto da atenção (Dehaene, 2020). Ela funciona, utilizando a metáfora trazida por Dehaene (2020), como um holofote, que “ilumina”, ou amplifica, o objeto que é mais relevante para as metas que se apresentam. Essa comparação, como salienta o próprio autor, é insuficiente, porque o cérebro trabalha tanto aumentando a incidência de “luz” sobre o item atendido quanto aumentando a “escuridão” sobre tudo aquilo que não interessa no momento.

Talvez a apresentação mais prototípica do sistema orientacional da atenção é o cenário em que uma pessoa move sua cabeça e seus olhos para a precisa localização de um estímulo visual de interesse. Mas o processo muitas vezes é bem

⁵ Do original: *A video game player manages to make ultra-fast decisions without compromising his or her performance* (Dehaene, 2020, p. 152).

mais elaborado, porque, como observa Dehaene (2020), ele opera em todas as modalidades sensoriais e sobre quaisquer representações que o cérebro consiga elaborar e categorizar. Ainda considerando estímulos visuais, manter os olhos completamente fixos não impede a pessoa de se concentrar sobre algum volume que esteja na visão periférica, a cor de um item, as partes de um objeto, e assim por diante. Isso está perfeitamente em linha com a atenção *top-down*, a supremacia dos fatores endógenos, conforme a classificação de Tsuchiya e Koch (2009) apresentada anteriormente. Também tem relação com o fenômeno chamado “cegueira desatencional” (Dehaene, 2014), que ocorre quando o cérebro está tão focado em algo que deixa de perceber coisas inesperadas que, sem sombra de dúvidas, entraram no campo de visão e foram recebidas adequadamente pelos olhos. Um exemplo célebre desse efeito é demonstrado pelo experimento do “gorila invisível”, desenvolvido por Simons e Chabris em 1999 (Dehaene, 2014, 2020), em que parte dos participantes, completamente focados em contar o número de passes entre jogadores de basquete em um filme, sequer toma conhecimento do absurdo aparecimento de um homem fantasiado de macaco que se movimenta e ocupa o centro da tela em algum momento.

Para pensar sobre o sentido da audição, um exemplo que é muito interessante e ilustrativo é o “efeito coquetel” (*cocktail party effect*, em inglês), que acontece, como explicado por Dehaene (2014, 2020), quando uma pessoa é capaz de se concentrar em apenas uma conversa em meio a uma dúzia de outras vozes que falam simultaneamente ao seu redor, em uma cena estereotípica de uma festa. Considerando-se que o efeito ocorre até mesmo quando a intensidade da voz alvo e das intrusas está muito próxima, a orientação da atenção não se dá puramente pela característica básica do volume, mas sim por construções mais complexas, como significado e timbre.

O último dos três sistemas de atenção listados por Dehaene (2020) é o controle executivo, que, por sua vez, determina *como* a informação atendida será processada. Essa rede é um tanto quanto diferente das outras duas em sua natureza e complexidade. Conforme Petersen e Posner (2012), sua conceptualização foi sofisticada de forma considerável desde que foi rotulada como “detecção de alvo” em artigo dos mesmos autores vinte anos antes (Posner; Petersen, 1990). O controle executivo é, segundo Dehaene (2020), também um sistema atencional porque seleciona uma entre várias opções disponíveis; a grande

diferença é que a decisão é entre operações mentais que podem ser usadas no momento, e não estímulos vindos de fora.

Ainda pelo entendimento de Dehaene (2020), essa rede atencional funciona como um painel de comando, supervisionando os “programas” mentais executados a cada momento, monitorando erros e adaptando estratégias. O exemplo que o autor apresenta é o da realização de qualquer operação matemática feita de cabeça, evidenciando que, nesse caso, o que existe é um processo serial, muito mais lento, e que precisa ser governado com precisão, passo a passo, guardando resultados intermediários para aplicá-los aos cálculos subsequentes.

O controle executivo, como informa Dehaene (2020), às vezes também é chamado de “executivo central”, o que sugere a ligação com a memória de trabalho (ver subseção 2.2.4), principalmente por necessitar do armazenamento de uma série de dados bastante efêmeros para atuar satisfatoriamente. De igual forma, Dehaene (2020) o coloca como quem faz o controle das entradas e saídas do “espaço de trabalho neuronal global” (ver subseção 2.1.3), um sistema teórico que age como roteador das informações mentais e que está no centro daquilo que experimentamos como consciência (Dehaene, 2014). Vale notar também que o controle executivo, tal como apresentado aqui, tem várias sobreposições com o que Tsuchiya e Koch (2009) entendem como as atribuições da consciência.

Em algum ponto do circuito mental, antes do acesso consciente, existe um gargalo que somente libera a passagem para uma linha de pensamento por vez, e essa limitação fisiológica acaba com qualquer pretensão de uma pessoa ser plenamente capaz de executar múltiplas tarefas. O que existe, conforme atesta Dehaene (2020), é uma ilusão que se justifica pela própria limitação da mente (o cérebro, por óbvio, não tem como se distanciar e perceber aquilo tudo que está escapando de sua própria percepção) e explicada de, pelo menos, três formas. Se existe uma tarefa consciente sendo executada como a principal, ou a tarefa secundária é executada satisfatoriamente por já estar automatizada, portanto, operada de forma inconsciente; ou ela sofre redução no desempenho, especialmente na sua velocidade, por gerar uma fila para o limitado processamento consciente; ou ela é absolutamente ignorada. Dehaene (2020) chega a admitir a possibilidade, ainda em debate entre psicólogos, de casos pouco comuns em que a atenção é comutada rapidamente entre tarefas. Ainda assim, o comportamento típico testado nas multitarefas segue os três padrões mencionados.

Essas limitações explicam os fenômenos da “cegueira desatencional”, apresentado há pouco, e, talvez mais diretamente, o do “pisca atencional”, que acontece quando um segundo alvo, tão visível quanto o primeiro, é totalmente perdido pela mente (é ela quem “pisca”), evidenciando a limitação da percepção consciente no eixo do tempo (Dehaene, 2014, 2020). Buscando mais um exemplo da audição, estudos experimentais com pilotos de avião demonstraram que não é possível prestar atenção em duas conversas diferentes, apresentadas uma para cada ouvido, independentemente de quão bem treinados sejam os sujeitos (Broadbent, 1958; Dehaene, 2014).

O treinamento, contudo, é uma parte fundamental para que seja possível lidar adequadamente com situações em que a disputa por atenção é razoavelmente gerenciável. Identificado por Petersen e Posner (2012) como uma das extensões do quadro teórico original das três redes principais de atenção (alerta, orientação e controle executivo), a função de *autorregulação* é a chave para entender como a mente pode controlar a si mesma. Tratada na literatura científica a partir de vários sinônimos, como “atenção executiva”, “concentração” e “autocontrole”, a autorregulação consiste na habilidade de controlar voluntariamente as emoções e a cognição, ou, ainda em outras palavras, a capacidade de selecionar e amplificar respostas menos dominantes sobre as mais dominantes e outros reflexos (Petersen; Posner, 2012; Dehaene, 2020). Uma demonstração clássica que evidencia o autocontrole em ação é o “efeito Stroop”, que representa um aumento no tempo de resposta e/ou na quantidade de erros quando um sujeito precisa informar a cor objetiva em que uma palavra está escrita, sendo que essa palavra é o nome de uma cor diferente, por exemplo, “vermelho” impresso em tinta verde (Petersen; Posner, 2012; Dehaene, 2020). Do ponto de vista fisiológico, segundo Dehaene (2020), as áreas neurais responsáveis pela autorregulação levam de quinze a vinte anos para estarem completamente maduras; no entanto, é um circuito plástico que pode e deve ser exercitado desde cedo. Um autocontrole mais eficiente, entre um grande número de outros benefícios, reduz distrações; e o inverso também se aplica: a redução de distrações ajuda a melhor formar a autorregulação (Dehaene, 2020).

É oportuno agrupar e definir mais claramente alguns conceitos com nomenclaturas parecidas e que representam funções aproximadas. Fazendo uma metáfora simples, há um “volante” que determina a direção dos pensamentos e ações de cada indivíduo. Esse volante costuma receber o nome geral de *funções*

executivas na Psicologia, também aparecendo recorrentemente sob a identificação *controle cognitivo*, ou mesmo, mais popularmente, “força de vontade” (Petersen; Posner, 2012). Como há casos em que o volante pode ser girado de forma involuntária, por processos inconscientes (van Gaal *et al.*, 2010; Dehaene, 2014), é importante ter um conceito que isole aquilo que é puramente controlado pela vontade. Para isso, Dehaene (2014, p. 85) aplica o termo “executivo central” (o mesmo que dizer *controle executivo*; terceiro circuito da atenção, discutido acima), que é exclusivamente consciente. A ideia que parece mais difícil de situar nesse arranjo é a da *atenção executiva*, aprofundada no parágrafo anterior. É evidente que ela realiza funções que, no mínimo, sobrepõem-se às do controle executivo. A diferença que se pode encontrar é que a atenção executiva, presumivelmente, é capaz de operar no nível metarreflexivo: em outras palavras, pensando sobre o próprio pensamento. Assim, a noção de “autocontrole” se aplica plenamente a ela, contrastando com o simples “controle”; voltando à metáfora do volante, tem a ver com, além de girá-lo, ponderar por que foi girado para aquela direção e daquela maneira.

2.1.2 Processamento inconsciente

Considerando que a consciência responde por uma parcela muito pequena, ainda que de altíssima relevância, de operações mentais humanas – e, de forma muito mais rudimentar, até mesmo de um seleto grupo de espécies animais, como macacos e golfinhos (Dehaene, 2014), algo debatível e que precisa de maior comprovação –, todo o universo restante de processos desenvolvidos pelo sistema nervoso é feito de forma inconsciente.

Pelo quadro teórico de Dehaene (2014), as representações inconscientes são divididas em cinco categorias. A primeira é a das informações pré-conscientes, aquelas que apresentam alta força e estão corretamente codificadas, mas que não recebem amplificação pelos sistemas de atenção; se não fosse por isso, se converteriam em conscientes. A segunda é composta pelos estímulos subliminares, ou seja, aqueles sinais fracos demais para avançar até os centros mais elevados do cérebro, não importa o quanto se tente prestar atenção. O terceiro grupo é formado pelos padrões desconectados – do qual a ação de respiração é um bom exemplo –, que são fortes e duradouros, mas simplesmente estão restritos a circuitos que não

se comunicam com a rede mental maior, que sustenta a consciência. A quarta categoria é a da informação diluída, na qual o padrão de disparo dos neurônios é complexo demais para ser tratado pela mente consciente. Dehaene (2014) a exemplifica com tipos muito específicos de estímulos visuais, que ou oscilam muito rápido ou possuem detalhes muito refinados, e que, então, avançam com energia pelos caminhos cerebrais, mas são, de alguma maneira, incompreensíveis. Por fim, o quinto conjunto é o das conexões latentes, inacessíveis à mente consciente por não estarem codificadas no formato esperado, muito embora representem um vasto conhecimento composto por memórias dormentes e intuições inconscientes, todas guardadas pela forma como o cérebro de cada indivíduo foi criando seu “cabeamento” particular de sinapses. A propósito dessa categoria, Dehaene (2014, p. 197) traz o oportuno exemplo das regras gramaticais pelas quais processamos a fala, as quais se mantêm largamente despercebidas. A aritmética, em comparação, é altamente consciente, controlada de forma direta e conduzida passo a passo, de maneira explícita.

Para entender o funcionamento, a importância e a localização do processamento inconsciente no grande esquema da mente, é extremamente conveniente a metáfora que Dehaene (2014) cria para sua explicação, comparando a macroestrutura neural com um grande órgão público (no texto do autor, o *FBI*, *Federal Bureau of Investigation*, a polícia federal dos EUA). Nesse cenário, o inconsciente representa a grande massa de funcionários dividida em setores nas partes baixas da hierarquia. Eles dão seus pareceres individuais até que se chegue a um consenso, que é repassado a um superior imediato (tudo ainda no nível inconsciente). Em um arranjo piramidal, os chefes de setor passam pelo mesmo processo, algo que se repete em vários níveis, avançando por conselhos cada vez mais altos e restritos. No topo, o porta-voz (a consciência) recebe um *briefing* completo e, enfim, pode informar o presidente (o controle executivo), que tomará a decisão ou ação suprema. Sendo assim, os funcionários mais baixos são os neurônios das áreas sensoriais primárias; as reuniões e deliberações setoriais são as chamadas “assembleias neuronais”; e é por meio de análises estatísticas feitas entre as células nervosas que todo o trabalho desenvolvido no inconsciente se processa (Dehaene, 2014).

As operações inconscientes, como se pretendeu mostrar até aqui, compõem, de longe, o maior volume das funções mentais. Convém, neste momento, destacar

alguns processos dessa categoria que são de especial importância para esta dissertação. Um deles é a atenção, que foi discutida em detalhes na parte inicial deste capítulo. Como bem argumenta Dehaene (2014), os sistemas atencionais devem operar, na maior parte dos casos, de maneira inconsciente: apenas por meio de monitoramento constante é que o cérebro pode se proteger de eventuais riscos ou buscar oportunidades melhores para sua meta naquele momento. Se não fosse o caso, seria absolutamente impossível perceber e reagir a um estrondo alto enquanto se está concentrado em uma leitura, já que o limitado espaço da consciência estaria lotado. A mesma lógica, Dehaene (2014) aplica à vinculação de experiências multissensoriais em uma representação única e coerente. Como demonstrado a partir dos estudos do “efeito McGurk” – pelo qual um som de uma sílaba e a leitura labial de outra sílaba diferente se fundem, gerando, ainda, um terceiro som distinto, mas apenas quando audição e visão são acionadas no mesmo momento –, os dados já chegam à consciência “empacotados” e reconstruídos, não como estímulos crus totalmente soltos.

Apesar dessas surpreendentes habilidades do inconsciente, certamente há limitações naquilo que é capaz de realizar. Após apresentar uma síntese de vários experimentos que buscaram entender em que nível mental o processamento semântico de uma palavra ocorre, Dehaene (2014) defende mais uma conclusão intrigante: o significado pode ser capturado mesmo sem a pessoa estar consciente de que leu uma palavra. Mas isso quer dizer que as mensagens subliminares fazem exatamente o que as teorias da conspiração alegaram por tanto tempo? Longe disso: essa percepção inconsciente é muito fraca e não escapa de alguns circuitos mais avançados, porém locais e distantes da consciência; portanto, há pouquíssimo poder de influência sobre o comportamento (Dehaene, 2014). Captar a ideia geral de uma cena visual exibida por um tempo extraordinariamente curto (na casa de dezenas de milissegundos) também é possível, como prova a pesquisa de Fei-Fei *et al.* (2007). No entanto, é bastante revelador o quanto são mais integradas e corretas as percepções derivadas das fotografias apresentadas por meio segundo: imagens que ficaram em tela por apenas 27 ms geraram relatos muito rudimentares e sustentados por aspectos visuais básicos, como luminosidade e formas.

Isso posto, seria um erro subestimar o inconsciente naquilo em que ele é realmente bom: velocidade, processamento paralelo e análise estatística. Um pianista consegue executar perfeitamente uma passagem tecnicamente difícil de

uma música porque não precisa se concentrar no movimento de cada dedo individualmente (Dehaene, 2014). Pensar de forma detalhada e metódica sobre movimentos em esportes como golfe e futebol é, na mesma linha, algo que reduz o desempenho de esportistas experientes (Beilock *et al.*, 2002). Curiosamente, a mesma condição não piora o desempenho de iniciantes, tampouco de praticantes experientes usando seu lado (braço ou perna) menos treinado. Ter conhecimentos automatizados, para além de favorecer a excelência de uma ação bem aprendida, também é a única forma razoável de viabilizar multitarefas suficientemente bem executadas, como discutido na subseção 2.1.1. Por último, é interessante relatar um experimento que concluiu que um grupo de pessoas foi mais eficaz escolhendo o melhor carro em uma lista justamente quando foi distraído dessa tarefa de escolha (Tsuchiya; Koch, 2009; Dehaene, 2014). Os sujeitos tiveram um primeiro acesso a um catálogo de automóveis e suas correspondentes especificações, porém deveriam definir o veículo superior usando apenas sua memória. O resultado indica que decisões complexas mais proveitosas podem ser tomadas quando se deixa o processo de deliberação a rodar apenas no inconsciente.

2.1.3 Pensamento consciente

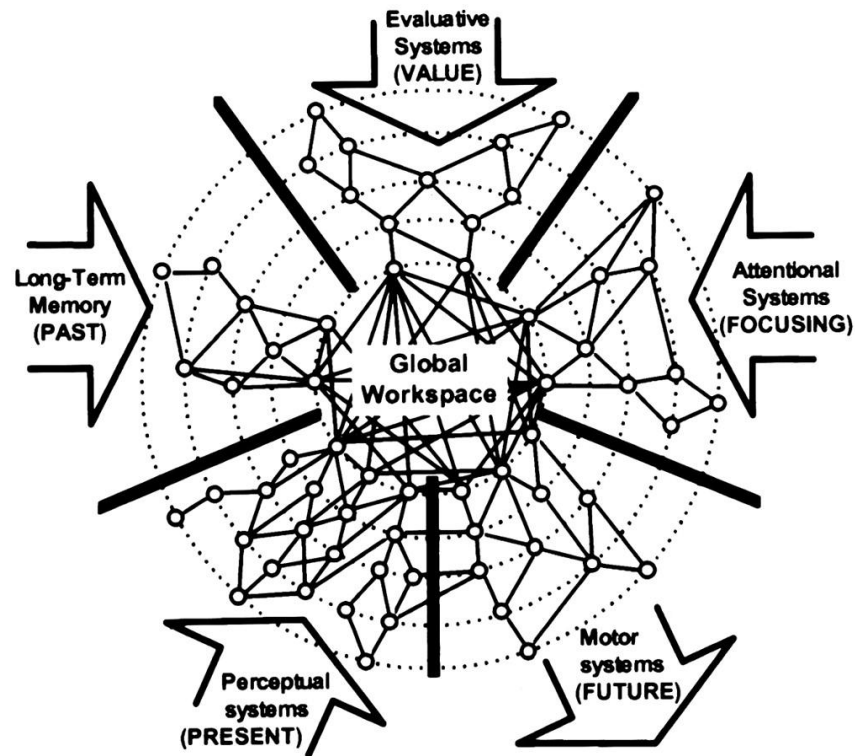
Neste momento, é importante retomar e sintetizar o que foi discutido sobre a consciência ao longo deste capítulo e delimitar qual o entendimento que se utilizará neste trabalho. Considerá-la a partir do conceito de acesso consciente, tal qual especificado por Dehaene (2014), é o que parece tornar a ideia de consciência o mais distinta possível do restante dos fluxos mentais que poderiam deixar as fronteiras mais difusas. Assim, o pensamento consciente será tomado exclusivamente como aquele que a introspecção pode “segurar”, sobre o qual o controle voluntário e a manipulação podem ser exercidos, e que pode ser externado de forma explícita. Desse jeito, excluem-se funções listadas por Tsuchiya e Koch (2009) que, em última análise, quase certamente são ancoradas em algum mecanismo inconsciente, como a sumarização de todas as condições presentes internas (do organismo) e externas (do ambiente). As visões de Tsuchiya e Koch (2009) e Dehaene (2014) não diferem muito, senão na especificidade: este último considera a consciência só a “ponta do iceberg”, o absoluto produto final, gerado

apenas ao fim de uma longa cadeia de processos inconscientes que ficam, incrementalmente, mais elaborados.

Admitir que acesso consciente é o mesmo que a consciência permite assumir os demais pressupostos da teoria da estação de trabalho neuronal global de Dehaene (2014). Por esse quadro conceitual, a “[...] consciência é apenas compartilhamento de informação por todo o cérebro”⁶ (Dehaene, 2014, p. 165). As dúzias de módulos especializados do sistema nervoso são interligados por um sistema de comunicação – a *área de trabalho global* – que é capaz de preservar a informação selecionada, provinda de algum ponto da rede, e repassar de forma incrivelmente flexível a quaisquer outros módulos que estão igualmente interconectados. Dehaene, Kerszberg e Changeux (1998) identificam cinco categorias principais de processadores que se projetam até e a partir da área de trabalho global (Figura 1): circuitos de percepção, que acessam o estado presente do mundo externo (incluindo compreensão de linguagem, por conter a área de Wernicke, por exemplo); circuitos de memória de longo prazo (ver subseção 2.2.3), que permitem recuperar eventos e percepções do passado; circuitos de avaliação, ocupados com variados julgamentos de valor; circuitos de atenção (ver subseção 2.1.1) capazes de mobilizar os recursos para focar representações de módulos neurais específicos, por meio de aplicações e atenuações de sinais; e circuitos de programação motora, que guiam o comportamento futuro intencional (abrangem a área de Broca e, portanto, têm relação com a produção de fala).

⁶ Do original: [...] *consciousness is just brain-wide information sharing* (Dehaene, 2014, p. 165).

Figura 1 – Área de trabalho global



Fonte: Dehaene, Kerszberg e Changeux (1998, p. 14530).

A teorização de Dehaene (2014) sobre a estação de trabalho neuronal global possui não só grande consistência no ponto de vista conceitual, como também se sustenta em muita pesquisa empírica. A partir de medições que tentavam corroborar as hipóteses idealizadas, Dehaene (2014) identificou quatro assinaturas que o pensamento consciente apresenta. A primeira é uma ignição intensa de muitas áreas cerebrais, para além dos circuitos locais imediatos, causada a partir de uma grande amplificação de sinais, que culminam invadindo os lobos parietal e pré-frontal do encéfalo. A segunda é uma onda, chamada "P3", que emerge de forma relativamente lenta (cerca de 300 milissegundos) após o estímulo primário e que pode ser medida por instrumentos mais simples, como o eletroencefalograma (EEG). Em terceiro lugar, a marca da consciência é um transbordamento longo de oscilações em alta frequência no cérebro, vibrações que ocorrem mesmo sob estímulos subliminares, mas que só se sustentam com força quando há o acesso consciente da informação. A quarta e última assinatura trata-se da sincronização de muitas áreas distantes do cérebro, estado atingido pela comunicação bidirecional entre essas regiões.

Por causa da menor complexidade para medição e da boa confiabilidade da onda P3, que representa a segunda assinatura da consciência, Dehaene (2014) se concentrou sobre ela para constatar acesso consciente em sujeitos que não podem reportar seu estado interno, tais quais bebês, pacientes em coma e até macacos e camundongos. Um instrumento que o pesquisador emprega com frequência, o “teste local-global”, permite que se entenda de forma simples a diferença entre um processamento inconsciente e um pensamento consciente. O experimento, que pode ser elaborado com sons simples, busca mensurar uma resposta diferente que só ocorre quando o sujeito percebe a quebra no padrão de uma sequência mais ampla, ou seja, global. No caso de uma violação de regras locais, como quando um quinto som destoa dos quatro idênticos que o precederam, não importa quantas vezes seja ouvida essa sequência de 4 + 1, o final dela sempre trará uma surpresa moderada, passível de ser medida em cérebros muito rudimentares ou debilitados, indicando uma resposta automática do inconsciente. Porém, o consciente, que integra e analisa as informações de maneira bastante mais elaborada, rapidamente percebe que a sequência 4 + 1 é o “novo normal”, de modo que para de responder emitindo a onda P3 ao final de cada frase sonora. Mais do que isso: essa onda P3 volta a aparecer no cérebro consciente quando a sequência é alterada para cinco sons idênticos, pois, curiosamente, essa sequência monótona passou a ser uma novidade.

Na subseção anterior, discutimos as desvantagens da consciência, por operar de maneira lenta e serial, trabalhando com uma informação de cada vez. E é claro que ela tem muitas limitações, mas elas só nos surpreendem porque estamos habituados, por intuição e por instrução, a pensar frequentemente nos benefícios dos pensamentos conscientes. Como relembra Dehaene (2014), existem operações mentais que só podem ser conduzidas conscientemente, a exemplo de raciocínios estratégicos e cálculos matemáticos com múltiplas etapas. O autor argumenta que a informação consciente é estável, pois consegue estender a vida útil das percepções e representações, e também comprimida eficientemente, já que um símbolo mental empacota muitos atributos relevantes. Esses símbolos – a “linguagem do pensamento”, como Dehaene (2014) os chama – transitam livremente e são entendidos nos mais diferentes módulos mentais. A arquitetura dos dados conscientes, em última análise, é o que permitiu e favoreceu a criação das linguagens humanas, especialmente a verbal, que nos capacitaram a comunicar

pensamentos também para fora do nosso mundo interior, criando laços e cooperação, conquistando feitos impossíveis individualmente (Dehaene, 2014).

O pensamento consciente, como será aprofundado na seção 2.4, é a parcela do processamento cognitivo indispensável para que exista a compreensão linguística, e, em última instância, é justamente ele que este estudo pretendeu investigar se recebe algum tipo de reforço ou se é sobrecarregado e mal aproveitado pelo uso de múltiplas modalidades na apresentação de um único texto. O processamento inconsciente, nesse contexto, representa a vasta gama de estímulos simultâneos que a cognição do leitor precisa filtrar, empregando suas estratégias e seus sistemas de atenção, para que a compreensão do material linguístico se dê de forma eficaz.

2.2 Memórias

Dando sequência aos temas fundamentais para esta dissertação, avançamos em direção às capacidades que são eminentemente humanas. Se a atenção pode ser algo extremamente basilar, tendo relação com necessidades elementares de sobrevivência e autopreservação dos organismos, e se a consciência pode, no limite, ser analisada como uma máquina que processa uma parcela minúscula da realidade – o tempo presente –, o que a memória adiciona a esses sistemas altamente complexos é uma maneira de guardar acervos de experiências passadas que podem ser recuperadas quando for útil.

A exemplo do pensamento consciente, há animais – todos eles, segundo Lent (2010, p. 651) – que evidenciam possuir algum sistema de memória; porém, mais uma vez, é algo muito mais limitado e rudimentar em comparação ao ser humano, bem como é pouco acessível à ciência, já que não há como contar com relatos verbais que pudessem dar melhores pistas dessas outras subjetividades e introspecções. Indícios robustos de que há espécies que não vivem exclusivamente confinadas no “agora” são verificados a partir de comportamentos bastante simples: ratos que encontram seu caminho em um labirinto, cães que salivam ao ver uma luz que costuma sinalizar a chegada da comida, ou gatos que aprendem a acionar alavancas para sair de uma gaiola (Lent, 2010). Lent (2010) afirma que a aprendizagem dos animais ocorre por mecanismos muito primários. O tipo menos elaborado é o das memórias não associativas, isto é, aquelas em que não há a

mediação ou associação com um segundo estímulo arbitrário. Então, nesse grupo, há os casos em que a repetição de um evento acaba parando de provocar um reflexo esperado (habituação) ou, de outro lado, passa a provocar uma resposta instintiva que não existia antes (sensibilização; por exemplo: sons de passos começam a ser desproporcionalmente tratados tal como se um invasor já tivesse sido avistado na prática). As memórias associativas, por sua vez, envolvem maior complexidade: no condicionamento clássico, algo inócuo como uma campainha que costuma soar antes do almoço já passa a abrir o apetite, mesmo que a comida não venha; e já no condicionamento operante ou instrumental, o animal aprende a preferir realizar ações relacionadas a recompensas e evitar comportamentos que geram punições.

Como se pode perceber por essa rápida listagem dos tipos de memória que os animais apresentam, são muitas as possibilidades de recortes teóricos nesse conceito muito abrangente. Assim, tomando por base a linha de raciocínio de Izquierdo (2018), intitulamos esta seção como “memórias”, no plural, para sinalizar a vasta quantidade de arquivos, tipos e mecanismos que há, ainda mais para conseguir dar conta do universo conhecido da muito mais sofisticada mente humana.

Quais as vantagens que ter sistemas de memória tão desenvolvidos nos dá? Para deduzir a resposta, basta que lembremos do que foi discutido na seção anterior: o mecanismo do pensamento consciente proporcionou uma revolução para a nossa espécie, começando por uma organização e otimização sem precedentes dos dados que transitam internamente e culminando numa complexa e avançada sociedade, coordenada pela comunicação dos pensamentos estruturados que são trocados entre pessoas. A memória, em um sentido amplo, eleva o poder da consciência para outra escala, pois permite armazenar conhecimento acumulado (também na forma de um sem número de processos inconscientes), o que evita que toda e qualquer ação tenha que surgir do zero e que habilite para o estabelecimento de associações significativamente mais rebuscadas. É claro que se trata de uma habilidade que nos libera do presente em direção ao passado, mas, talvez algo não intuitivo, é que ela também nos permite pensar o futuro. Nesse sentido, Oliveira (2007) recorda uma das concepções fundamentais da psicologia cognitiva soviética, que é a de que o comportamento humano não depende só das aprendizagens ocorridas, mas também da história futura que se pode projetar a partir delas. É a

partir desse referencial que tomam forma os planos e as intenções. Ribeiro (2019), quando reflete sobre os sonhos – que são construídos a partir das memórias que reverberam pelo inconsciente, mesmo em um estado de vigília, quando se “sonha acordado” – defende que eles podem apontar para o nosso futuro, já que as inúmeras combinações que se podem estabelecer dentro desse acervo de informações, para além de mostrar o que fomos, são capazes de mostrar quem podemos vir a ser.

2.2.1 Classificação das memórias pelo seu conteúdo

Mais adiante nesta seção, investigaremos com maior detalhamento as memórias classificadas a partir da sua duração temporal, já que esse é um ponto de especial importância para o trabalho a ser desenvolvido nesta dissertação. Assim sendo, traremos, a seguir, recortes teóricos da memória que consideram suas diferenças quanto ao seu conteúdo. Não é demais lembrar que todo arquivo de memória é, necessariamente, classificado conjuntamente em termos de tempo e conteúdo; então, podemos ter memórias episódicas de curto prazo, semânticas de longo prazo, e uma série de outras combinações.

Uma das categorias essenciais é a das episódicas. Como o nome sugere, esse tipo de memória guarda episódios específicos ocorridos, ou, nas palavras de Lent (2010, p. 649), “[...] envolve eventos datados, isto é, relacionados ao tempo”. Ainda por esse autor, essa categoria lida com fatos sequenciados e, por ser tão detalhada, é também bastante particular a cada indivíduo. A leitura de Baddeley (2011a) e Izquierdo (2018) é consideravelmente similar à apresentada, com a exceção de que este último considera explicitamente que memórias episódicas e memórias autobiográficas são a mesma coisa. Sobre isso, Baddeley (2011a) faz um comentário bastante contundente, argumentando que a biografia de cada um contém informações como o próprio nome e lugar onde moramos, que, então, não são episódios propriamente ditos.

Isso nos conduz ao segundo tipo de memórias de acordo com seu conteúdo: as semânticas. São informações que assumem uma representação mais abstrata, suficientemente desligadas de etiquetas de tempo e espaço. Lent (2010) as considera como uma memória cultural, já que sua característica de atemporalidade e generalização permite que possam ser compartilhadas por muitas pessoas.

Baddeley (2011a) segue a mesma linha e faz um destaque importante: a memória semântica não se restringe a conhecimento de palavras, ela também avança para conhecimentos multimodais (ver seção 2.3), de aspectos sensoriais variados, e envolve construções mentais realmente amplas, como o funcionamento da sociedade ou como se comportar em determinado local ou situação. Izquierdo (2018, p. 41) comenta que “Quase sempre, as memórias semânticas [...] são adquiridas por meio de memórias episódicas” e – apesar de ser uma boa discussão se o “quase” usado no início da citação realmente se aplica, porque, salvo melhor leitura, toda aprendizagem poderia ser rastreada até o episódio objetivo em que primeiro ocorreu –, essa noção ajuda a compreender que o conhecimento se torna geral assim que vai perdendo seus detalhes muito específicos. Aliás, é interessante pensar no processamento que ocorre na informação: ela perde alguns metadados (data e local) e ganha uma série de outras associações mentais, como outras representações sensoriais ou conceituais.

Essas duas classificações explicadas até aqui, semânticas e episódicas, compõem o grupo maior das memórias explícitas ou declarativas, isto é, aquelas que podem ser pensadas e comunicadas utilizando-se palavras ou outras representações simbólicas (Lent, 2010; Baddeley, 2011a; Izquierdo, 2018). Esse enquadramento, cruzado com a exigência de reportabilidade potencial para que a consciência seja constatada (Dehaene, 2014), permite que se entenda que só as memórias episódicas e semânticas são plenamente conscientes. Há que se considerar casos não tão evidentes, como quando uma pessoa é capaz de explicar um aprendizado por habituação (por exemplo, “não perco a concentração porque já me acostumei com os sons de aviões passando”), mas é possível se admitir que, nessa situação, essa memória mais primária possa ter sido elaborada e convertida para uma semântica. E, em todo caso, uma diferença importante é que o indivíduo não conseguiria editar essa memória com facilidade e usando apenas o pensamento consciente, considerando que ainda se trata de uma resposta eminentemente automática.

No outro lado do espectro, com relação às declarativas, está o grupo das memórias implícitas. Neste, pela divisão concebida por Lent (2010), está a recém mencionada habituação e todos os outros mecanismos de aprendizagem dos animais que foram listados no início desta seção (e que também fazem parte do acervo dos humanos, não custa reforçar). As memórias implícitas, também

chamadas de não declarativas, de acordo com Lent (2010), têm a desvantagem de requerer mais tempo e treinamento para se formar, mas o lado positivo é que persistem de forma mais duradoura. Um subtipo muito importante das implícitas é a memória de procedimentos – tipicamente ilustrada com habilidades como andar de bicicleta, dirigir ou tocar um instrumento musical –, que engloba também hábitos e regras, todos conhecimentos difíceis de traduzir em palavras claras e que sejam suficientes (Lent, 2010; Izquierdo, 2018).

As memórias implícitas, especialmente as de procedimentos, pelo que se pode entender, operacionalizam as mesmas tarefas automatizadas que Dehaene (2020) aceita que poderiam ser executadas juntamente e sem prejuízos (idealmente) com uma atividade consciente. Recuperando as categorias pelas quais Dehaene (2014) divide o processamento inconsciente, os chamados padrões desconectados representariam um cenário extremo, pois ações como a respiração estão tão protegidas nos circuitos do cérebro que não podem ser interrompidas por quaisquer ações simultâneas. Porém, o grupo das conexões latentes parece ser o que comporta os casos mais típicos, que são os dos aprendizados motores e outras habilidades. A leitura também se enquadraria aí, mas apenas a considerando na sua parcela de decodificação grafema-fonema que, como defendem Gabriel, Morais e Kolinsky (2016), ao ser automatizada, libera a atenção e o processamento consciente para operações mais complexas de compreensão de texto.

Para dar uma ideia do quanto a aprendizagem que gera memórias implícitas é mais poderosa do que imaginamos, é válido refletirmos sobre a aquisição da língua materna. Como afirma Dehaene (2014) e foi tratado na seção anterior, a fala é produzida a todo momento, em formulações corretas e aceitas pelos interlocutores, sem que se tenha que consultar o conjunto de regras gramaticais de forma consciente. Isso seria impraticavelmente lento, idêntico ao que faz um sujeito tentando conversar em um idioma pouco dominado por ele, precisando conferir cada palavra, flexão e frase. De que maneira, então, é possível guardar um número infinito de possibilidades de uma língua? Será que o inconsciente guarda cópias de todas as construções frasais que algum dia ouvimos ou lemos? Não é o que parece. Pela interessante explicação de Koch e Cunha-Lima (2011), simulações computacionais usando redes neurais comprovaram que é plausível que o sistema cognitivo humano extraia de dados considerados “pobres” (por exemplo, poucas frases ditas a uma criança) muito mais informações do que se imagina. Essa ideia,

de acordo com as autoras, enfraqueceriam a concepção inatista ou gerativista, que, conforme Ilari (2011), postulava que a rapidez e a perfeição com que a criança adquire sua língua materna só seria explicada pela existência de um equipamento biológico prévio nela, que seria a “gramática universal”.

2.2.2 Operações sobre as memórias

A aprendizagem, mencionada ainda há pouco e em outros pontos desta seção, é uma das operações conduzidas nos sistemas de memória. Juntamente com ela, também referida como aquisição, há a retenção, a consolidação, a evocação e o esquecimento das informações (Lent, 2010; Izquierdo, 2018). Nesta subseção, serão considerados tipos ideais de memória, que geram arquivos de longa duração, já que são mais previsíveis e evitam complicações desnecessárias para a explicação proposta agora.

Começando pelo primeiro processo mnemônico, a aquisição ou aprendizagem envolve o ingresso da memória e seu momento inicial de gravação. Lent (2010) a define como a entrada de qualquer evento memorável no sistema, o que se estende de estímulos externos até pensamentos. O autor destaca um aspecto fundamental, que é o da seleção:

[...] como os eventos são geralmente múltiplos e complexos, os sistemas de memória só permitem a aquisição de alguns aspectos mais relevantes para a cognição, mais marcantes para a emoção, mais focalizados pela nossa atenção, mais fortes sensorialmente, ou simplesmente priorizados por critérios desconhecidos (Lent, 2010, p. 647).

O que é importante sublinhar é que a realidade, além de recortada pelos sistemas de atenção (discutidos na seção anterior), chega ainda mais comprimida até a memória, tendo muito mais a ver com uma representação elaborada do que uma gravação fidedigna. Outro ponto relacionado à eficiência é que as informações que entram a cada momento são comparadas com o acervo de memórias que existem, descartando-se o que é redundante (Izquierdo, 2018).

A propósito da aprendizagem em um sentido mais prático, como o da sala de aula, Dehaene (2020) reforça a importância da atenção, afirmando que ela é um ingrediente vital para que o aluno aprenda com sucesso, principalmente porque

somente pode ser aprendido aquilo que é percebido. Talvez em uma interpretação mais aberta, provida até mesmo por Dehaene (2014), pode-se pensar cenários em que aprendizados, tais como os implícitos e de procedimentos, podem ser agregados pelo processamento inconsciente em pessoas não tão focadas. A própria repetição mecânica de alguma informação – possivelmente feita com atenção, mas duvidosamente tratada conscientemente durante toda a seção de estudo – é tida por Baddeley (2011d) como uma estratégia útil para alguns aprendizados específicos (como pronúncia de idiomas); no entanto, não é o melhor método de forma geral: o autor considera muito mais produtivo que o aluno faça elaborações no seu acervo mental entre materiais novos e já conhecidos.

A segunda fase pela qual passa a memória é a retenção, uma espécie de “purgatório”. De acordo com Lent (2010), é um armazenamento temporário e sujeito a limitações de espaço, bem como é muito suscetível à distorção pela presença de elementos distratores. Em concordância, Izquierdo (2018) informa que as memórias nessa fase são muito lábeis, de modo que podem ter drásticas mudanças pela ocorrência de traumas físicos, abuso de álcool ou outras substâncias, estresse e até mesmo interferência com outras memórias existentes ou que competem no mesmo momento. Distorções ocorridas nessa etapa são as principais produtoras do fenômeno das falsas memórias, que são informações incorretas que se guardam e se recuperam com grande convicção na sua veracidade (Izquierdo, 2018). A janela de tempo pela qual as memórias ficam retidas dessa forma é, segundo Izquierdo (2018, p. 63), a partir de estudos dos mecanismos fisiológicos e moleculares envolvidos, de entre três a seis horas.

Tendo sobrevivido durante a retenção, com ou sem alterações, a memória chega até o estágio da consolidação, ou seja, quando um evento é retido de forma prolongada, podendo chegar a ser permanente (Lent, 2010), adquirindo a forma mais estável que terá, enquanto arquivo de longo prazo (Izquierdo, 2018). Baddeley (2011d) propõe que, apesar de ser discutível que se possa aprender dormindo, há evidências de que o sono ajuda na consolidação do que já foi aprendido e retido. Isso é algo que faz bastante sentido, considerando que uma pessoa que dorme estará muito menos exposta a interferências.

Chega-se ao momento fundamental, que basicamente resume a utilidade da memória: a evocação. Também chamada de lembrança, como diz Lent (2010), é o acesso feito pela mente à informação guardada. É relevante destacar que essa

operação pode ser feita tanto com as memórias consolidadas quanto com as retidas. E, ainda que a lembrança seja recuperada, na maioria das vezes, de maneira voluntária e completa, existe também o fenômeno do *priming*, que é quando um fragmento mínimo de uma memória não prontamente disponível, ao ser apresentado externamente como “dica”, é suficiente para que se retome a informação completa, como lembrar a letra de uma música só depois de ouvir o primeiro verso (Izquierdo, 2018).

O processo final é o do esquecimento. Apesar de ser popularmente tratado como algo fundamentalmente negativo, tal como é o caso de situações patológicas de amnésia e demência, a verdade é que é uma propriedade normal da memória e, mais do que isso, é o que torna possível conduzir uma vida funcional, filtrando o que importa e evitando a sobrecarga do sistema (Lent, 2010; Izquierdo, 2018). Izquierdo (2018) explica que o esquecimento é o apagamento completo de informações, acontecido por morte neuronal ou desmantelamento das sinapses envolvidas. O autor, na mesma linha do que afirma Lent (2010), coloca que as memórias são conservadas se o organismo as considera relevantes; portanto, a evocação continuada de uma memória é, provavelmente, o que mais reduz as chances de ser perdida.

Izquierdo (2018) separa outros dois fenômenos daquilo que é considerado esquecimento, já que estes não conduzem as memórias ao apagamento irreversível. O primeiro deles é a habituação (ou extinção), comentada no início desta seção, um processo que inibe ou suprime respostas instintivas ou aprendidas. Segundo o autor, essas memórias seguem latentes e podem ser reativadas com uma intensificação do estímulo ao qual o sujeito se habituou. Além do mais, a habituação nada mais é do que uma memória adicional, aprendida para modificar um reflexo primário. O segundo fenômeno elencado por Izquierdo (2018) é a repressão, isto é, quando o organismo, voluntária ou involuntariamente, decide ignorar ou suprimir a evocação de lembranças por longos períodos, geralmente devido a elas representarem eventos desagradáveis ou prejudiciais. A exemplo do que ocorre com a habituação, as memórias inibidas seguem intactas aqui, dependendo de um gatilho certo para regressarem.

Todas as operações realizadas sobre as memórias, como destacam Lent (2010) e Izquierdo (2018), sofrem modulações pelo estado presente do organismo, considerando especialmente as variações hormonais e de neurotransmissores,

expressas na forma de emoções e outros comportamentos. A repressão de lembranças, recém comentada, é um exemplo de prejuízo ao desempenho da evocação. Izquierdo (2018) comenta que o estresse, a depender da dose, pode tanto melhorar a consolidação (lembranças vívidas de um acidente, por exemplo) quanto, no extremo, provocar amnésia. No lado das modulações positivas, um dos casos mais referidos é, certamente, o da maior facilidade que um aluno tem para aprender quando está com um estado de ânimo elevado (Izquierdo, 2018), de modo que a motivação é um elemento importante para os estudantes (Baddeley, 2011d; Dehaene, 2020).

2.2.3 Memórias de longo e curto prazo

Considerando as memórias pelo seu tempo de duração, iniciamos pelo seu formato mais estereotípico, já exposto muitas vezes ao longo deste texto. As memórias de longo prazo, ou longa duração, são os engramas – “‘unidade’ teórica da memória” (Lent, 2010, p. 645) – que são conservados por um período de horas até anos (Lent, 2010; Baddeley, 2011a; Izquierdo, 2018). Retomando a esquematização que Dehaene, Kerszberg e Changeux (1998) desenharam para a estação de trabalho neuronal global, esse repositório de informações é um dos cinco grupos de circuitos principais que enviam e recebem dados de e para a consciência. Assim sendo, representam a parcela do pensamento mais voltada à introspecção e também ao tempo passado. Esse acervo das memórias de longo prazo é composto por muitas lembranças e conhecimentos de todos os tipos, tais como os algoritmos mentais que usamos para soluções de problemas e também padrões de comportamento ou de respostas complexas que apresentamos (Oliveira, 2007).

Deixando a abordagem puramente conceitual das memórias de longa duração, é relevante considerar qual a forma concreta que assumem esses arquivos. Lent (2010) narra que houve uma linha teórica inicial que considerava a existência de um centro especializado em guardar os engramas no cérebro, e Izquierdo (2018) afirma que uma concepção clássica entendia que uma estrutura específica, o hipocampo, era o solitário responsável por fazer as gravações iniciais e depois encaminhá-las ao arquivo permanente. A respeito da localização centralizada das memórias, Lent (2010) apresenta várias e robustas evidências de que é um equívoco: as memórias estão por todo o cérebro. Como aprofunda Izquierdo (2018), elas são guardadas por

modificações duradouras que ocorrem na função e na forma das sinapses das redes neurais envolvidas nos eventos específicos; memórias semânticas de muita complexidade, como o conhecimento de toda uma língua, envolvem vários bilhões de conexões espalhadas por múltiplas áreas neurais. Refletindo o mesmo entendimento, Dehaene (2014) afirma que a memória consciente é uma reconstrução de um momento que também foi consciente no passado, que, ao ser evocado, é reencenado com o mesmo padrão de disparos neuronais, de maneira bastante aproximada. Sobre a importância do hipocampo na aquisição e evocação das memórias, Izquierdo (2018) aponta que a área ainda é compreendida como tendo protagonismo nesses processos, mas não atua isolada.

Uma vez que avançamos para examinar os entendimentos sobre as memórias com menor durabilidade, há menos consenso. A memória de trabalho, talvez o conceito mais controverso, será discutida na subseção seguinte. A seguir, exploramos o que podemos identificar como memória de curta duração.

Chama a atenção o quão cedo no desenvolvimento da Psicologia moderna surgem considerações a respeito de memórias de curto prazo. Em 1890, o filósofo e psicólogo norte-americano William James já falava sobre uma “memória primária” (Izquierdo, 2018), que se diferenciava da memória secundária (a “autêntica”), por alongar uma percepção consciente do presente, que, portanto, não é interrompida ou recuperada, necessariamente (James, 1931). Baddeley (2011b), por sua vez, narra outro registro do final do século XIX, sobre John Jacobs, um diretor de escola que desenvolveu um experimento simples para verificar a capacidade de os alunos recordarem sequências incrementalmente maiores de dígitos numéricos recém ouvidos, um teste que se tornou muito usado e que recebeu o nome de *digit span* (ou extensão de dígitos). Algumas décadas mais tarde, por volta de 1960, comenta Baddeley (2011a) que a memória era pensada a partir do “modelo modal”, que era alinhado com teorias do processamento de informação e previa um encadeamento serial de processos, tendo a memória de curto prazo antecedendo o destino final da informação na memória de longo prazo. Esse armazenamento de pouca duração aparece na literatura científica assumindo várias características, e, assim, é interessante adotar uma definição neutra proposta por Baddeley (2011a, p. 21), que o considera “retenção temporária de pequenas quantidades de material sobre breves períodos de tempo”. O autor também acrescenta que a memória de curto prazo é um sistema puramente passivo, que não realiza processamento de dados.

Diante do exposto, provavelmente o conceito mais concreto que há para a memória de curta duração vem da neurobiologia. Izquierdo (2018), embasado em uma série de pesquisas das quais participou no final dos anos 1990, tem total convicção de três princípios sobre as memórias de curta duração: elas são entidades próprias (não a mera ativação de partes das memórias de longo prazo), elas são muito distintas do que se entende por memória de trabalho e, por último, elas são processadas paralelamente (e não em cadeia sequencial) com os arquivos de longa duração. Um dos importantes estudos do pesquisador (Izquierdo *et al.*, 1998) descreve resultados com testes em ratos, nos quais a injeção de determinados fármacos em áreas cerebrais centrais para a manipulação de memórias, entre outros efeitos, foi capaz de prejudicar a lembrança de curto prazo sem causar qualquer efeito sobre a recordação posterior, vinda do longo prazo. De acordo com a explicação oferecida e o método rigoroso utilizado, Izquierdo e colaboradores (1998, p. 365) atestam que sua manipulação foi efetuada sobre o mecanismo de formação das memórias, excluindo a possibilidade de que representasse falhas somente na recuperação delas ou no desempenho. Dessa maneira, se a destruição da formação de memórias de curta duração não causa problemas à criação de arquivos de longo prazo, existe algum tipo de sistema temporário que atua de forma autônoma.

Ao assumir os posicionamentos de Izquierdo (2018), podemos considerar as memórias de curto prazo como aquelas que ficam vivas por até seis horas, servindo como uma cópia utilizável enquanto a versão definitiva ainda está sendo consolidada para durar indefinidamente. É possível que esses achados científicos tenham maior relevância para outras disciplinas e outros níveis de análise do problema, lidando com processos moleculares, porque é de se imaginar que, em situações típicas, o sujeito nem precisaria e nem seria capaz de saber se sua lembrança já é a permanente ou sua cópia. Ainda assim, Izquierdo (2018) faz com que se olhe com mais cuidado esse sistema intermediário: uma vez que a memória de trabalho é extremamente curta (na ordem de segundos, como veremos a seguir), as memórias de curto prazo estão envolvidas para operacionalizar grande parte das atividades cotidianas, como conversas e leituras, que envolvem mais tempo e mais informações.

2.2.4 Memória de trabalho

Esse é um conceito muito discutido e, ao mesmo tempo, muito impreciso quando se considera as várias formas nas quais diferentes autores o empregam. Para ilustrar duas representativas posições antagônicas: o neurocientista (vindo da Medicina e da Farmacologia) Izquierdo repreende dizendo que “um erro comum entre os psicólogos norte-americanos é a confusão entre memória de trabalho e memória de curta duração” (2018, p. 97); já o psicólogo britânico Baddeley (2011a, p. 15) rejeita que possa existir uma única “interpretação ‘correta’” quando se trata de teorias psicológicas que transitam por vários níveis de análise: essa conduta geralmente leva ao reducionismo, que não aceita explicações a não ser aquelas microscópicas que vêm das leis da Química e da Física.

Começando pela perspectiva neurobiológica (conforme classificam Gabriel; Morais; Kolinsky, 2016), capaz de oferecer um quadro teórico relativamente mais concreto, Izquierdo (2018) indica que a memória de trabalho é um construto totalmente distinto das memórias de curto e longo prazo: “[...] não só ocupa outras estruturas neurais (fundamentalmente o córtex pré-frontal), como também tem uma farmacologia molecular totalmente diferente [...]” (Izquierdo, 2018, p. 97). Em termos mais simples, o autor também explica que a memória de trabalho não gera arquivos duradouros e, por isso, se mantém ativa por um tempo muito curto, tipicamente alguns segundos, podendo chegar a três minutos, no máximo.

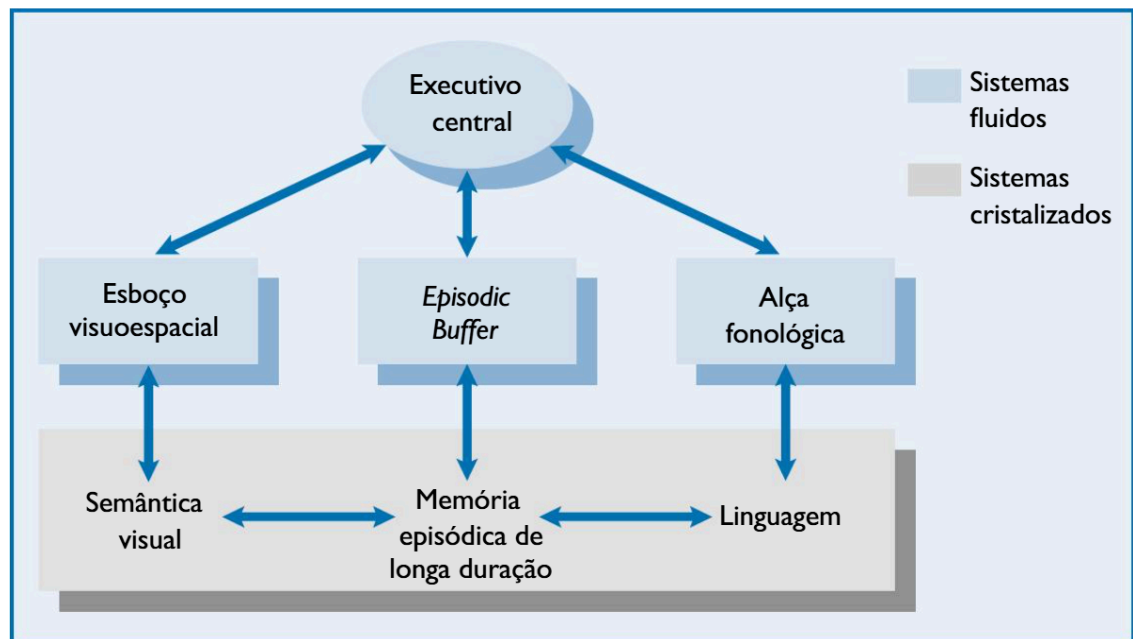
Passando à tradição teórica do campo da Psicologia, o principal paradigma usado é derivado do modelo clássico da memória de trabalho, apresentado por Baddeley e Hitch (1974). Aqui, introduz-se uma capacidade de processamento ao sistema, que acompanha o já bastante convencional armazenamento de curta duração, compondo um sistema único para realizar tarefas complexas (Baddeley, 2011b). O modelo, pois, é multicomponente, formado, quando da sua concepção original:

“[...] por um executivo central e por dois sistemas subordinados, a alça fonológica (*phonological loop*) e o esboço visuoespacial (*visuo-spatial sketchpad*), sendo o primeiro responsável pelo armazenamento temporário de informação verbal e acústica, e o segundo responsável pelo armazenamento temporário de informação visuo-espacial” (Gabriel; Morais; Kolinsky, 2016, p. 64, grifo dos autores).

O próprio Baddeley (2011c, p. 57) elabora um exemplo prático que demonstra como esses três elementos da memória de trabalho se relacionam para uma tarefa de raciocínio. Para uma pessoa lembrar quantas janelas a sua casa tem, primeiro usa o esboço visuoespacial para fazer um desenho mental de todo o seu lar; em seguida, a alça fonológica é colocada a trabalhar para contar, utilizando a voz interna da mente, o número de janelas visualizadas; e o executivo central é o gerenciador que, com seus recursos limitados de atenção, selecionou e executou todas essas estratégias ao longo do processo. Como se pode perceber, é difícil separar o que especificamente é processamento e o que é pura memória de curta duração nos dois sistemas subordinados, o auditivo e o visual. E é muito interessante constatar que – considerando a relevância do modelo multicomponente de Baddeley e Hitch (1974) por tantas décadas e até hoje – toda a experiência introspectiva consciente, aparentemente, pode ser traduzida em sons e imagens.

O esquema da memória de trabalho de Baddeley e Hitch (1974) sofreu algumas modificações para se adequar a críticas conforme foi colocado à prova em variadas pesquisas. Uma importante revisão feita em uma publicação de Baddeley no ano 2000, segundo Gabriel, Morais e Kolinsky (2016), foi a inclusão de um *buffer* episódico, que funcionaria como uma interface temporária capaz de integrar as informações da alça fonológica, do esboço visuoespacial e, especialmente, da memória de longo prazo, que se teorizava desconectado do sistema até então (Figura 2).

Figura 2 – Esquema revisado da memória de trabalho



Fonte: Baddeley (2011c, p. 71). Memória de trabalho de múltiplos componentes segundo a versão de Baddeley (2000).

Propomos, agora, a seguinte indagação: a memória de trabalho e a estação de trabalho neuronal global (Dehaene, 2014) não parecem ter muitos pontos em comum? De certa forma, o modelo de três componentes de Baddeley e Hitch (1974) pode ser visto como um recorte, ainda que muito relevante, de um sistema amplamente maior. Se for considerado o esquema desenhado por Dehaene, Kerszberg e Changeux (1998) e simplificadas as possíveis conexões, a memória de trabalho pode ser um subconjunto que abarque parte dos circuitos de percepção, essencialmente os de manejo da visão e da audição, e pelo menos um dos sistemas presentes nos circuitos de atenção, manejando entradas e saídas na compacta, mas altamente conectada, área de trabalho compartilhada da consciência. Relembrando o que discutimos no final da subseção 2.1.1, o termo “executivo central” é considerado por Dehaene (2020) exatamente o mesmo que “controle executivo”, justamente o sistema atencional que atua como o painel de controle do pensamento consciente.

Nessa compatibilização entre os modelos teóricos discutidos até aqui nesta dissertação, uma vez mais, a “atenção executiva” (Petersen; Posner, 2012; Dehaene, 2020) parece um pouco avulsa. Baddeley (2011c) auxilia, em alguma medida, ao trazer uma leitura posterior que aplicou ao seu conceito de executivo

central: este combinaria dois modos de controle – propostos por Norman e Shallice, em 1986 –, um automático (ou semiautomático), mais simples e baseado em hábitos, e o segundo é o sistema atencional supervisor (SAS), que intervém para resolver conflitos que se criem entre as ações automáticas. Baddeley (2011c) considera que o executivo central e o SAS são análogos. Entretanto, é evidente o quanto as funções do SAS se aproximam da ideia de autocontrole que define a “atenção executiva”. Considerando a dificuldade de se encontrar uma divisão clara entre o que seria controle executivo e o que seria atenção executiva, o termo “executivo central” será admitido aqui como o conceito que reúne tudo aquilo que é controlado pela vontade do indivíduo, das ações, aos pensamentos simples e à metacognição.

2.2.5 Chunking

Refletir a respeito das memórias de curta duração e a forma como são processadas nos conduz a pensar sobre como se estruturam as informações nos seus níveis mínimos e mais elementares na memória de trabalho e, ao mesmo tempo, sobre quais os limites do sistema para lidar com essas unidades. Como recordam Gabriel, Morais e Kolinsky (2016), um artigo de George A. Miller, publicado em 1956, endereçou essas dúvidas e tem se mantido relevante até os dias de hoje. O artigo científico expõe já no título a sua conclusão fundamental: o limite do processamento da informação pela memória de curto prazo é o “número mágico” sete, com possibilidade de variar duas unidades para mais ou para menos. O número médio, portanto, seria de sete itens (dígitos numéricos ou letras do alfabeto avulsos, por exemplo); entretanto, essas sete “gavetas” poderiam ser ocupadas de forma mais eficiente se cada item for, na verdade, um *chunk*, “[...] que seria uma reorganização da informação destinada a ampliar a quantidade de informação recuperada ou mantida na memória imediata” (Gabriel; Morais; Kolinsky, 2016, p. 64). Um exemplo consideravelmente trivial desse processo de *chunking* é considerar que é muito mais fácil gravar uma palavra, por maior que seja, do que suas muitas letras individuais embaralhadas, sem formar sentido. Indo diretamente a Miller (1956), vale notar que ele utiliza para a menor unidade possível o termo *bit* (muito empregado na computação, equivalente a um “bocado”) e, como já mencionado, o agrupamento de *bits* é um *chunk* (algo como um “naco”). Na tradução de Baddeley

(2011b) para o português, *bits* são itens, *chunks* são blocos e *chunking* é a formação de blocos.

Os *chunks*, de acordo com Gabriel, Morais e Kolinsky (2016), são elaborados a partir de uma recodificação que gera blocos cada vez mais informativos. Segundo os autores, essa compactação fica incrementalmente mais poderosa conforme o cérebro aprende e percebe (mesmo que não necessariamente de forma consciente) padrões. Assim, somando, também, a interpretação concordante de Baddeley (2011b), trata-se de uma integração entre a memória imediata e o vasto acervo das memórias de longa duração. Alguns exemplos de melhoria no desempenho diretamente proporcionais à melhoria no *chunking* podem ser o caso de um operador de código Morse – que, inicialmente, só consegue escutar dois tipos de som (o longo e o curto, cada um, um *chunk*) diferentes em sucessão aleatória e, com a prática, passa a perceber padrões de letras, palavras e frases (Miller, 1956) – e o aprendizado da leitura fluente, também parecido, com cada letra exigindo muito processamento individual nas fases iniciais (Gabriel; Morais; Kolinsky, 2016).

A mente humana se aproveita dessa arquitetura de processamento de dados para desenvolver estratégias de memorização, muitas vezes até de maneira espontânea. Baddeley (2011b) elenca casos bastante básicos, como quando uma pessoa precisa memorizar um número de telefone ou uma senha e o faz recitando os dígitos para si mesmo (em voz alta ou mentalmente) dividindo em grupos menores, colocando pausas em pontos da sequência e dando entonações diferentes para os itens. Essa maneira consideravelmente intuitiva que todos têm para agrupar sequências sem sentido *a priori* é uma das fortes evidências que levaram Cowan (2000) a revisar o “número mágico”, isto é, o limite de itens na memória de trabalho, e reduzi-lo para quatro (mais ou menos um; de três a cinco, portanto). O autor concluiu isso após extensa revisão bibliográfica, ao passo que também encontrou uma explicação plausível para os números de Miller (1956), que, de fato, são medidos em vários testes diferentes. Cowan (2000) argumenta que, quando não se isolam totalmente outras variáveis, o sujeito pode ser capaz de usar estratégias (como o ensaio subvocal, repetindo informações com a voz mental) para efetuar o *chunking* de maneira simples. Assim, pode-se pensar que o indivíduo é capaz de recordar três blocos com três itens cada, chegando, na melhor das hipóteses, a nove elementos recordados corretamente. Mesmo em casos de *experts* em memorização, capazes de recuperar séries de 80 dígitos, percebe-se a mesma tendência de

agrupar números em conjuntos de três ou quatro, indicando que essas pessoas montam supergrupos de *chunks* compostos (Cowan, 2000).

Considerando operações mais elaboradas do que a recordação de sequências de dígitos, Miller (1956, p. 95) relata o quanto a linguagem é um potente “reempacotador” de *chunks* maiores e mais densos de informações – basta pensar que um simples nome, que pode ser falado mentalmente, usando apenas a alça fonológica, é suficiente para recuperar todo tipo de informação sensorial, memórias episódicas e semânticas, e assim por diante. As imagens (passíveis de serem carregadas no esboço visuoespacial) também foram consideradas por Miller (1956) como uma forma de recodificação, embora ele não tenha se estendido nesse ponto por considerar mais difícil de testar experimentalmente. Aproveitando essas ideias, que, no fundo, representam apenas as aptidões naturais do sistema cognitivo humano, Farmer e Matlin (2019) comentam técnicas mnemônicas que são usadas com frequência, até mesmo de forma intuitiva. Elaborando imagens mentais, pode-se imaginar em uma mesma representação visual todas as formas que se quer recordar. Utilizando estratégias de organização (que podem ser baseadas igualmente, em imagens ou sons), algumas das possibilidades são a formação de diagramas mentais que coloquem os elementos em uma relação hierárquica, a montagem de uma palavra composta pelas iniciais de todos os itens que se quer memorizar, e a criação de narrativas que vinculem, na ação, todos os elementos relevantes (Farmer; Matlin, 2019).

Se o limite da memória de trabalho é, então, por volta de quatro blocos de informação (Cowan, 2000), Dehaene (2014) esclarece que o pensamento consciente só consegue lidar com um *chunk*⁷ por vez, o que é uma forma mais específica de conceber o espaço limitado da consciência. Referenciando alguns conceitos tratados na seção 2.1, um aspecto relevante é que o *chunking* pode ser operado

⁷ Parte-se do entendimento que Dehaene (2014) utiliza o termo “chunk” para designar o mesmo conceito empregado por Miller (1956) e Cowan (2000). Alguns trechos, do original (Dehaene, 2014) que parecem corroborar essa compreensão: *We are fundamentally reduced to just about one conscious thought at a time (although a single thought can be a substantial “chunk” with several subcomponents, as when we ponder the meaning of a sentence)* (p. 20-21); *This registering appears tightly limited: at any given time, only one chunk of information can go through it. Meanwhile everything else in the visual scene remains unperceived* (p. 33); *Does dual task performance refute the idea that our conscious awareness is structurally limited to one chunk at a time? No. The evidence shows that, even in such cases, we are still tightly limited. We never really process two unrelated items consciously at exactly the same moment* (p. 33); *Leaving aside this discredited brain localization, Descartes was right: our conscious brain cannot experience two ignitions at once and lets us perceive only a single conscious “chunk” at a given time* (p. 125).

voluntariamente, mas há inúmeras situações em que se dá de forma espontânea e automática (Dehaene, 2014; Gabriel; Morais; Kolinsky, 2016). Como aponta Dehaene (2014), a formação de blocos se torna, em grande parte dos casos, um reflexo, muito útil para otimizar ações em tempo real, inclusive aquelas que dependem de mais de uma representação sensorial, já que a união dos *inputs* da visão, audição e outros tipicamente acontece antes de chegar à consciência. Ao apresentar um exemplo de um jogador experiente de xadrez que compreende rapidamente a sua situação no jogo a partir de uma olhada veloz na complexa organização de peças sobre o tabuleiro, Dehaene (2014) atesta o quanto o *chunking* é eficaz, sobretudo para atividades típicas e muito bem aprendidas. No caso de associações menos rotineiras, ainda de acordo com o autor, considera-se que são acionadas coleções de neurônios inusitadas e, por isso, precisa haver a mediação de um estado mental mais consciente. Uma última reflexão que se integra bem com os conceitos de Dehaene (2014) é que os processos de *chunking* coincidem adequadamente com a metáfora do sistema nervoso como se fosse o FBI: os relatórios individuais e setoriais que os neurônios vão combinando chegam na forma de um *chunk* para a apreciação final da consciência.

O *chunking* parece ser a peça fundamental para entender como o sistema cognitivo pode lidar de maneira eficiente com a densidade de informações que chegam pelos sentidos do leitor/ouvinte/espectador em uma situação de processamento de texto multimodal. Esse recurso, operando otimamente, pode criar pacotes de dados ricos e complexos que aproveitam de maneira mais vantajosa os espaços limitados da memória de trabalho na atividade on-line de compreensão e que são mais bem preparados para o armazenamento em arquivos de memória duradouros que serão evocados depois, seja para construir o modelo completo do texto ou para recuperar informações pontuais para responder a um questionário de compreensão. Para o estudo conduzido nesta dissertação, considerando a natureza das informações utilizadas (passagens narrativas) e o tempo curto durante o qual os participantes foram expostos a elas (teste com duração total próxima a 30 minutos), a classificação de memória mais relevante é a episódica de curto prazo, sem contar a memória de trabalho, que tem grande proximidade com o pensamento consciente, discutido na seção 2.1. Contudo, a memória semântica de longo prazo tem um papel relevante, seja na extração da ideia central do texto, seja disponibilizando o acervo de conhecimentos prévios ou de fundo que cada leitor carrega e que foi aproveitado,

em maior ou menor grau, para estabelecer associações que facilitam a memorização (p. ex., recordar com mais prontidão o nome de uma cidade com a qual o sujeito já teve uma relação próxima) e auxiliar de maneira difusa na dedução das respostas aos questionários.

2.3 Multimodalidade e linguagem

Nesta seção, progrediremos, de forma mais aguda e definitiva, para a discussão de habilidades exclusivas dos seres humanos. Isso ficará evidente quando chegarmos à subseção 2.3.2, que abordará a linguagem verbal. No entanto, iniciamos tratando sobre as dinâmicas das associações entre modalidades sensoriais, algo que outros seres vivos também experimentam, e depois refletimos a respeito das várias linguagens que existem para que se estabeleçam diferentes formas de comunicação.

Em primeiro lugar, o termo “modalidade” necessita de uma definição precisa. Esta dissertação utiliza essa palavra e suas derivações (modo, unimodalidade, bimodalidade e multimodalidade) em um sentido bastante basilar, alinhado com o que se costuma encontrar em publicações científicas da área da Psicologia e Neurociência descrito como *modality*. A ideia de “modalidade sensorial”, conforme identificada na tradução em português do texto de Gardner e Gardner (2023), é o que se pretende dizer. Para os autores, as modalidades são o mesmo que os sentidos, ou seja, sistemas com células especializadas para lidar com estímulos específicos e transformá-los em sinais elétricos que são conduzidos pelo sistema nervoso até os centros de maior hierarquia no encéfalo. São cinco os sentidos clássicos, tal como elencados por Aristóteles: visão, audição, tato, olfato e paladar, respectivamente recebidos pelos órgãos: olhos, ouvidos, pele, língua e nariz. Além desses – que, sem dúvidas, são os mais lembrados e compreendidos popularmente –, Gardner e Gardner (2023) destacam que a neurociência contemporânea identifica outras modalidades sensoriais:

[...] as sensações *somáticas* de dor, coceira, temperatura e propriocepção (postura e movimento do nosso próprio corpo); sensações *viscerais* (conscientes e inconscientes) necessárias para a homeostase; e os sentidos *vestibulares* de equilíbrio (a posição do corpo no campo gravitacional) e movimento da cabeça (p. 1398, grifo dos autores).

Tomando uma interpretação muito extensiva, até mesmo organismos como plantas são dotados de sensores primitivos que provocam respostas com vistas a se adequar às condições que o ambiente apresenta (Lamers; van der Meer; Testerink, 2020). Seguindo para uma leitura menos abrangente, Lent (2010, p. 502) confirma que “[...] todos os animais possuem um sistema nervoso”, e, aí sim, é possível perceber uma relevante semelhança entre o funcionamento dos nossos sentidos e os desses outros seres. É trivial notar que um animal, como um cachorro, é capaz de ver e ouvir; e é do conhecimento geral que há situações para as quais esse animal é mais bem equipado (p. ex., um olfato superior) e outras para as quais é pior servido (p. ex., enxerga menos cores). Essas variações assinalam que os sistemas de percepção evoluíram em seres diferentes para atender às necessidades que os variados ambientes exigem de cada organismo.

Outra vez, o verdadeiro diferencial do ser humano fica no seu processamento muito mais complexo das informações. Porém, a especificidade ainda não está na capacidade de integrar informações multimodais, ou seja, de modalidades sensoriais distintas, já que Small e Prescott (2005) afirmam que isso é uma propriedade geral dos sistemas nervosos de todos os mamíferos, provavelmente servindo ao propósito de aumentar ou aprimorar a detecção de um estímulo, fundamentalmente em casos em que a unimodalidade traria apenas dados menos perceptíveis, incompletos ou ambíguos. O fato de essas integrações de modos, muitas vezes, acontecerem nos níveis inconscientes e automáticos (Dehaene, 2014; Gabriel; Moraes; Kolinsky, 2016), conforme recém apresentamos na subseção 2.2.5 sobre *chunking*, permite a dedução de que é um mecanismo um pouco mais simples, que prescinde da exclusividade da evoluída consciência humana.

A respeito do que foi testado em termos de integração multimodal em seres humanos, um interessante relato de Petersen e Posner (2012) relaciona-a com o funcionamento da atenção. Segundo eles, ainda que o foco seja operado considerando somente a modalidade mais relevante para o momento, ocorrem sinergias importantes entre os modos: há muitos casos em que orientar a atenção para um local ou aspecto prioriza não apenas o sentido esperado, mas também outras informações presentes nas demais modalidades do recorte amplificado. Adaptando essa ideia para um exemplo simples: olhar para a tela da televisão facilita também ouvir o que está sendo dito ali. Recuperamos, também, o interessante “efeito McGurk” (ver subseção 2.1.2), que cria um meio-termo de

interpretação para lidar com duas percepções sensoriais que não concordam entre si.

Nota-se o quão rico é o tema da multimodalidade, até mesmo porque é possível analisá-lo a partir de inúmeras combinações entre todas as modalidades sensoriais. Na sequência desta dissertação, contudo, toda a proeminência será dada para os modos visual e auditivo. Há dois motivos essenciais para isso: primeiro, é para esses sentidos que quase a totalidade das mídias que temos até os dias de hoje foi e é desenhada; e segundo, conforme dito na subseção 2.2.4, o modelo da memória de trabalho de Baddeley e Hitch (1974) sugere que os dados fonológicos e visuais-espaciais ocupam uma posição central na nossa cognição.

Antecipando um pouco alguns temas da Linguística, é importante comentar outra acepção bastante recorrente na literatura para a modalidade. A linha de pensamento defendida por Ribeiro (2013) trata todo e qualquer texto como multimodal, sem que exista a possibilidade de uma comunicação que não venha acompanhada de outros modos semióticos. Apesar de ser uma posição bastante sensata – afinal, a escolha de uma tipografia diferente para um texto impresso ou a adoção de outra entonação para uma fala certamente podem modificar o sentido de textos que seriam idênticos –, aderir a ela poderia gerar confusões para o trabalho aqui proposto. Dessa forma, preferimos acrescentar outra noção de Gardner e Gardner (2023), a de submodalidades, para identificar variações de propriedades em um mesmo canal sensorial (p. ex., cores e formas, ambas dentro do modo visual).

2.3.1 Signos e linguagens não verbais

Torna-se possível reconhecer as diferenças mais importantes entre humanos e animais quando se analisa as formas vastamente mais complexas de comunicação que aqueles desenvolveram e seguem utilizando. Isso não significa dizer que muitas outras espécies não sejam capazes de trocar informações de maneiras realmente notáveis, fazendo uso de meios especializados vocais, ultrassônicos, gestuais, olfativos, entre outros (Kaplan, 2014). O que ocorre, como explica Lopes (2000), é que essa linguagem animal tem características bastante distintas da humana. Os animais, segundo o autor, realizam uma mera comunicação de comportamentos, por meio de habilidades herdadas diretamente da genética; ou seja, não apresentam variações no sentido do discurso (nem no tempo, nem no espaço), não são

aprendidas culturalmente e não são compostas de convenções acordadas entre a comunidade.

O que configura a linguagem humana como algo fundamentalmente distinto é que ela faz o uso de signos, unidades nas quais “[...] o significado é diferente (isolável) da substância do elemento material que o expressa (seu significante)” (Lopes, 2000, p. 37-38). É preciso dar um salto momentâneo e adiantar algumas ideias que serão tratadas de maneira específica na próxima subseção, pois assim nos habilitamos a usar a definição clássica de Saussure, pai da Linguística moderna, destinada especificamente para o signo linguístico; partindo dela, em seguida, será possível generalizar o conceito de forma mais clara. Saussure (2006) define o signo linguístico como uma entidade psíquica com duas faces intimamente unidas que, na terminologia que ficou consagrada, são o significado e o significante. Para o significado, Saussure (2006), primeiramente, usa a palavra “conceito”, que, explicado de outra forma, é a ideia que se tem de algo (p. ex., uma árvore tal como é percebida pelos sentidos no mundo real). Para o significante, o termo que utiliza é “imagem acústica”, que vem a ser a representação virtual e ideal da palavra que nomeia o conceito (p. ex., a palavra “árvore” dentro da mente, não necessariamente proferida pela fala). Saussure (2006, p. 80) insiste nesse ponto a fim de sublinhar a complexidade do processo, que não se reduz à simplória união de uma coisa qualquer a uma palavra: o que se interliga são o conceito e a imagem acústica. Deixando de lado, por enquanto, o aspecto puramente verbal, podemos aplicar a equação $signo = significante + significado$ para outras convenções. Um exemplo de Lopes (2000, p. 42) ajuda a compreender outras possibilidades: um motorista que encontra em seu caminho uma placa que informa sobre a interdição da pista à frente muda de trajetória tanto quanto como se enxergasse o objeto concreto que está interrompendo a estrada; ele não vê só a placa: ele vê o próprio conceito/significado.

Como se pode entender, as relações sgnicas se estendem para muito além das palavras. Essa noção já era clara para Saussure (2006), que considerava a Linguística como uma parte de uma ciência mais abrangente, que ainda viria a ser desenvolvida para considerar os outros signos existentes na vida social e que faria parte da Psicologia geral. Antecipada como “semiologia” pelo autor, a teoria geral dos signos acabou assumindo oficialmente a denominação “semiótica”, favorecendo o termo preferido pelos estudiosos anglófonos e alemães (Nöth, 2003). Sendo uma disciplina tão ampla, não é de se surpreender que existem várias escolas de

pensamento sobre o que exatamente é a Semiótica e qual o seu objeto de estudo. Para as finalidades desta dissertação, a concepção pluralista de Nöth (2003, p. 17) é suficiente e adequada: “A semiótica é a ciência dos signos e dos processos significativos (semiose) na natureza e na cultura”.

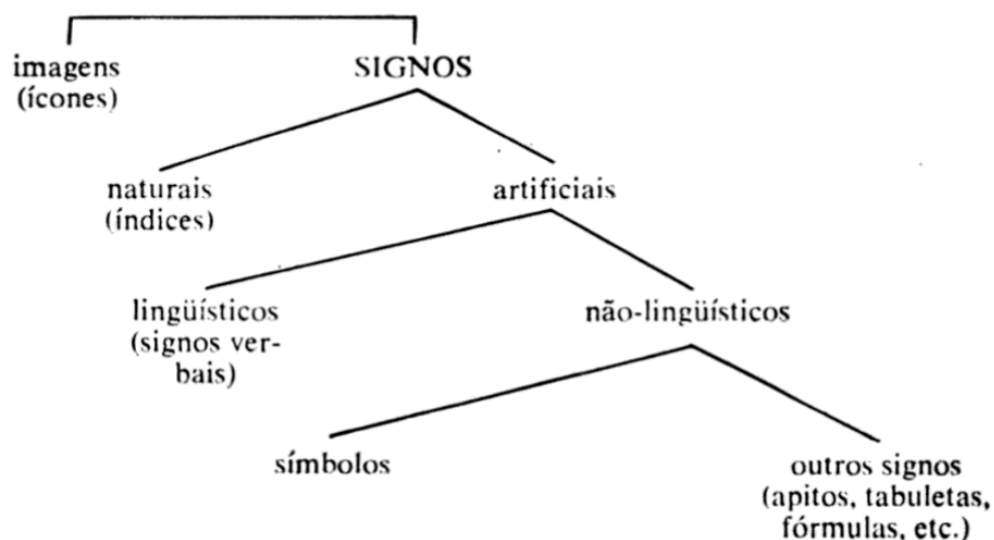
Essa noção ampla está de acordo com a esquematização do simbolismo linguístico elaborado por Lopes (2000, p. 46). Por meio dela (Figura 3), o autor classifica todos os signos em um diagrama de árvore, com categorias que possuem subdivisões, totalizando seis denominações diferentes. De fato, há uma sétima categoria que se localiza na periferia, fora do esquema principal: os *sinais não-sígnicos*, também referidos como “imagens” ou “ícones”. Ali, estão colocados os artefatos que são produzidos a partir da mente humana, mas que têm uma relação de total semelhança entre significante e significado; em outras palavras, eles não derivam de qualquer convenção cultural e, portanto, inexistente o processo de semiose. O exemplo categórico que Lopes (2000) usa é que uma foto de um amigo representa esse próprio amigo. É uma visão bastante extrema e, como será comentado a seguir, as imagens também podem habitar pelo menos outra divisão do signo. Vale destacar algumas posições teóricas divergentes que Nöth (2003) compila, dando conta de que a iconicidade não seria capaz de escapar totalmente das convenções culturais e que, no limite, praticamente qualquer imagem pode assumir qualquer significado. Outro ponto interessante sobre os sinais não-sígnicos é trazido por Lopes (2000, p. 46) quando comenta as onomatopeias, que não fariam parte desse grupo. Apesar de existir uma imitação feita pelo significante para se assemelhar ao significado, se não houvesse uma relação pelo menos parcialmente arbitrária, nada explicaria por que um mergulho em português faz “*tchibum*” e em inglês faz “*splash*”.

Indo, agora, diretamente para o topo do diagrama de árvore feito por Lopes (2000), os signos se dividem em dois grandes grupos: naturais (índices) e artificiais. Da primeira, não se desdobram subcategorias. Os signos *naturais* são produzidos por forças não culturais (elementos da natureza) e, então, a semiose ocorre só em uma das pontas do processo, com o humano interpretando relações de causa e efeito entre a fumaça e o fogo, por exemplo, existentes de forma não-convencionada entre os falantes de uma comunidade linguística.

Os signos *artificiais* são aqueles que Lopes (2000, p. 42) define como os “signos propriamente ditos”. Dividem-se em duas categorias: linguísticos (signos

verbais) e não-lingüísticos. Sobre os primeiros, já discutimos brevemente a partir da definição de Saussure (2006) e aprofundaremos o assunto mais adiante. Os segundos, os signos *não-lingüísticos*, são importantes para entender as outras tantas formas de comunicação humana que não passam diretamente pela mediação das palavras. Nessa classificação, o autor faz a última divisão em duas classes, sendo uma delas intitulada apenas “*outros signos*” (contendo exemplos como placas de trânsito, apitos e fórmulas matemáticas) e a outra representando os *símbolos*. Estes últimos são os que podem comportar as imagens sob outro entendimento, já que uma fotografia pode não apenas ser o ícone (sinal não-sígnico) do amigo retratado, mas sim simbolizar a amizade de maneira geral. De acordo com Lopes (2000, p. 44), “os símbolos são objetos materiais que representam noções abstratas [...]”, e um exemplo básico dessa relação é a cruz, que carrega a ideia do cristianismo. O conceito representado é sempre muito mais amplo do que o símbolo apresenta concretamente, que é apenas uma parte do todo, e, sendo assim, esse tipo de signo comunica uma ideia “[...] deficiente ou inadequada parcialmente [...]” (Lopes, 2000, p. 44). Assim, a representação está sujeita a condições de polissemia e de sinonímia (muitos símbolos podem significar o mesmo conceito). Outro ponto destacado pelo autor é que o símbolo e o conceito têm uma relação que é parcialmente motivada, de modo que é comum que ambos tenham vários traços afins (p. ex., caveira designando a morte).

Figura 3 – Esquema do simbolismo lingüístico



Essa esquematização do simbolismo linguístico, tal como intitulada por Lopes (2000), é muito útil para entender as várias faces que o signo pode apresentar; no entanto, há que se colocar em perspectiva algumas das nomenclaturas usadas pelo autor. Ele, conforme exposto acima, tratou os signos artificiais linguísticos como sendo exatamente o mesmo que signos verbais. Dessa forma, nota-se que linguagem, para o autor, é um rótulo que apenas se aplica para aquela comunicação que se utiliza de palavras. A posição é reforçada por Lopes (2000, p. 35) quando explica que chamar de “linguagem animal” as formas de comunicação dos outros seres é um uso metafórico (e não próprio) do conceito “linguagem”, o mesmo raciocínio que se aplicaria a outras manifestações humanas, como as artísticas e as gestuais. O conceito estrito de linguagem para Lopes (2000) está em consonância com o que Saussure (2006) conceptualiza mais especificamente como “língua”, que é o objeto particular da disciplina Linguística, restando para a “semiologia” (posteriormente, Semiótica) estudar os demais signos nas relações humanas (conforme discutido anteriormente, definições pluralistas, como a de Nöth (2003), levam em conta todo e qualquer processo de significação existente). Por mais que possamos falar, sem prejuízo, em “semioses” para designar os processos de comunicação na esfera mais ampla, consideramos mais claro e intuitivo chamar de “linguagem” também as outras trocas de signos, reservando “língua” ou “linguagem verbal” para delimitar o que é feito exclusivamente usando palavras.

Dito isso, a ideia de “linguagem não verbal” é um grande guarda-chuva que comporta, sobretudo, o que Lopes (2000) chamou de “signos artificiais não linguísticos” e, em menor grau, os signos naturais (índices) e os ditos sinais não-sígnicos (imagens/ícones). Já fizemos alguns comentários sobre o uso de imagens (como as fotografias) para comunicação, então o que é importante considerar agora são as outras linguagens não verbais que são primordiais e que acompanham intimamente a comunicação verbal oral. Como explica Dionísio (2004), fala-se tanto com a voz quanto com o corpo, tanto que recursos como expressões faciais ou entonações específicas podem desde contribuir para a construção do sentido a até mesmo substituir um enunciado verbal. Entre todos os cinco recursos não verbais que Steinberg (1988, p. 3) identifica que ocorrem normalmente em conversas, estão três que nos interessam de forma mais relevante: a cinésica, todos os movimentos corporais, como postura e olhar; o silêncio, que, apesar de ser a ausência de construções, comunica algo justamente pelo contraste que estabelece

com o conteúdo auditivo regular, além de dar a oportunidade para que os outros signos e códigos se tornem mais evidentes; e a paralinguagem, isto é, a emissão de sons pelo aparelho fonador mas que não pertencem à língua usada, como assobios e outros ruídos que não formam palavras.

Há um quarto elemento, não tratado pontualmente por Steinberg (1988) – está embutido na categoria da paralinguagem –, que é preciso destacar: a prosódia, ou seja, o contorno melódico no qual a fala se materializa. Dionísio (2004) a situa em um local curioso, ao mesmo tempo tendo natureza linguística e caráter não verbal. É pertinente, portanto, recuperar definições a respeito de paralinguagem e extralinguagem, que são termos que a literatura trata com alguma imprecisão entre as suas fronteiras. Laver (1994) oferece uma classificação bastante objetiva, considerando o nível linguístico como aquele constituído por unidades fonológicas e gramaticais; também é destacado que não se resume à oralidade, abrangendo a escrita e as línguas de sinais. A paralinguagem, que, para o autor, é onde está enquadrada a entonação vocal, é tida como similar à linguagem no sentido de ser codificada (ou seja, é produzida e interpretada de acordo com convenções culturais, e não universalmente) e dissimilar a ela por não ter a possibilidade de comunicar sentido pela forma como é estruturada (diferentemente da sintaxe verbal, mudar a ordem dos elementos paralinguísticos não altera significados por si só). O nível extralinguístico é, como explica Laver (1994), todo o resíduo que fica ao se retirarem as codificações linguísticas e paralinguísticas. Resta, dessa forma, uma parte informativa (que agrega dados novos, assim como a parte comunicativa, mas, diferentemente desta, não é intencional) que pode ser bastante rica, trazendo elementos que ajudam a identificar o falante, como o seu gênero, porte físico, ânimo, entre outros (Laver, 1994).

Barbosa (2012) trabalha a linguagem, a paralinguagem e a extralinguagem a partir de uma óptica bastante próxima à de Laver (1994) e utiliza essas divisões para situar a prosódia em um espectro mais amplo e, aparentemente, mais adequado para se pensar sobre a influência difusa que esse recurso carrega:

[...] a prosódia está, no cenário de pesquisa atual, associada a fatores linguísticos como acento, fronteira de constituinte, ênfase, entonação e ritmo, a fatores paralinguísticos como marcadores discursivos (e.g., “né”, “entendo”, “an-han”) e atitudes proposicionais (e.g., “confiante” e “duvidoso”) e sociais (e.g., “hostil” e “solidário”), além de tratar de fatores extralinguísticos como as emoções. Todos esses fatores se combinam com

aspectos sociais e biológicos indiciais como gênero, faixa etária, classe social, nível de escolaridade, entre outros (Barbosa, 2012, p. 13).

Os elementos prosódicos podem ser entendidos como uma ou, mais exatamente, várias submodalidades (conforme conceito de Gardner; Gardner, 2023) comunicadas para o canal da audição. As propriedades básicas envolvidas, não surpreendentemente, são as mesmas que delineiam a linguagem da música, entre elas: sensação de volume (*loudness*), sensação de altura ou afinação (*pitch*), duração (*duration*), ritmo, entonação, acentuação e timbre ou qualidade da voz (Kompe, 1997; Barbosa, 2012).

Embora possa ser um aspecto menosprezado ou esquecido, até mesmo por se estar constantemente exposto a ele, a prosódia (em união com a paralinguagem gestual) é um dos instrumentos mais potentes e versáteis para se fazer a comunicação em uma língua pouco ou nada conhecida. Basta visualizar um cenário muito comum: ao encontrar um estrangeiro, a pessoa recorre a mímicas e a produções de som não verbais (às vezes, podem até ser palavras do seu próprio idioma, ainda que com chances remotas de serem compreensíveis) e bastante exageradas em seu contorno, na esperança de transmitir algo que seja corretamente decodificado pelo interlocutor. Uma situação mais corriqueira ainda de dificuldade de comunicação é aquela pela qual todos os seres humanos passam: a aquisição da língua materna. Em um quadro que descreve o percurso da criança para o processamento de material linguístico, Corrêa (2008) demonstra o quanto os elementos prosódicos são reconhecidos muito cedo: com um ou dois dias de idade, o bebê já reconhece o timbre de voz da sua mãe (mesmo que seja consideravelmente diferente, em termos de espectro de frequência, do som que o feto ouvia dentro do útero); com um mês, discrimina sons da fala a partir de parâmetros acústicos e articulatórios; com seis meses, são sensíveis às fronteiras entre orações na fala; com nove a dez meses, já produzem balbucios que os adultos reconhecem como pertencentes à sua língua. O desenvolvimento é consistente e a curva de aprendizado é íngreme, com a criança exponencialmente entendendo os padrões mais complexos, logo desenvolvendo sintaxe e largo vocabulário; mas é importante lembrar que tudo se inicia rapidamente na parte mais primordial da comunicação auditiva.

2.3.2 Linguagem verbal

De maneira diluída e paralela, muito já discutimos nesta seção sobre a língua ou linguagem verbal. Seria difícil que fosse diferente, já que essa categoria é aquela sobre a qual se estuda há mais tempo e sobre a qual existem os conhecimentos mais sistemáticos. Isso acaba forçando que todas as outras linguagens, das mais diferentes semioses, sejam comparadas com o desenvolvimento supremo da comunicação: as palavras. A trajetória teórica que desenvolvemos no trabalho até aqui finalmente chega a um ponto no qual os humanos estão indiscutivelmente isolados e muito avançados com relação aos outros animais.

O artigo de Hockett (1960) desvenda de forma mais específica por que a linguagem verbal é tão única. O pesquisador identificou treze características que estruturam sistemas de comunicação, analisando uma série de comunidades animais, que se estende desde abelhas até macacos. Concluiu que nove das funcionalidades costumam ocorrer nas formas de linguagem animal mais elaboradas, mas que há outras quatro adicionais que apenas estão presentes na totalidade nos sistemas humanos (isto é, alguns outros animais exibem uma ou duas delas, nunca todas simultaneamente). Essas características especiais são as seguintes: deslocamento (*displacement*), a capacidade de falar sobre coisas em tempos e/ou espaços remotos; produtividade (*productivity*), a possibilidade de falar coisas novas, jamais ditas, e ainda ser compreendido; transmissão tradicional (*traditional transmission*), ou seja, o ensinamento e a aprendizagem são feitos culturalmente; e dualidade de padrões (*duality of patterning*), por meio da qual um mesmo conjunto limitado de unidades discretas (segmentos; como fonemas ou morfemas) pode ser arranjado em outra ordem e gerar significados totalmente distintos (p. ex., “ovo” e “voo”).

Se a linguagem verbal é manifestamente exclusiva dos humanos, também é virtualmente impossível considerar que a consciência altamente desenvolvida dessa espécie encontre qualquer paralelo no reino animal (conforme visto na seção 2.1). Isso não parece, de maneira alguma, um acaso ao se considerar a interdependência e colaboração que há entre a língua e o pensamento consciente. Dehaene (2014), com a finalidade de explicar o que tornaria a introspecção humana algo tão superior à das outras espécies, conjectura uma resposta que coloca a linguagem no centro de tudo. O autor credits ao que ele chama de “linguagem do pensamento” – ao que

tudo indica, o mesmo que a linguagem verbal humana em estado potencial, seja trafegando apenas dentro da mente, seja ainda em desenvolvimento na criança – a capacidade de gerar pensamentos complexos, recursivos, aninhados e, acima de tudo, altamente interconectados. Como propõe Dehaene (2014, p. 251), pensamentos aparentemente triviais combinam domínios absolutamente diferentes: tamanho, espaço, pessoas, objetos, cor, lógica, ação, etc. (p. ex., “a camiseta azul do João lhe serviu bem e não vai ter que ser trocada na loja da esquina”). Cada tipo de dado é processado pelo módulo especializado no cérebro, e tudo se encontra na estação de trabalho neuronal global (Dehaene, 2014), sendo que todas as informações têm que ser convertidas para uma “moeda” que pode ser recebida e trocada livremente no sistema: para isso, símbolos como as palavras são o formato mais versátil. E, contando com as habilidades comunicativas mais evoluídas elencadas por Hockett (1960), esses signos ainda podem ser desmontados, remontados, combinados e transportados para outros contextos.

A maleabilidade extrema das palavras, seja na sua existência dentro da mente ou fora dela, na sociedade, entra em choque com algumas concepções tradicionais da Linguística. Lopes (2000) considera o signo linguístico verbal como sendo unívoco: “[...] a representação do signo verbal é perfeitamente adequada: a palavra *casa* recobre, com precisão, o seu significado” (p. 45, nota de rodapé 10, grifo do autor). A partir dessa propriedade, o autor também conclui que a inteligibilidade verdadeira só existe quando todos os signos são convertidos para a forma verbal, já que todos os outros artefatos comunicativos não são consistentes. Desdobramentos teóricos mais recentes da Linguística disputam essa certeza de que haveria de um sentido único para o signo verbal. Ferrari (2011) explica que a semântica cognitiva defende uma concepção enciclopédica do significado linguístico, na qual a visão de dicionário (reducionista; um termo é igual a outro termo) está incluída, mas não é a totalidade. Sendo assim, seguindo com a autora, a Linguística Cognitiva rejeita que um item lexical tenha apenas um aspecto “essencial” que o defina, de modo que o conhecimento de mundo, o contexto e o uso são todos corresponsáveis por construir os significados.

É importante ter em mente que a linguagem verbal é o que realiza a mediação do pensamento consciente. Uma visão teórica extrema chegou a postular que a maneira como falamos define a forma como pensamos. Denominada como “hipótese de Sapir-Whorf” ou “relativismo linguístico”, a tese propunha que uma determinada

língua restringe o recorte da realidade feito, mentalmente, pela comunidade que a fala, necessariamente direcionando a visão de mundo (Ferrari, 2011). Esse recorte, de acordo com a explicação mais técnica de Ferrari (2011, p. 31), é o que se chama de categorização: “[...] o processo através do qual agrupamos entidades semelhantes (objetos, pessoas, lugares etc.) em classes específicas”. Como acrescenta a autora, essa estratégia mental inata está diretamente relacionada com as limitações que temos para processar dados e armazenar e recuperar memórias. Voltando à hipótese do relativismo linguístico, Ferrari (2011, p. 34) aponta que pesquisas subsequentes – fundamentalmente, a desenvolvida por Berlin e Kay em 1969, que analisou a nomenclatura de cores em mais de noventa línguas – sugerem que o processo ocorre mas não é tão determinante quanto o imaginado, já que há indícios de pontos centrais compartilhados na experiência perceptual de diferentes povos, e o que se altera mais significativamente entre culturas são as fronteiras das categorias ao redor desses marcos principais.

A respeito da categorização realizada por meio da linguagem verbal, é interessante recuperar o conceito de *chunk* de memória (subseção 2.2.5) e também a noção de Miller (1956) sobre língua ter uma capacidade tremenda de reempacotar informações em unidades pequenas e ricas em informações. É fascinante perceber que um dos cerca de quatro compartimentos disponíveis na memória de trabalho (Cowan, 2000) pode ter seu espaço igualmente esgotado seja com o algarismo “3”, seja com o termo “Brasil”, com todas as possíveis interpretações, sensações e associações que este conceito possa trazer consigo.

2.3.3 Redundância, irrelevância e interferência entre modalidades

Nesta seção, trabalhamos dois eixos complementares da transmissão de informação: por um lado, as modalidades sensoriais, que servem de canais de comunicação, e, por outro, os signos, que comprimem e codificam os dados. A partir da união dessas duas dimensões e da ponderação sobre as incalculáveis possibilidades de interações entre elas, pode-se considerar as implicações práticas que essas combinações de parâmetros têm sobre os processos de significação.

Quando mais de uma modalidade ou submodalidade trabalham com a mesma informação, pode-se considerar que são congruentes. Logicamente, exames mais profundos no campo da Semiótica e da Linguística podem invalidar a possibilidade

de mais de um signo carregarem uma informação idêntica; entretanto, partindo do pressuposto de que essa discussão vai mais além do que o necessário para o nível de análise que propomos para esta dissertação, uma boa concessão é que trabalhemos com a noção de signos que, no mínimo, representam a intenção de se parecer muito e, no limite ideal, são perfeitamente equivalentes.

Nesses cenários de redundância, aparentemente está a única possibilidade positiva de associação entre modalidades que ocorre puramente pelo reforço da informação, não por outros fatores paralelos. O estudo de Mousavi, Low e Sweller (1995) sugere que materiais educativos (problemas de Geometria) foram mais bem entendidos quando as ilustrações vieram acompanhadas por instruções narradas em áudio e não escritas ao lado dos desenhos, dando a vantagem para a apresentação em modalidades sensoriais distintas. Mayer e Moreno (1998) tiraram conclusões parecidas, pois seus testes indicaram que vídeos instrucionais (infográficos animados) tiveram maiores índices de compreensão no grupo que assistiu aos materiais com explicações faladas e não em forma de legendas. A pesquisa de Brasel e Gips (2014), a partir de audiovisuais publicitários, constatou que os espectadores foram mais hábeis, no geral, para lembrar informações relevantes (marca do carro anunciado e atributos do veículo) nos testes em que legendas redundantes, com o exato mesmo texto falado no filme, foram exibidas na tela.

Convém analisar outras interações que têm o propósito de melhorar a memorização e a aprendizagem. Um ponto é que há a possibilidade de até mesmo dados redundantes gerarem piores desempenhos, provavelmente por causarem uma sobrecarga à memória de trabalho. O mesmo estudo de Brasel e Gips (2014) citado acima traz achados importantes nesse sentido: uma peça publicitária é tornada pior com a inclusão de legendas redundantes caso já seja muito complexo visual e verbalmente; e algo curioso é que legendas, por mais que sejam recebidas pelo canal visual, aumentam somente a complexidade percebida pelos espectadores da dimensão verbal, sugerindo que esses signos acabam convertidos e ocupando espaço da alça fonológica, juntamente com todo o material nativamente auditivo. Outro ponto é que há indícios para considerar que o uso engenhoso de submodalidades pode não só promover o reforço de uma informação, mas, até mesmo, estendê-la. Schön e colaboradores (2008) observaram em sua pesquisa sobre a percepção de padrões de organização de sílabas nos sons de uma língua artificial que os participantes que ouviram sequências de áudio falado de forma

convencional não foram tão bem sucedidos quanto o grupo que foi exposto a versões cantadas (portanto, com informação adicionada na dimensão musical) do mesmo material.

Desviando da congruência em direção à incongruência de informações, uma possibilidade de associação entre modos e signos é a irrelevância. Se for levado em consideração o conceito que aparece na literatura científica como “efeito sonoro irrelevante” (*irrelevant sound effect*) (Page; Norris, 2003; Baddeley, 2011b), na verdade, já estamos tratando de interferência, pois os sons secundários usados nessas pesquisas só são considerados irrelevantes porque não dizem respeito à tarefa principal que se tenta executar, o que não significa que eles não sejam nocivos para o desempenho. Uma possibilidade alternativa é conceituar que alguns sons e outros estímulos concorrentes possam ser irrelevantes só sob o recorte das informações, conservando outros efeitos paralelos. Isso poderia proporcionar um quadro teórico útil para considerar, por exemplo, quando músicas supostamente ajudam nas tarefas reduzindo o estresse ou melhorando o ânimo do sujeito, como nos estudos de Jiang *et al.* (2013) e Perham e Currie (2014). Talvez, apenas se possa falar em irrelevância autêntica nos casos em que o estímulo que distrairia o indivíduo acaba bloqueado pelos mecanismos de atenção seletiva, e, então, ele é tratado da mesma forma como se nem existisse em primeiro lugar.

Levando em conta as informações incongruentes recebidas ao mesmo tempo, o pior cenário é a interferência. Fazendo uma breve digressão, pode-se conceber situações nas quais a interferência – mas, aí, operando no nível físico, não necessariamente perceptual – poderia ser benéfica: um som menos disruptivo, como um ruído da natureza, poderia ser usado para mascarar, acusticamente, vozes conversando, que provavelmente incomodariam mais o ouvinte (Vasilev; Kirkby; Angele, 2018). Mas, voltando especificamente às interferências entre signos e/ou modalidades com mensagens incongruentes, o prejuízo da distração é a ocorrência mais verificada pelos estudos. Na metanálise de Vasilev, Kirkby e Angele (2018), que compilou achados de 65 estudos sobre distração auditiva durante a leitura, o que foi encontrado é um efeito pequeno, mas negativo e confiável, de quaisquer tipos de som de fundo, sendo que fala inteligível e música com letras são as piores distrações.

O artigo de Vasilev, Kirkby e Angele (2018) também é muito relevante porque reúne as teorias sobre a distração sonora e tenta relacioná-las com os dados

quantitativos que foram encontrados. Primeiramente, os autores excluem a possibilidade de a *hipótese da interferência fonológica*, um paradigma bastante influente nas pesquisas desse campo, explicar sozinha os efeitos percebidos, já que essa concepção considera que o atributo problemático fundamental seria a similaridade dos sons que habitam ao mesmo tempo a alça fonológica do sujeito (p. ex., a pseudopalavra “troqua” atrapalharia mais do que qualquer palavra real a recordação do dígito “4” apenas porque aquela é feita de fonemas semelhantes e confundiria o ensaio subvocal). Os resultados encontrados por Vasilev, Kirkby e Angele (2018) não permitem conceber que o sentido das palavras seja indiferente ao ouvinte e, assim, a *hipótese da interferência semântica* e também a *interferência por processo* (uma das metades da *teoria duplex da distração auditiva*) ganham força, pois ambas consideram que conteúdos carregados semanticamente, tais como falas e músicas vocais, comprometem outras atividades que simultaneamente estejam extraíndo o significado de outros signos. Apesar da aplicabilidade dessas duas últimas hipóteses, elas não conseguem explicar por que sons não verbais também causam perturbações, ainda que em menor grau. Em outra linha, a segunda parte da teoria duplex da distração auditiva, que é chamada *captura da atenção*, não chegou a ser avaliada profundamente pelos autores por causa de dificuldades metodológicas inerentes a ela, mas tem potencial para, futuramente, explicar os fenômenos medidos, relacionando-os com os temas dos sistemas de atenção, que tratamos na subseção 2.1.1. O último quadro teórico que os autores examinaram é a *hipótese da alteração de estado*, que põe a perturbação na conta apenas das variações no sinal acústico, estando também correlacionada a desvios de atenção. Essa teoria se mostrou relevante, porém insuficiente, porque não é capaz de explicar como um ruído artificial constante atrapalha tanto quanto uma música instrumental, que é muito mais imprevisível.

As possibilidades de redundância, irrelevância ou interferência entre modalidades representam o cerne da nossa pesquisa, uma vez que são os efeitos que se procurou identificar no processamento de textos equivalentes apresentados em diferentes formatos com densidade e qualidade variáveis de elementos sensoriais. Algumas das condições experimentais concebidas e desenvolvidas – nomeadamente, aquelas com vídeos e imagens estáticas – impõem que se pense a respeito de como são recebidos e empacotados pelo sistema cognitivo (relembrando o conceito de *chunking* detalhado na seção 2.2) os sinais que compõem as

linguagens não verbais. Já os signos da linguagem verbal, sejam comunicados pela escrita ou pela oralidade, configuram-se como o conteúdo concreto que o leitor/ouvinte deveria compreender em nossos experimentos e que foi avaliado em termos de quão próximas as respostas dadas pelo sujeito estavam das informações apresentadas nos textos que precediam os questionários.

2.4 Leitura

Nesta última seção do capítulo de referencial teórico, apresentaremos a leitura, o quarto eixo conceitual principal desta dissertação. Mantendo a progressão que partiu dos aspectos cognitivos elementares, portanto, mais ligados à natureza, e avançando para as criações verdadeiramente humanas, ou seja, culturais, a capacidade de ler e escrever é o ponto culminante.

Ao se admitir que a linguagem verbal é um artefato exclusivo da nossa espécie, conforme discutimos na seção anterior, e que sua manifestação primordial é pela comunicação oral, presente em todas as comunidades humanas conhecidas (Hockett, 1960), até mesmo por ser uma capacidade básica transmitida por genes específicos de acordo com alguns pesquisadores (Wolf, 2019, é bastante enfática nesse ponto bastante debatido), a leitura de signos linguísticos verbais é ainda mais restrita. A escrita é uma invenção, comparativamente, muito recente, tendo em torno de 5,4 mil anos (Dehaene, 2012) e não é uma capacidade dominada por todos os povos. Os dados que existem são pouco definitivos, mas ajudam a ter uma estimativa: de acordo com a metodologia e os levantamentos de Eberhard, Simons e Fennig (2025), das mais de 7 mil línguas vivas catalogadas, acima de 3 mil, provavelmente, não possuem um sistema de escrita. Efetuando outro recorte e por uma métrica mais familiar, a Unesco (2024) constatou que, em 2023, 13% dos jovens e adultos do mundo eram iletrados, ou seja, de acordo com a definição de Moraes (2019, p. 9) para o sentido mais elementar de letramento, não possuíam habilidade de ler e escrever.

Obviamente, quando se fala em leitura, o habitual é se pensar nos sistemas de escrita alfabéticos ou que se servem de outros grupos de sinais verbais. Aliás, o entendimento de “leitura” tende a ser tão restritivo que não costuma ser aplicado nem mesmo a alguns tipos de linguagem verbal que não fixam signos escritos. É o que acontece com a Língua Brasileira de Sinais (Libras), por meio da qual a

comunidade surda brasileira se comunica verbalmente, sinalizando e compreendendo (Mertzani; De Monte; Fernandes, 2023; Gomes; Nascimento; Lopes, 2023), e não necessariamente lendo.

Tal como já foi feito nesta dissertação a respeito da linguagem, é válido colocar a leitura em uma perspectiva mais ampla. Mengarda (2012, p. 123) defende que, muito antes de o ser humano ler os códigos de escrita, ele passou milhares de anos lendo os sinais da natureza, desde as direções informadas pelas estrelas até a identificação de animais que deixavam marcas no ambiente. Aproveitando-se a classificação do simbolismo linguístico de Lopes (2000), o que esses ancestrais faziam era extrair o sentido de signos naturais, ou índices. A própria escrita, se for compreendida unicamente como a impressão de uma marca duradoura executada por uma pessoa, pode ser rastreada até eventos muito mais antigos, sendo que Dehaene (2012, p. 198) conta que a Caverna de Chauvet, na França, uma das mais antigas grutas ornadas conhecidas, apresentava “[...] formas muito sofisticadas de arte gráfica” 33 mil anos antes da nossa era. É também Dehaene (2012, p. 331) que comenta que “Não se conhece praticamente nenhuma sociedade humana que não pratique o desenho ou a gravura”, e que muitas das culturas desenvolveram desenhos com forma simbólica sofisticada.

Não é raro encontrar a palavra “leitura” usada para designar a operação de extração de sentido a partir de artefatos visuais, sejam fotos ou ilustrações, estáticas ou em movimento. De acordo com Sardelich (2006, p. 453), “A expressão leitura de imagens começou a circular na área de comunicação e artes no final da década de 1970, com a explosão dos sistemas audiovisuais”, e, até hoje, não é um termo consensual entre os variados campos de estudo que se ocupam da imagem: boa parte considera impossível que se “leia” esse tipo de signo. Oportunamente, voltamos às discussões da subseção 2.3.1, considerando que os processos significativos, ou semioses, ocorrem em todo lugar e, se é possível admitir o uso extensivo de “linguagem” para as comunicações visuais ou artísticas, por exemplo, também parece razoável chamar amplamente de “leitura” a decodificação desses e de outros signos. Sardelich (2006), inclusive, cita que uma forma tradicional de se ler imagens é pela abordagem formalista que, justamente, foi influenciada pela Semiótica. Essa variedade de leitura, ainda segundo a autora, incorpora outras terminologias diretamente conectadas com a linguagem verbal, como “alfabetismo visual” e “sintaxe visual”, e também trabalha com noções de denotação (aquilo que é

visto de forma “objetiva”) e conotação (as apreciações do intérprete sobre o que é sugerido). É claro que, caso o desejo seja aprofundar esse pensamento, pode-se chegar a concluir que uma imagem não tem a capacidade de conter elementos denotativos puros, mas, aí, é um entendimento que também se aplicaria facilmente à leitura de textos com pretensão literária ou mesmo, no limite, a qualquer texto verbal. Para o nível de análise proposto para esta dissertação, é importante assumir que os signos verbais e os não verbais possam ser suficientemente objetivos para a leitura, ou extração de sentidos, feita por quem os ouve ou vê.

Consideradas as várias leituras possíveis em sentido amplo, torna-se claro que a ação de ler é, com frequência, tratada como sinônima de interpretar ou compreender. A propósito da distinção entre estas duas últimas, Livingstone (2008) coloca a interpretação em um nível mais elaborado – conotativo –, trabalhando com convenções, conexões culturais e significados míticos ou ideológicos. Já a compreensão está relacionada ao que o texto efetivamente transmite de informação, podendo o receptor dele ser avaliado em termos de tê-lo entendido de forma certa ou errada. Em linha com o que foi exposto no parágrafo anterior, os aspectos interpretativos não serão aprofundados nesta dissertação; e a parte da compreensão será abordada na subseção seguinte, tendo em vista sua importância central para este trabalho.

Se as variadas leituras das inúmeras linguagens podem ser tratadas, sem prejuízo, como compreensão e interpretação (em diferentes níveis de equilíbrio), há de se especificar o que há de único na leitura *stricto sensu*, dos signos linguísticos verbais na sua forma escrita. Tomando como partida a perspectiva de Hoover e Gough (1990) da “visão simples de leitura” (*simple view of reading*), conforme explicado por Freitag (2020), há dois processos conduzidos ao mesmo tempo pelo leitor do texto verbal: compreensão linguística e decodificação. Como dito anteriormente, as habilidades relacionadas ao compreender, além de estarem presentes, de uma forma ou outra, em todos os processos de semiose, também serão analisadas na próxima subseção. Então, a diferença fundamental que se encontra na leitura dos signos verbais escritos e que cabe apreciar agora é o seu processo de decodificação ou, colocando de forma mais específica, a conversão de sinais visuais para suas correspondentes formas sonoras e representações semânticas.

Aproveitando um diagrama que esquematiza a visão simples da leitura de Hoover e Gough (1990), o modelo de cordas de Scarborough (2001; Freitag, 2020) concebe a leitura fluente como sendo um espesso cordão (Figura 4), que é formado pela união de dois cadarços (os dois processos; compreensão e identificação de palavras), cada um destes formado pelo entrelaçamento de alguns fios mais finos. Atentando apenas para o cadarço que representa a identificação das palavras, são três os fios que o compõem: consciência fonológica, que envolve o reconhecimento de sílabas, fonemas e outras estruturas afins; decodificação, relacionada à correspondência entre grafemas e fonemas (o “princípio alfabético” nos sistemas de escrita que utilizam o alfabeto); e o reconhecimento direto e visual de palavras familiares. Essas três habilidades, segundo Scarborough (2001), devem se tornar cada vez mais automatizadas para que seja possível chegar à leitura fluente.

Figura 4 – Modelo de cordas de Scarborough (2001)



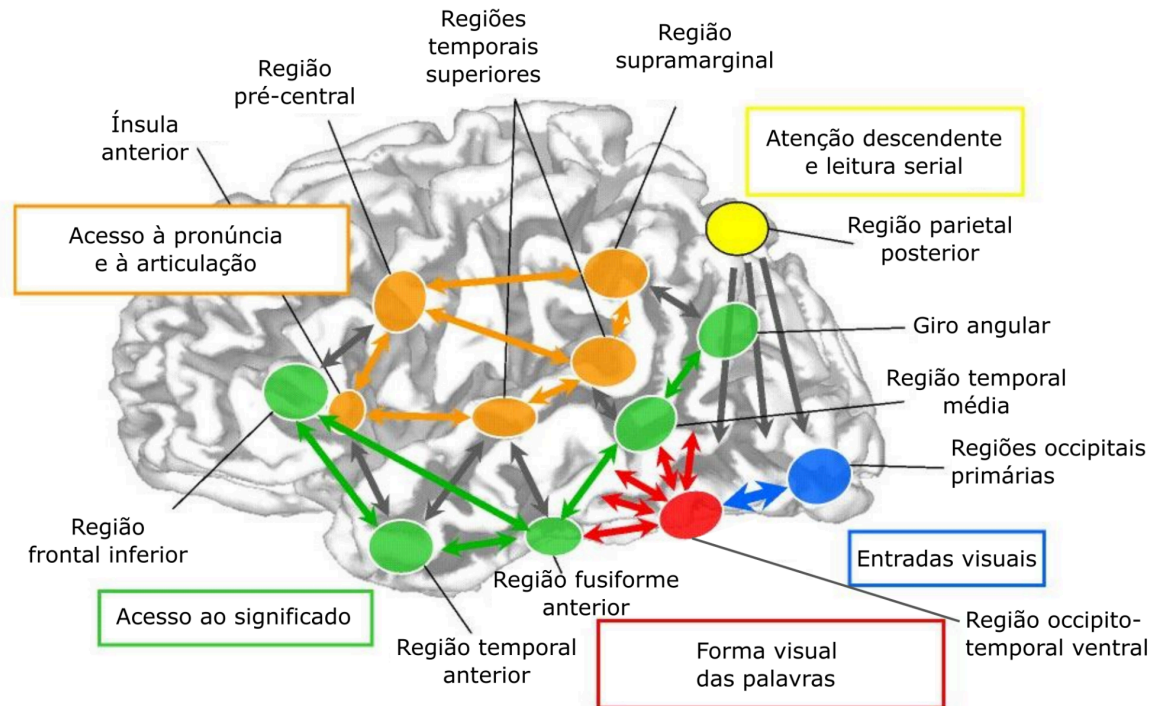
Fonte: Brasil (2019, p. 29). Baseado em Scarborough (2001).

As operações primordiais que o cérebro efetua e que fazem ser possível ler estão relacionadas com a função básica de transformar caracteres gráficos (letras, no sistema alfabético) em sons. Dehaene (2012, p. 168) resume bem a revolução que isso representa: “Eis que há cinco mil anos, os primeiros escribas descobriram um poder escondido no cérebro humano, o de aprender a transmitir a linguagem através dos olhos”. Levando em conta que o sistema nervoso humano, há muitos

milênios, de acordo com uma linha teórica dos estudos da aquisição da linguagem, já contém circuitos próprios para desenvolver a linguagem verbal naturalmente e que a leitura é demasiado recente para que a evolução biológica pudesse incluí-la na genética da mesma maneira (Dehaene, 2012; Wolf, 2019), a capacidade de ler surge do uso de outros artifícios. É com o termo “reciclagem neuronal” que Dehaene (2012, p. 166) nomeia este processo muitíssimo mais veloz, por meio do qual uma invenção cultural se aproveita da plasticidade do cérebro e adapta circuitos pré-existentes para novas tarefas. Essa reciclagem, no entanto, não é, de maneira alguma, ilimitada, tanto que uma das principais teses de Dehaene na sua obra “Os neurônios da leitura” (2012) é que não foi o córtex que evoluiu para a leitura, mas sim os sistemas de escrita que evoluíram para funcionarem adequadamente com as potencialidades e limitações existentes no cérebro.

O que ocorre, então, é que o aprendizado da leitura coloca, em conexão direta, os circuitos neuronais ocupados com o tratamento de estímulos visuais e todo o sistema cerebral da linguagem. Duas áreas são fundamentais para o estabelecimento dessa interface. Uma delas, de acordo com a explicação de Dehaene (2012, p. 77), fica na região occípito-temporal esquerda e se especializa “[...] na análise da forma das letras, seu reconhecimento e sua reunião em palavras”, recebendo o nome de região da forma visual das palavras – vale destacar que Gabriel, Morais e Kolinsky (2016, p. 71) apresentam outras nomenclaturas que também aparecem na literatura: *Visual Word Form Area* (VWFA) ou área da forma visual da palavra. A outra região, conforme Dehaene (2012, p. 121), na parte superior do lobo temporal, é conhecida como *planum temporale* e “[...] reage à compatibilidade entre as letras e os sons [...]”, sendo o ponto onde a parte visual e a auditiva da leitura se encontram (Figura 5).

Figura 5 – Redes corticais da leitura



Fonte: Dehaene (2012, p. 78). A *Visual Word Form Area* (VWFA) está indicada pela cor vermelha, e o *planum temporale* se localiza onde está a elipse laranja da direita do grupo identificado como “Regiões temporais superiores”.

De quase nada serviria a leitura se ela não alcançasse o significado das palavras. Resulta que o destino final do processo é a área semântica, e um modelo teórico propõe que se pode chegar a ela por dois caminhos: pela rota fonológica ou pela rota lexical do Modelo de Leitura de Dupla Rota, tradução informada por Freitag (2020) para a concepção de *Dual-Route Cascaded model* desenvolvida por Coltheart *et al.* (1993). A compilação de pesquisas apresentada por Dehaene (2012) dá forte sustentação a essa hipótese, já que exames de imageamento cerebral (tais como os presentes no estudo de Simos *et al.*, 2002) acusam diferentes padrões de ativação neural, em termos de localização, intensidade e tempo, a depender de como os signos verbais são tratados. Assim como defende Dehaene (2012), o indivíduo letrado em qualquer idioma, ao se deparar com um signo verbal escrito que seja identificável por ele, inicia o processo da mesma forma, usando as áreas visuais primárias e, logo em seguida, encaminhando para a região da forma visual das palavras. A partir daí, a destinação varia de acordo com o material lido e também a habilidade do leitor. A rota fonológica é a mais longa e também a mais básica, considerando que – certamente em um sistema alfabético e provavelmente

em todos os outros sistemas de escrita – toda a palavra é, inicialmente, para alguém em processo de letramento, uma pseudopalavra que precisa ser convertida sistematicamente em sons para que seja comparada com os nomes já conhecidos em linguagem oral. Como indica Dehaene (2012), é essa trilha que aciona o *planum temporale* e, fazendo uma retomada de Saussure (2006), ativa a representação da imagem acústica do signo. Se a palavra não tiver significado, o processo pode terminar aí ou ser encaminhado para áreas motoras de fala, por exemplo, para que seja recitada; agora, se houver conteúdo semântico, segue-se para as regiões neurais responsáveis pela recuperação de sentido, uma operação que depende de muitos circuitos diferentes e espalhados, mas que pode ser vista convergindo para as regiões frontais e temporais esquerdas do encéfalo (Dehaene, 2012, p. 127).

A outra possibilidade é que a leitura faça uma espécie de ligação direta e avance, sem escalas, da região da forma visual das palavras até os centros semânticos, ou seja, use a rota lexical. Esse trilho que faz emergir antes o conceito do que a imagem acústica ou o significante – nos termos de Saussure (2006) –, é o que Scarborough (2001) coloca como a capacidade de reconhecimento visual direto na identificação das palavras. Isso acontece para vocábulos familiares, como uma forma de otimização da leitura, e também, como explica Dehaene (2012), para palavras com ortografia irregular ou sinais gráficos mais abstratos (p. ex., de idiomas como o chinês), ambos os quais não são passíveis de serem recuperados por uma conversão direta e simples entre caracteres e sons: nesses dois casos, a escrita precisa buscar primeiro o sentido para que, depois, venha a lembrança da correta pronúncia. Um destaque importante feito por Dehaene (2012) é que as leituras em todas as línguas fazem uso tanto da rota lexical quanto da fonológica, variando o equilíbrio que cada uma daquelas faz entre essas duas estratégias.

É fácil esquecer o quanto todos esses processos são sofisticados e o quão duro costuma ser o aprendizado da leitura, afinal, ela se torna uma habilidade como tantas outras, disponível para uso em condições normais e com grandes chances de ser utilizada todos os dias. A metáfora dos “óculos letrados” que não podemos retirar, relatada por Kolinsky e Morais (2018) e Morais (2019), é muito acertada para dar a dimensão do quanto ler se torna uma segunda natureza. Dehaene (2012) faz um pertinente comentário nesse sentido, dizendo que, apesar de Saussure (2006) estar correto sobre o signo linguístico ser arbitrário na sua construção (no sentido que imagem acústica e conceito não precisam ter qualquer semelhança inerente,

apenas precisando ser convencionados dentro de uma comunidade), ele deixa de sê-lo depois que a pessoa aprende uma palavra. Uma prova contundente disso é que um sujeito letrado pode acessar o significado de um termo mesmo em condições em que este seja exibido de forma subliminar (Dehaene, 2012, 2014), tamanha a automatização dessa capacidade de identificação e decodificação das palavras.

2.4.1 Compreensão leitora

Indo para além das questões específicas da decodificação da parte visual dos signos verbais escritos, o componente fundamental da leitura que precisa ser abordado é a compreensão. Retomamos a visão simples da leitura de Hoover e Gough (1990), e, mais especificamente, sua diagramação no modelo de cordas de Scarborough (2001; Freitag, 2020), concentrando-nos no segundo cadarço que, entrelaçado com o primeiro, o da identificação das palavras, compõe a leitura fluente. Esse feixe é, oportunamente, rotulado com o nome “compreensão da linguagem”, o que permite entender que é algo que vai além da atividade de leitura tradicional, envolvendo competências supramodais que são usadas para entender a comunicação oral e outras formas de manifestação da linguagem verbal (como a sinalização para surdos). Extrapolando o modelo teórico, é lícito, também, identificar aspectos que podem ser relacionados a outras linguagens ou semioses. Dos cinco fios que integram o cadarço da compreensão da linguagem, apenas um pode ser considerado mais limitado à leitura em sentido estrito, que é o que Freitag (2020) traduz como sendo o nível de alfabetização (*literacy knowledge*, para Scarborough, 2001), que envolve conceitos sobre a escrita, gêneros literários, entre outros. Os demais quatro fios são, em maior ou menor grau, igualmente aplicáveis à linguagem verbal (e, em maior ou menor grau, a outros sistemas de signos): conhecimento de fundo, sobre fatos e conceitos; vocabulário, concernente à amplitude, precisão e articulação; estruturas da linguagem, relativas à sintaxe, à semântica e outros níveis linguísticos; e raciocínio verbal, que lida com inferências, metáforas e demais construções contidas em textos.

Se a parcela da identificação de palavras é treinada para ser cada vez mais automática, a porção de habilidades de compreensão da linguagem é desenvolvida a fim de se tornar cada vez mais estratégica (Scarborough, 2001; Freitag, 2020). A

partir desse entendimento do ato de ler dividido em duas metades (aparentemente assimétricas), um raciocínio é permitido. Já discutimos um tópico central desta dissertação, sobre a leitura quando realizada juntamente com outra tarefa simultânea; no entanto, a própria leitura já é, em si, multitarefa. Conforme discutido na seção 2.1, Dehaene (2020) só considera possível a execução de múltiplas atividades adicionais sem perda de desempenho quando elas são desempenhadas de maneira inconsciente, só havendo espaço para um trilho de informação na consciência. Assim se sucede com a leitura fluente ideal: decodificação totalmente automática e compreensão e interpretação conduzidas pelo pensamento consciente. Condições menos ideais no departamento de identificação das palavras, como aquelas empreendidas por pessoas iniciando o seu letramento, precisam de muito processamento consciente e, por isso, a compreensão do texto se torna menos eficiente (Gabriel; Morais; Kolinsky, 2016).

Não se poderia, então, automatizar a própria ação de compreender? Ao que tudo indica, é impossível, a não ser quando se levam em conta eventos minúsculos quanto à sua expressão, sem qualquer utilidade prática para quem queira entender um texto inteiro, por exemplo. Das cinco categorias de processamento inconsciente concebidos por Dehaene (2014; ver subseção 2.1.2), apenas duas parecem que se enquadrariam no caso da leitura, nomeadamente as informações pré-conscientes e os estímulos subliminares. Contudo, se o texto estiver em estado pré-consciente, ou seja, codificado corretamente mas sem receber atenção, de nada adianta, porque, como sustenta Dehaene (2020), o que não pode ser percebido, seja por distrações ou outros empecilhos, não pode ser aprendido (admitindo, aqui e provisoriamente, compreensão e aprendizagem como ideias bastante aproximadas); e se o estímulo com significado a ser extraído estiver em nível subliminar – em outras palavras, fraco demais, mesmo recebendo atenção –, o seu efeito no cérebro é muito tímido e limitado. Assim como visto na subseção 2.1.2, conteúdos semânticos são acessados de maneira subconsciente (reforçando a metáfora de Kolinsky e Morais (2018) dos “óculos letrados” não-removíveis), mas, seja em linguagem verbal ou em conteúdo pictórico, um signo só será entendido muito superficialmente e sem chance de influenciar um comportamento caso seja recebido com pouca intensidade ou por pouco tempo (Fei-Fei *et al.*, 2007; Dehaene, 2014).

Antes de prosseguir, é relevante aprofundar e estender alguns conceitos tratados de forma passageira ainda há pouco. A reflexão de Back (2023, p. 6)

autoriza que, no limite, “Compreender já equivaleria a aprender”, havendo distinção pouco clara ou até nenhuma diferença entre os dois verbos. Parece mais útil, entretanto, aproveitar outras definições de Back (2023, p. 5) que estabelecem contornos ligeiramente distintos para cada uma:

Compreender significa basicamente tornar familiar algo ainda estranho. E algo apenas se torna familiar para alguém quando ele/a pode relacioná-lo com outras coisas já familiares. [...] O que caracteriza a aprendizagem? Ela exige de quem está aprendendo a habilidade para realizar por si próprio certas atividades relacionadas à coisa compreendida.

Por um lado, portanto, Back (2023) coloca a compreensão como uma atividade intelectual, empenhada em encontrar sentido para a informação nova, e, por outro, a aprendizagem tem uma dimensão mais prática. Por essa linha, não se pode negar que exista um componente de memória, independentemente de qual for sua duração, para a composição de uma aprendizagem, já que o sujeito precisa carregar em si o que foi compreendido para que seja possível executar a devida ação correspondente. Abrir esse flanco para a consideração da memória nesses processos nos habilita a conceber uma terceira vertente, que Back (2023, p. 5) sugere ao dizer que “[...] ninguém pode ser, primeiro, desprovido de compreensão e, só depois de aprendê-la, passar a compreender as coisas”. Uma aprendizagem sem compreensão, sendo assim, poderíamos considerar como uma simples memorização mecânica, algo decorado de forma superficial, a exemplo das formas de “aprendizagem” animal (Lent, 2010; conforme discutido na seção 2.2), ou incorporado ao inconsciente na forma de conexões neurais latentes (Dehaene, 2014) que armazenam as memórias de procedimentos (Lent, 2010; Izquierdo, 2018).

Compreender, como se torna claro, é uma ação cognitiva que depende de que informações sejam integradas e não apenas estocadas. Tratando especificamente de aprendizagem – que consideraremos uma compreensão memorizada –, Dehaene (2020) a define como a criação de um modelo mental interior do mundo exterior, uma representação que é atualizada e ajustada conforme se incluem novas informações naquele, providas de testes de hipóteses geradas e projetadas. A compreensão leitora, analisando pela conceptualização de Kintsch (1998, 2005; Kintsch; Rawson, 2005), que é bastante compatível com a visão de Dehaene (2020) recém apresentada, é formada por estágios de construção e de integração de informações em diferentes níveis.

O Modelo de Construção-Integração (CI) de Kintsch é uma série de processos de compreensão textual explicados da seguinte forma em Kintsch e Rawson (2005). Na base, está o processamento linguístico, que lida com as já discutidas etapas de decodificação dos caracteres escritos e reconhecimento de palavras, avançando um pouco mais até a análise sintática (*parsing*), pela qual são atribuídas as devidas funções às unidades na estrutura das frases. Em seguida, vem a análise semântica, na qual os sentidos das palavras são, nas condições determinadas e permitidas pelo texto como um todo, convertidas em proposições que, por sua vez, são interconectadas em uma rede complexa, que vem a ser a microestrutura do texto. Essa construção é, a seguir, organizada em unidades mais abrangentes, como tópicos ou ideias principais, e esse esquema mais global é a macroestrutura do texto. Esta e a microestrutura juntas compõem o texto-base (*textbase*), a representação fidedigna daquilo que é expressamente comunicado pelo texto. Como colocam os autores, neste ponto da assimilação da leitura, já existe compreensão, mas é um entendimento superficial, que, no máximo, permitiria que o leitor fosse capaz de apenas reproduzir o que leu.

Continuando com Kintsch e Rawson (2005), a compreensão profunda só advém da construção do modelo situacional, ou seja, “[...] um modelo mental da situação descrita pelo texto”⁸ (Kintsch; Rawson, 2005, p. 211). Aqui, aparecem de forma mais evidente as operações de integração, já que o texto-base é combinado com o conhecimento prévio do leitor e suas intenções ou metas, recuperando informações que podem extrapolar o domínio verbal, como imagens ou emoções. Entretanto, é evidente que a construção e a integração se complementam em todas as etapas, pois, como explica Kintsch (1998), o processo construtivo é caótico e indiferente ao contexto, apenas podendo ser ordenado pela satisfação das restrições impostas pelas operações integrativas. Dessa maneira, mesmo no nível linguístico básico, múltiplas percepções construídas inconscientemente devem ser reforçadas e atenuadas para se manter ativas apenas aquelas que se ajustam com as informações que vão chegando e com aquelas já contidas na memória. É muito semelhante ao funcionamento das assembleias neuronais (Dehaene, 2014; ver subseção 2.1.2), com vários processos paralelos inconscientes que culminam em uma representação informativa consciente.

⁸ Do original: “[...] a mental model of the situation described by the text (Kintsch; Rawson, 2005, p. 211).

As hipóteses que são geradas de forma paralela e a todo momento durante o processo de leitura são as chamadas inferências, utilizadas para preencher as lacunas da compreensão. Algumas delas são rapidamente descartadas ou confirmadas por informações que vão aparecendo explicitamente no transcorrer das frases. Mas, como é quase impossível que um texto seja absolutamente explícito, a inferenciação tem um papel importante até mesmo no nível da microestrutura, por exemplo, para recuperar referentes ligados a pronomes, a sinônimos ou a outros termos omitidos ou não tão detalhados (Kintsch; Rawson, 2005). A propósito do termo “inferência” como qualquer preenchimento de espaços do entendimento, Kintsch (1998) o considera um tanto quanto indiscriminado, porque é usado tanto para operações muito triviais de recuperação de dados quanto para as inferências lógicas autênticas, que lidam com geração de novas informações derivadas tendo por base o texto; no entanto, o autor cede em favor dessa nomenclatura empregada de maneira mais livre.

As inferências são especialmente importantes para a construção do modelo situacional. Aí, entra em jogo, de maneira mais óbvia, o equilíbrio entre os processos que Kintsch (2005) classifica como *bottom-up* (ascendentes; originados nas entradas sensoriais) e os *top-down* (descendentes; vindos da memória ou do conhecimento). O autor é definitivo ao atestar que nada acontece sem ambas essas metades, de modo que a interação normal é que o componente descendente guia a compreensão e a parte ascendente é quem a restringe àquilo que é aceitável. É fácil visualizar casos típicos em que a desarmonia desses processos gera leituras extremamente falhas: quando domina o *bottom-up*, pode no máximo ocorrer uma memorização pobre do texto; e quando praticamente só se apresenta o constituinte *top-down*, o leitor é capaz de produzir entendimentos que sequer têm relação com o que está escrito, já que as informações estão sendo puxadas quase somente da memória. Desequilíbrios menos aberrantes, contudo, são bastante comuns em diferentes cenários de leitura, sendo que especialistas tipicamente leem de maneira bem mais fluida os textos de assuntos nos quais são proficientes, uma vez que as informações chegando pelos olhos encontram muita afinidade com o que já foi aprendido previamente (Kintsch; Rawson, 2005).

Há, então, algumas considerações sobre a eficiência da compreensão leitora que devem ser feitas. Para além desses casos exemplares de peritos consumindo material de dentro do seu campo de domínio, existem outras maneiras de se ter um

bom desempenho. Em linha com o que propõe Scarborough (2001), sobre as capacidades da compreensão da linguagem funcionarem tanto melhor quanto forem mais estratégicas, Kintsch e Rawson (2005) defendem que o complexo processamento da leitura se torna mais eficiente quando o indivíduo tem um bom acervo de conhecimentos transversais, desde um rico vocabulário repleto de interassociações até a destreza para conseguir montar a macroestrutura textual.

Ainda assim, não parece haver escapatória para os limites biológicos do ser humano, que já discutimos anteriormente: os cerca de quatro *chunks* de informação simultâneos na memória de trabalho (Cowan, 2000) e o único bloco que pode ser ocupado na área de trabalho do pensamento consciente (Dehaene, 2014). Por essa óptica, é extremamente pertinente a ponderação de Kintsch e Rawson (2005) a respeito de esses sistemas mencionados serem muito escassos para dar conta de toda a complexidade de processamento que o leitor tem que articular para entender um texto longo e complexo. O que não parece mandatário é introduzir um novo conceito, que esses autores rotulam como “memória de trabalho de longo prazo”, para comportar o que de fato transcorre. Isso não quer dizer que seja uma ideia equivocada tida por eles: aparentemente, é só um recorte diferente, que afasta mais o *zoom* para além do restrito modelo clássico da memória de trabalho (Baddeley; Hitch, 1974) e que, se continuar olhando de mais longe e enquadrando circuitos adicionais, chegará a englobar toda a estação de trabalho neuronal global (Dehaene; Kerszberg; Changeux, 1998). O fundamental é admitir que a leitura, ocorrida na prática e em tempo real, não tem como se sustentar sem conexões com as memórias de longo prazo. Resta, enfim, a dúvida sobre como os gargalos da mente são superados. Pode ser que existam processos muito rápidos e elaborados, que enviam para a consciência só aquilo que tem que ser tratado em detalhes em um preciso momento e que, ao mesmo tempo, ocupam dinamicamente os supostos quatro compartimentos da memória de trabalho enquanto vão sendo construídas novas memórias mais estáveis (de curto e longo prazo), que podem ser guardadas ou ativadas, alternadamente, conforme os modelos textuais e situacionais vão incorporando novas informações.

Sobre problemas de compreensão leitora, Rayner *et al.* (2016) comentam que o indivíduo volta atrás no texto quando precisa recuperar informações perdidas ou esquecidas, uma observação aparentemente singela, mas que nos coloca em contato direto com os temas tratados nas outras seções desta dissertação. Se algo

foi perdido, então não foi lido com atenção e não teve a chance de ser tratado conscientemente. Se algo foi esquecido, a informação até pode ter chegado à consciência, mas não fixou memórias porque, talvez, não tenha sido associada a conhecimentos prévios pertinentes ou não foi um processo modulado favoravelmente pelo estado de ânimo do leitor. Avançando um pouco, podemos pensar, igualmente, em um problema vindo da incorreta interpretação do texto, e, aí, é uma falha no nível da linguagem. Associada a todos esses percalços, está a multimodalidade sensorial, quer agindo como um auxílio para a correção de inconsistências, quer atuando como algo excedente que apenas preenche espaços mentais que seriam mais vantajosamente deixados livres para o processamento da tarefa principal pura, livre de qualquer tipo de competição.

2.4.2 Instrumentos de avaliação da compreensão leitora

É evidente que a compreensão leitora é um processo complexo e multifacetado, o que leva a pensar como pode ser possível medi-la de maneira acurada e representativa. Não é uma tarefa banal fazê-lo, de modo que não há um método incontestável disponível atualmente para gerar medidas informativas e precisas para todas as situações.

Em um primeiro momento, é válido olhar de maneira mais crítica para testes que podem, facilmente, ser empregados de forma muito equivocada. Nesse grupo, estão incluídas as técnicas que mensuram o tempo de leitura e as que avaliam tão somente a memória. A velocidade com que um texto é lido é um dado certamente informativo, mas não espelha, sozinha, a compreensão leitora. Uma leitura rápida fala muito mais da eficiência dos processos locais de decodificação. A pesquisa de Salles e Parente (2002) com 76 crianças de segunda e terceira séries do Ensino Fundamental sustenta o que parece incontroverso: uma boa precisão na leitura de palavras isoladas está ligada à maior velocidade de leitura textual, mas esta tem uma associação que é apenas moderadamente fraca com a compreensão textual. Tratar rapidez como sinônimo de boa compreensão dá margem para alegações ultrajantes de pessoas que se dizem capazes de ler um livro da série *Harry Potter* em 47 minutos, o que equivaleria a uma taxa de 4.200 palavras por minuto (ppm), e outros que poderiam processar até 30.000 ppm (Rayner *et al.*, 2016). Para colocar em perspectiva os tempos típicos de processamento de linguagem – que, quase

certamente, indicam os limites da mente humana para essa habilidade –, Rayner *et al.* (2016) informam os seguintes valores em ppm: fala confortável, 150-160; fala rápida, 250-400; áudio acelerado, por volta de 300; e leitura silenciosa hábil, 200-400. Mesmo valores mais típicos declarados para leitura rápida ficam em taxas bem mais modestas de 500-750 ppm.

A outra abordagem reducionista da compreensão leitora que destacamos é tratá-la como se fosse representada pela simples recuperação de memórias. Esse ponto já era visto como problemático até mesmo por Baddeley e Hitch (1974, p. 63), que entendiam haver uma distinção a ser mais bem explorada entre retenção e compreensão, mesmo que ambas parecessem muito proximamente relacionadas. A exemplo do que foi discutido no parágrafo anterior sobre o tempo de leitura, a capacidade de armazenar e recuperar memórias a partir de um texto é, sem sombra de dúvidas, um indício muito pertinente, mas ele não basta para caracterizar compreensão profunda. Voltando a Kintsch e Rawson (2005), o leitor só tem uma visão sólida do texto quando constrói e integra um modelo situacional com seu conhecimento prévio. Caso o sujeito não atinja esse nível, ele ainda pode recuperar, de forma mais superficial e mecânica, informações textuais macroestruturais, na melhor das hipóteses, e, mais tipicamente, aspectos tão microscópicos como palavras isoladas.

Mantendo em mente as limitações inerentes para se efetuar medições de compreensão leitora, é fundamental conhecer a compilação de métodos de pesquisa do tema elencados e classificados por Tomitch (2008). Segundo a autora, todas as metodologias podem ser enquadradas em dois grandes grupos: dados comportamentais e dados cerebrais. Começando pelos últimos, ali estão medições que procuram ativações de regiões do cérebro que sejam correlatos da cognição humana durante a atividade de leitura. Para tanto, utilizam-se ferramentas consideravelmente sofisticadas, tais como eletroencefalograma (EEG), imagem por ressonância magnética funcional (*fMRI*; da sigla em inglês) e tomografia por emissão de pósitrons (*PET*; da sigla em inglês). Pela natureza dessas medições, todas envolvendo capturas de “fotografias” de um ponto no tempo do decorrer da tarefa, podemos incluí-las integralmente em uma outra subcategoria proposta por Tomitch (2008), a do foco sobre a compreensão enquanto processo, que será detalhada a seguir.

Prosseguindo com a explicação de Tomitch (2008), o segundo grande grupo de metodologias de pesquisa em leitura é o da coleta de dados comportamentais. Esses são relativamente mais acessíveis e, então, são usados mais habitualmente. A autora elabora uma subsequente divisão em três classes aqui, cada uma com o foco em um recorte diferente: a situação de leitura, o processo e o produto final. A primeira abordagem é mais holística, considerando leitor, texto e contexto, e analisa a situação global por meio de entrevistas, diários e questionários. A segunda, preocupada com compreensão enquanto processo, intenciona aferir dados que caracterizem a leitura ocorrendo *on-line*, e pode utilizar instrumentos como a leitura em voz alta, também sob a variação de “[...] protocolos verbais produzidos durante a leitura (*think aloud procedure*)” (Tomitch, 2008, p. 39). A terceira e última subclasse concebida pela autora é a da compreensão como produto, a abordagem que é a mais usual entre todas. Nela, os dados são coletados de maneira *off-line*, após a execução da leitura, e o propósito que se tem é medir o que o indivíduo foi capaz de guardar ou assimilar. As tarefas e questionários aplicados referentes a essa categoria são amplamente conhecidos, até mesmo por acompanharem a educação formal dos alunos: múltipla escolha, verdadeiro ou falso, relacionar itens dispostos em colunas diferentes, escrita de resumo e verbalização livre (*free recall*). Talvez o instrumento menos familiar nesse grupo – isso considerando essencialmente seu nome técnico no âmbito da pesquisa científica – seja o teste de Cloze, que envolve completar lacunas inseridas de forma sistemática em uma passagem textual. Essa técnica, desenvolvida, testada e primeiramente apresentada por Taylor (1953), é definida por Tomitch (2008) como o preenchimento de espaços em um texto com cerca de 300 palavras, com a primeira e a última frase inalteradas, e com o apagamento das palavras ocorrido em intervalos regulares (tradicionalmente, nas posições múltiplas de cinco no texto: quinta, décima, décima quinta palavra, e assim por diante).

Certas de que não há consenso sobre a forma mais apropriada de avaliar a compreensão leitora, Sousa e Hübner (2015) identificam as limitações inerentes a cada tipo de tarefa. No grupo das avaliações que elas chamam de receptivas (onde estão múltipla escolha, verdadeiro ou falso e verificação de sentenças), as autoras concordam que são mais fáceis de se corrigir e avaliar objetivamente, mas isso vem às custas dos riscos de o leitor poder simplesmente adivinhar as respostas, basear-se praticamente apenas na memória ou empregar estratégias e inferências

customizadas unicamente para resolver os problemas específicos presentes no teste. As tarefas de avaliação produtivas (questões abertas, resumo e reconto), por outro lado, sofrem com problemas de subjetividade da correção, possíveis falhas em habilidades paralelas do leitor – ainda que ele entenda perfeitamente o texto, pode ter dificuldade para se expressar na forma escrita ou oral, bem como pode não ser familiarizado com o gênero textual que é solicitado a entregar – e, também, possível sobrecarga indevida da memória. Inserido no mesmo grupo das tarefas produtivas por Sousa e Hübner (2015), está o teste de Cloze, mas sua deficiência é um tanto peculiar: ele é inferior para avaliar a compreensão global, já que se presta mais a examinar processos locais ligados a vocabulário e sintaxe.

Um destaque importante que se faz necessário é a respeito da metodologia de rastreamento ocular (*eye tracking*), uma tecnologia que Riess e Gabriel (2020) identificam como promissora para as pesquisas em leitura e cognição. A partir da categorização de Tomitch (2008), apresentada há pouco, fica claro que é uma abordagem do subgrupo que estuda os processos conduzidos pelo leitor, uma vez que são gravados os movimentos dos olhos efetuados sobre o texto exibido em tela. O que é menos evidente é a classificação desse procedimento em um dos grandes grupos, de dados cerebrais ou comportamentais. Obviamente, o dispositivo de *eye tracking* não demonstra de forma direta as áreas do cérebro que vão se ativando, mas ele se aproxima disso, admitindo os comportamentos oculares como indícios muito representativos e autênticos das operações cognitivas realizadas ao fundo, até mesmo porque a movimentação ocular, quase na totalidade, decorre fora do controle consciente (Riess; Gabriel, 2020).

A maneira mais elementar e objetiva de considerar o que revelam os movimentos oculares é pensar em termos daquilo que foi discutido na subseção 2.1.1, fundamentalmente sob a óptica da rede de atenção orientacional (Dehaene, 2020). Considerando a importância central que o sentido da visão tem para os seres humanos, basta averiguar para onde a cabeça e os olhos de um indivíduo estão apontando e ali, na vasta maioria das vezes, estará a atenção. Só não se pode perder de vista que há exceções, como quando é dito popularmente que a pessoa “vê, mas não olha”, ou, mais tecnicamente, os casos do piscar atencional, da cegueira desatencional, e quaisquer outros que, por limitações dos sistemas cognitivos ou por presença de metas discrepantes na intenção do indivíduo, o

pensamento que ocupa a consciência não corresponde àquilo que está no foco do campo visual concreto.

Feitas essas ressalvas, é possível concentrar-se sobre o mecanismo da visão operando em condições normais, sendo que são esses comportamentos que o *eye tracker* busca registrar e avaliar. O aspecto central é apresentado por Dehaene (2012): de toda a imagem que é recebida pela retina, apenas a parte central – na região chamada fóvea e que corresponde a cerca de 15 graus do campo visual – tem resolução alta suficiente para servir à leitura. Em função disso, os olhos têm que se mover muito e a todo momento e, sendo assim, de acordo com Riess e Gabriel (2020), os movimentos oculares fundamentais na leitura são as fixações e as sacadas, ou seja, quando os olhos estacionam para examinar uma posição e quando eles avançam para um novo ponto de parada. Rayner *et al.* (2016) asseguram que as fixações duram, em média, 250 milissegundos para leitores adultos experientes e, curiosamente, há indícios de que só uma pequena fração desse tempo é ocupada pela percepção visual pura, restando um tempo considerável para a execução de processamentos cognitivos e linguísticos de mais alta ordem. Segundo esses mesmos autores, as sacadas para a frente (*forward saccades*), tendo a leitura em língua inglesa como parâmetro, duram algo entre 20 e 30 milissegundos e abrangem a distância de sete letras, ou seja, geralmente saltam de uma palavra para a próxima. Rayner e colaboradores (2016) fazem uma pontuação interessante, dizendo que as sacadas para a frente, em contraste com o que ocorre com as fixações, têm sua variação de duração mais ligada à distância física a ser percorrida pelos olhos do que à complexidade do conteúdo.

Indo para além das fixações e das sacadas, Rayner *et al.* (2016) especificam alguns outros movimentos oculares. Os saltos (*skips*) são muito comuns e são feitos sobre palavras muito pequenas, frequentes ou previsíveis. As refixações (*refixations*) são paradas sobre a mesma palavra, algo comum para vocábulos longos, com mais de sete letras de comprimento. As regressões (*regressions*) são movimentos no sentido inverso do texto, para refixar uma palavra e melhor processá-la, algo que mesmo leitores experientes fazem de 10% a 15% do tempo (Rayner *et al.*, 2016, p. 9). E as varreduras de retorno (*return sweeps*) são os movimentos muito recorrentes que se passam quando se chega ao final de uma linha e move-se a visão para o início da próxima. Um útil resumo do que caracteriza um bom leitor (ou, pelo menos, um rápido) é oferecido por Rayner e colaboradores (2016, p. 13): “[...] fixações mais

curtas, sacadas mais longas e menos regressões [...]”⁹. Abrindo para outras possibilidades de análise de dados que podem ser captados pelo dispositivo de *eye tracking*, Riess e Gabriel (2020, p. 110) adicionam que se podem medir as piscadas para “[...] avaliar fadiga, sonolência ou sobrecarga do trabalho mental”.

É oportuno realizar uma breve comparação para colocar a leitura estrita – atividade visual altamente especializada – em perspectiva com a visualização de imagens em contextos mais gerais. Aumont (2014) entende que o campo visual útil da fóvea é de 30 graus (ou seja, o dobro da angulação que Dehaene (2012) concebe útil para a leitura), considerando que os estímulos que não são signos linguísticos verbais escritos podem ser perceptualmente muito mais simples. O outro tópico essencial trazido por Aumont (2014) é que, diferentemente da leitura de textos, que segue um padrão direcional completamente previsível (da esquerda para a direita e de cima para baixo no nosso sistema alfabético da língua portuguesa), o olhar percorre outras imagens seguindo rotas muito variáveis, geralmente em torno de várias fixações sucessivas, de décimos de segundo, nas partes da cena mais providas de informação, isto é, que permitam a memorização do que foi visto e seu futuro reconhecimento (Aumont, 2014, p. 59). O traçado do olhar, ainda de acordo com o autor, pode variar de forma significativa se o observador tiver outros conhecimentos de fundo (p. ex., um especialista em artes gráficas) ou outras intenções, até mesmo vindas por orientações explícitas externas, indicando o quê e como deve ser olhado.

Seja fazendo uso de tecnologias de rastreamento ocular ou não, a posição defendida por Sousa e Hübner (2015, p. 42) é que a forma mais segura de compensar as inescapáveis limitações dos testes que avaliam o entendimento do leitor é “[...] combinar mais de um método e obter diversas amostras do desempenho da compreensão leitora de cada participante”, desde que os testes sejam concordantes com a concepção de leitura adotada e com os objetivos da avaliação. Na mesma linha, Tomitch (2008, p. 48) sustenta que “[...] fica claro que o melhor procedimento é sempre buscar aliar diferentes abordagens para que consigamos obter dados os mais fidedignos possíveis aos processos investigados”. Dessa maneira, o entendimento mais acertado é que o procedimento metodológico

⁹ Do original: [...] *shorter fixations, longer saccades, and fewer regressions* [...] (Rayner et al., 2016, p. 13).

recomendável para estudar a compreensão leitora deve conter uma razoável diversificação nos dados coletados e analisados.

A leitura, destacando, fundamentalmente, a dimensão da compreensão leitora, é o objeto básico sobre o qual este estudo lançou mão para investigar as interferências da multimodalidade na recepção, assimilação, retenção e recuperação de informações contidas em textos narrativos curtos. A complexa atividade de ler mobiliza todos os elementos trabalhados nas seções deste capítulo de referencial teórico: atenção, consciência, memórias e processamento linguístico e semiótico. Questões específicas referentes ao sentido da visão e à decodificação do sistema de escrita (como o alfabético) são uma exclusividade da leitura em sentido estrito, mas todos os demais aspectos podem ser reconhecidos, igualmente, na audição da linguagem verbal codificada no meio sonoro.

Os desafios práticos para a avaliação da compreensão leitora e linguística serão perceptíveis no próximo capítulo, que descreve o método usado para a pesquisa experimental desta dissertação. Considerando, ainda, que nosso objetivo exigiu a inclusão de outros fatores, sobretudo a multimodalidade sensorial, uma série de cuidados adicionais foi requerida para a elaboração e condução dos testes, conforme será relatado na sequência.

3 METODOLOGIA

Encerramos o capítulo de referencial teórico discutindo o tópico mais específico da avaliação da compreensão leitora, o ponto central para o método de pesquisa desta dissertação. Como se pretendeu evidenciar pela sequência dos conceitos abordados, a compreensão de textos escritos é uma construção complexa, que depende de aspectos fundamentais da cognição – atenção, consciência, memórias – e de habilidades avançadas e eminentemente humanas – a linguagem verbal, seja na modalidade oral seja na escrita, transmitida culturalmente.

O principal diferencial da abordagem que aqui propomos é a consideração da multimodalidade sensorial, ou seja, a recepção de informações por mais de uma entrada perceptual, inserida no processo da compreensão, algo que tanto nos permite quanto nos exige pensar a linguagem e a leitura em um escopo mais amplo, no contexto das semioses ou produções e extrações de significados. Concentramo-nos sobre a bimodalidade, no par visão e audição, com o intuito de identificar ganhos ou perdas que o aumento na densidade informacional pode trazer à compreensão linguística.

3.1 Objetivos

3.1.2 Objetivo geral

Analisar possíveis interferências, desde a promoção de auxílio até a geração de obstáculo, provocadas a partir da associação, em um único texto, das modalidades sensoriais visual e auditiva à compreensão linguística de estudantes universitários.

3.1.3 Objetivos específicos

- 1: Comparar a compreensão linguística de um texto bimodal (visual e auditivo) com a de um texto unimodal.
- 2: Medir se a multimodalidade pode sobrecarregar a memória de trabalho do leitor.

- 3: Comparar a compreensão linguística na condição de múltiplos fluxos de informação codificados no mesmo canal sensorial (submodalidades) com as situações de unimodalidade simples e bimodalidade em canais sensoriais diferentes.
- 4: Identificar se informações bimodais não redundantes afetam negativamente a compreensão leitora.
- 5: Investigar se estímulos que surpreendem o leitor (como notificações sonoras ou visuais) afetam a compreensão leitora.

3.2 Hipóteses

- H1: Textos que acionam os modos visual e auditivo simultaneamente com informações redundantes terão índices de compreensão linguística melhor, porque o aproveitamento de um segundo canal sensorial otimiza o uso dos recursos cognitivos disponíveis para o processamento de um mesmo conteúdo-base. Apesar de esse ser o ponto que representa a maior controvérsia, estudos (Mousavi; Low; Sweller, 1995; Mayer; Moreno, 1998; Brasel; Gips, 2014) sugerem que há a possibilidade de melhorar a compreensão quando informações congruentes são dispostas em mais de uma modalidade.
- H2: Determinadas situações de multimodalidade sensorial, mesmo com informações congruentes, podem sobrecarregar a memória de trabalho do leitor, tendo como consequência uma compreensão linguística inferior. Os mesmos estudos citados na hipótese 1 identificaram casos em que pode haver um “congestionamento” já na entrada do canal sensorial ou, até mesmo, em um nível cerebral superior e mais integrativo, por exemplo, com legendas redundantes aumentando a complexidade percebida nas informações verbais, e não nas visuais (Brasel; Gips, 2014).
- H3: A codificação de mais de um fluxo de informação no mesmo modo sensorial (como é o caso do texto oral musicado, que usa apenas o canal da audição; ou vídeo somente com imagens e legendas, que ocupa unicamente o canal da visão) estará relacionada a um melhor desempenho de compreensão linguística do que a condição unimodal simples, mas pior do que a condição bimodal, já que a submodalidade adiciona um segundo fluxo

de informações coincidentes, o que é benéfico, mas não faz proveito de um segundo canal sensorial, que fica ocioso. Essa noção encontra amparo no estudo de Schön *et al.* (2008) e pode ser derivada da articulação entre as hipóteses 1 e 2.

- H4: A bimodalidade sensorial com informações não redundantes gerará índices inferiores de compreensão leitora do que as situações bimodais redundantes e unimodais, independentemente de estas utilizarem múltiplas submodalidades, dado que elementos incluídos não relacionados ao texto principal competirão pela atenção e pelo processamento consciente do indivíduo. A metanálise de Vasilev, Kirkby e Angele (2018), ao encontrar efeitos detrimenais consistentes – ainda que modestos – em todas as classes de distração auditiva, sustenta essa previsão.
- H5: Estímulos sonoros ou visuais que surpreendam o sujeito durante a tarefa de leitura, ou seja, desviem pontualmente a sua atenção, resultarão em prejuízo considerável à compreensão leitora, uma vez que ocuparão partes do processamento cognitivo do sujeito e causarão interrupções na atividade principal, que deverá ser retomada reiteradamente, aumentando a exigência sobre os sistemas de memória. Entende-se que essas perturbações são o pior cenário possível para a manutenção do foco na tarefa de compreensão textual e que os desvios afetarão a construção e a integração daquilo que Kintsch e Rawson (2005) definem como texto-base e modelo situacional.

3.3 Participantes

A população escolhida foi a de alunos universitários de Santa Cruz do Sul - RS, e a amostra consistiu em 25 sujeitos. A opção por alunos com esse grau de educação formal teve a intenção de selecionar leitores que provavelmente não teriam problemas em processos mais básicos de decodificação da escrita, permitindo analisar a compreensão leitora de forma mais isolada, com os dados estando menos sujeitos a ruídos.

De modo a tentar atingir uma maior homogeneidade na amostra, os critérios de exclusão definidos previram suprimir da pesquisa pessoas com língua materna diferente de português brasileiro, ou com algum tipo de condição cognitiva atípica,

ou com deficiências visuais ou auditivas não corrigidas, ou com idade fora da faixa compreendida entre 18 e 50 anos. Em compensação, não houve exclusões baseadas em quais cursos específicos os graduandos estavam matriculados: tentou-se chegar a um equilíbrio nesse sentido, principalmente, considerando o recorte por grandes áreas do conhecimento (Ciências Humanas, Exatas e Biológicas/Saúde).

O protocolo do estudo foi submetido ao Comitê de Ética em Pesquisa da Universidade de Santa Cruz do Sul em 16 de maio de 2025. Sua aprovação ocorreu em parecer emitido no dia 4 de junho de 2025. O Certificado de Apresentação para Apreciação Ética (CAAE) recebeu o seguinte número de identificação: 88460125.0.0000.5343. O Termo de Consentimento Livre e Esclarecido (TCLE), lido e assinado pelos participantes da pesquisa, pode ser visualizado no Apêndice A.

3.4 Instrumentos

Para que fosse possível testar as hipóteses propostas e atingir os objetivos que traçamos, foi utilizado um conjunto de instrumentos cujo cerne são os questionários para aferir a compreensão leitora. Juntamente com essas medidas off-line do produto final da compreensão dos textos (Tomitch, 2008), foi considerado o processo on-line, conforme registrado pelo dispositivo de rastreamento ocular. No sentido de complementar e contextualizar essas medidas, foram coletados dados sobre perfil e hábitos dos sujeitos, bem como a respeito de suas preferências percebidas a partir dos estímulos que viram e ouviram durante o teste principal.

3.4.1 Questionário de perfil do participante

Alguns dados preliminares foram coletados por um Questionário de Perfil do Participante (Apêndice B) para que se pudesse caracterizar a amostra de modo geral e ponderar fatores individuais que, possivelmente, influenciam os resultados nos testes de compreensão leitora. Esse instrumento foi elaborado por nós a partir de uma abordagem empírica, que levou em consideração os temas explorados no referencial teórico desta dissertação e as experiências práticas provenientes de outros estudos realizados pelo Grupo de Pesquisa Linguagem e Cognição (CNPq/UNISC), do qual fazemos parte. O Questionário foi construído usando-se a aplicação

Formulários Google e foi enviado para o endereço de e-mail de cada participante, uma informação que foi solicitada no TCLE (o espaço para ela foi incluído posteriormente, escrito à mão, porque sua necessidade somente foi percebida depois da impressão dos documentos).

Um primeiro grupo de perguntas se concentrou sobre tópicos mais comuns, como idade, gênero, curso e semestre que o indivíduo estava cursando na graduação. O outro conjunto tratou sobre preferências e comportamentos do sujeito: hábitos de leitura (frequência com que lê e quais tipos de materiais), consumo de mídias sonoras ou audiovisuais (rádio, *podcasts*, televisão, plataformas de *streaming* e redes sociais) e percepção subjetiva sobre aptidões ou dificuldades de aprendizagem (se entende melhor o conteúdo ouvindo ou lendo; se é preferível um ambiente de estudo ou trabalho que seja silencioso ou com determinados sons de fundo).

3.4.2 Textos para compreensão leitora

Esse instrumento, crucial para esta dissertação, é o que impôs os principais desafios para a sua elaboração satisfatória. Sousa e Hübner (2015), assim como identificam dificuldades para a confecção de testes de compreensão leitora, consideram uma fonte de incertezas a escolha do texto apropriado para as avaliações. Segundo as autoras, essa seleção requer que se tome cuidado com aspectos como leiturabilidade (complexidade textual), especificidades intrínsecas dos gêneros textuais, e temática, esta podendo interferir na motivação do leitor e também interagir de forma indevida e não controlada com o conhecimento prévio do sujeito. Fora isso, considerar um público adulto e de nível universitário no contexto brasileiro adiciona outros problemas. A revisão de literatura feita por Dias e colaboradores (2016, p. 126), que considerou o recorte 2009-2013 na produção científica nacional, encontrou uma carência de instrumentos para avaliação da leitura de adolescentes e adultos, limitado exclusivamente a testes de Cloze. O estudo de Carvalho e Souza (2023, p. 15), que considerou o intervalo dos anos 2014-2020, revela um quadro ainda similar, com uma minoria de estudos voltados aos leitores adultos proficientes.

A escassez de pesquisas contemporâneas sobre compreensão leitora com jovens e adultos com desenvolvimento cognitivo típico e que tiveram educação

formal regular – algo que também pode ser notado nos artigos internacionais (Carvalho; Souza, 2023) –, vem acompanhada de uma falta de opções de textos e testes (padronizados ou, pelo menos, usados com regularidade) para esse público-alvo. Uma opção à qual se poderia recorrer é a utilização de materiais presentes em grandes avaliações nacionais coordenadas pelo Governo Federal, como o Exame Nacional de Desempenho de Estudantes - ENADE ou o Exame Nacional do Ensino Médio - ENEM, mas, ao menos para a pesquisa que propomos, eles não foram considerados satisfatórios. A maior parte dos textos que compõem os enunciados das questões desses testes é muito direcionada a um conteúdo específico (como, por exemplo, História, Matemática ou Literatura) e, mesmo se fazendo uma seleção criteriosa dos poucos itens das provas que consideramos suficientemente universais, o conjunto que resta é bastante heterogêneo em termos de comprimento, complexidade, estilo e outros parâmetros.

Realizando uma varredura para averiguar os métodos usados na pesquisa nacional de compreensão leitora em adultos, foi possível encontrar alternativas, mas também inadequadas aos nossos objetivos. Para citar algumas abordagens, há a retirada de um texto diretamente de um jornal (Balzan, 2018), o que representa o perigo de favorecer o conhecimento de mundo de um subgrupo de participantes testados, e há a aplicação de um teste prático de leitura retirado do Indicador Nacional de Alfabetismo Funcional - INAF (Lustosa, 2014), que avalia processos de leitura bastante básicos em um conjunto muito variado de gêneros textuais, mas com apenas um exemplar de cada um desses. Uma estratégia aparentemente mais alinhada aos propósitos desta dissertação aparece em dois estudos (Alcará; Santos, 2015; Figueiredo *et al.*, 2016) que utilizaram o gênero crônica para os testes. O lado negativo é que ambos selecionaram um único texto (“Desentendimento”, de Luis Fernando Veríssimo, publicado em 1995) e avaliaram a compreensão leitora exclusivamente pela técnica de Cloze, dois aspectos que deixam o instrumento bastante restrito e pouco adaptável. Vale destacar, adicionalmente, que há estudos internacionais que utilizam testes padronizados para aferir a leitura de adultos, mas, além de não serem desenhados para o idioma português, eles são somente disponibilizados comercialmente, como é o caso do *York Assessment of Reading for Comprehension - YARC*, aplicado por Harvey e Walker (2018).

Consideradas todas essas dificuldades para a seleção de textos para avaliar a compreensão leitora de universitários, uma abordagem que se revelou um bom guia

para esta dissertação é a que foi utilizada por Daneman e Carpenter (1980), que construíram 24 pequenas passagens narrativas, com cerca de 150 palavras cada, para o teste específico do componente da compreensão. Desse conjunto de textos, apenas dois exemplos completos estão apresentados no artigo (Figura 6) e, lamentavelmente, os demais não puderam ser obtidos via pesquisas na internet ou mesmo contato por e-mail com as autoras¹⁰.

Figura 6 – Passagens narrativas e questionários de Daneman e Carpenter (1980)

SAMPLE PASSAGES AND QUESTIONS FROM THE COMPREHENSION TESTS

Teenager Passage (Distance 6)

Sitting with Richie, Archie, Walter and the rest of my gang in the Grill yesterday, I began to feel uneasy. Robbie had put a dime in the juke box. It was blaring one of the latest "Rock and Roll" favorites. I was studying, in horror, the reactions of my friends to the music. I was especially perturbed by the expression on my best friend's face. Wayne looked intense and was pounding the table furiously to the beat. Now, I like most of the things other teenage boys like. I like girls with soft blonde hair, girls with dark curly hair, in fact all girls. I like milkshakes, football games and beach parties. I like denim jeans, fancy T-shirts and sneakers. It is not that I dislike rock music but I think it is supposed to be fun and not taken too seriously. And here he was, "all shook up" and serious over the crazy music.

Questions

Pronoun	(1) Who was "all shook up" and serious over the music?
Fact	(2) Where was the gang sitting?
Fact	(3) Who put money in the juke box?"
Theme	(4) Provide a title for the passage that captures its theme."

Animal Passage (Distance 2)

It was midnight and the jungle was very still. Suddenly the cry of a wolf pierced the air. This anguished note was followed by a flurry of activity. All the beasts of the jungle recognized that an urgent meeting had been summoned by the lion, their king. Representatives from each species made rapid preparations to get to the river clearing. This was where all such emergency assemblies were held. The elephant and tiger were the first to arrive. Next came the gorilla, panther and snake. They were followed by the owl and the crocodile. The proceedings were delayed because the leopard had not shown up yet. There was much speculation as to the reasons for the midnight alarm. Finally he arrived and the meeting could commence.

Questions

Pronoun	(1) Who finally arrived?
Fact	(2) Where were these emergency meetings held?
Fact	(3) What broke the stillness of the night?"
Theme	(4) Provide a title for the passage that captures its theme."

" These questions were included only in Experiment 2.

Fonte: Daneman e Carpenter (1980, p. 455).

¹⁰ Enviamos e-mails para as Dras. Daneman e Carpenter no final de janeiro/2025, encaminhando-os aos endereços que fomos capazes de encontrar a partir de buscas na internet, explicando o contexto da pesquisa que pretendíamos desenvolver e solicitando o conjunto completo das passagens narrativas, informando que tínhamos o interesse em traduzi-las para o português e usá-las em nossos testes. Não recebemos retorno da Dra. Carpenter, e a Dra. Daneman nos respondeu esclarecendo que se encontrava aposentada e que já não mais possuía os materiais relativos ao artigo publicado em 1980.

Usar passagens narrativas curtas, tal como fizeram Daneman e Carpenter (1980), parece ser uma opção segura porque permite que se mantenha o conhecimento de mundo do leitor tão fora quanto possível dos testes, já que toda história traz elementos e estruturas desconhecidas até o momento da leitura. O uso de crônicas, considerando a característica do gênero textual, poderia ser adequado, mas há outros problemas de ordem prática: crônicas tão breves – com cerca de 150 palavras – não existem em grande quantidade, a ponto de que se pudesse descartar a maioria dessas por, digamos, exigirem níveis excessivamente altos de interpretação, e, se fossem de autores consagrados, como Luis Fernando Veríssimo, corria-se um grande risco de alguns leitores já conhecerem previamente os textos.

A decisão tomada, por fim, foi a de usar textos gerados por inteligência artificial, testados em estudos-piloto a serem detalhados na sequência. Essa tecnologia parece resolver um difícil paradoxo, escrevendo textos que são bons o bastante para serem considerados aceitáveis e corretos e, ao mesmo tempo, que são suficientemente descartáveis, sem marcas estilísticas acentuadas ou originalidade aparente. É um texto verdadeiramente médio ou, como colocam Santaella e Kaufman (2024), um sequenciamento razoável de palavras estatisticamente mais prováveis. Além disso, o recurso tecnológico da inteligência artificial facilita a obtenção de textos completos dentro do comprimento estipulado pelo usuário e, algo muito importante, constrói redações inéditas.

Os textos para compreensão leitora (Apêndice C) foram produzidos pelo software “ChatGPT”, versão GPT-4o (*Generative Pre-trained Transformer 4 Omni*), disponível gratuitamente para usuários cadastrados – com uma limitação diária de usos do seu algoritmo mais avançado. O *prompt* inicial que escrevemos pediu a criação de passagens narrativas que usassem os textos “Teenager Passage” e “Animal Passage” (Daneman; Carpenter, 1980, p. 455) como referência, solicitando também que os textos tivessem uma extensão entre 140 e 160 palavras. Utilizamos uma série de outros *prompts* para ajustar a escrita do *ChatGPT* aos nossos objetivos, pedindo: quatro personagens por história (a fim de apresentar boa quantidade de elementos para serem processados e memorizados pelo leitor), temáticas variadas, minimização de clichês (p. ex., nas iterações iniciais, todos os contos terminavam com suspense) e mais detalhamento dos cenários e personagens. Esse processo foi feito de maneira que houvesse um acervo de dezenas de textos, dos

quais selecionamos as seguintes nove passagens narrativas, conforme intituladas pelo software de inteligência artificial generativa (OpenAI, 2025a, 2025b):

- 1: A Pista de Patinação
- 2: O Café da Esquina
- 3: O Observatório Abandonado
- 4: O Jogo de Cartas
- 5: O Casarão da Rua Augusta
- 6: O Farol de São Tomé
- 7: O Festival de Balonismo
- 8: O Treino no Parque
- 9: A Loja de Discos

O número final de textos foi determinado pela quantidade de condições experimentais que planejamos: como decidimos pela apresentação de nove formatos diferentes (detalhados na subseção seguinte) para cada participante, foi necessário elaborar uma quantidade igual de narrativas, a fim de que os sujeitos não lessem histórias repetidas, que provavelmente seriam mais bem memorizadas do que as que tivessem visualização única, enviesando os testes. É importante destacar que fizemos pequenas edições nas histórias, praticamente todas limitadas à última frase de cada texto (embora algumas tenham tido que invadir outras das frases da parte final da narrativa, justamente para viabilizar que a conclusão da passagem fosse adequada), para que fosse possível montar cada teste com uma pergunta de referência pronominal (ver subseção 3.4.3).

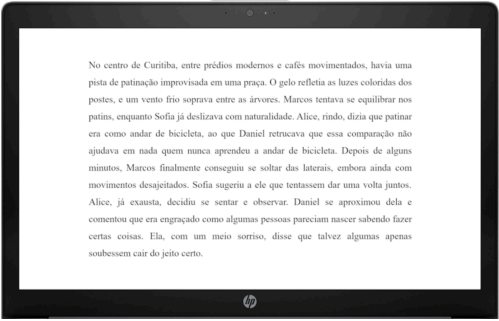
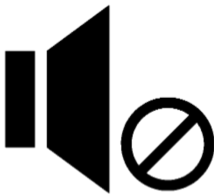


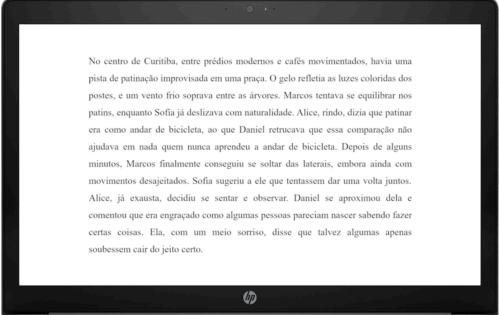



3.4.2.1 Modalidades e versões de apresentação dos textos

Cada uma das passagens narrativas para compreensão leitora foi adaptada para todas estas versões de apresentação, totalizando 81 estímulos (9 multiplicado por 9):

- A: apenas texto escrito em tela;
- B: apenas texto falado em áudio;
- C: texto escrito + texto falado;
- D: texto escrito (legendas) + imagens redundantes;
- E: texto falado + imagens redundantes;
- F: apenas texto cantado (em forma de música);

- G: texto escrito + música instrumental;
- H: texto escrito + sons de notificação (perturbações sonoras intermitentes);
- I: texto escrito + janelas *pop-up* (perturbações visuais intermitentes).

Figura 7 – Formatos de exibição de texto por condição experimental

Identificação	Conteúdo visual	Conteúdo auditivo
<p>A</p>	 <p>No centro de Curitiba, entre prédios modernos e cafés movimentados, havia uma pista de patinação improvisada em uma praça. O gelo refletia as luzes coloridas dos postes, e um vento frio soprava entre as árvores. Marcos tentava se equilibrar nos patins, enquanto Sofia já deslizava com naturalidade. Alice, rindo, dizia que patinar era como andar de bicicleta, ao que Daniel retrucava que essa comparação não ajudava em nada quem nunca aprendeu a andar de bicicleta. Depois de alguns minutos, Marcos finalmente conseguiu se soltar das laterais, embora ainda com movimentos desajeitados. Sofia sugeriu a ele que tentassem dar uma volta juntos. Alice, já exausta, decidiu se sentar e observar. Daniel se aproximou dela e comentou que era engraçado como algumas pessoas pareciam nascer sabendo fazer certas coisas. Ela, com um meio sorriso, disse que talvez algumas apenas soubessem cair do jeito certo.</p>	
<p>B</p>		
<p>C</p>	 <p>No centro de Curitiba, entre prédios modernos e cafés movimentados, havia uma pista de patinação improvisada em uma praça. O gelo refletia as luzes coloridas dos postes, e um vento frio soprava entre as árvores. Marcos tentava se equilibrar nos patins, enquanto Sofia já deslizava com naturalidade. Alice, rindo, dizia que patinar era como andar de bicicleta, ao que Daniel retrucava que essa comparação não ajudava em nada quem nunca aprendeu a andar de bicicleta. Depois de alguns minutos, Marcos finalmente conseguiu se soltar das laterais, embora ainda com movimentos desajeitados. Sofia sugeriu a ele que tentassem dar uma volta juntos. Alice, já exausta, decidiu se sentar e observar. Daniel se aproximou dela e comentou que era engraçado como algumas pessoas pareciam nascer sabendo fazer certas coisas. Ela, com um meio sorriso, disse que talvez algumas apenas soubessem cair do jeito certo.</p>	
<p>D</p>	 <p>Marcos tentava se equilibrar nos patins.</p>	

E



F



G

No centro de Curitiba, entre prédios modernos e cafés movimentados, havia uma pista de patinação improvisada em uma praça. O gelo refletia as luzes coloridas dos postes, e um vento frio soprava entre as árvores. Marcos tentava se equilibrar nos patins, enquanto Sofia já deslizava com naturalidade. Alice, rindo, dizia que patinar era como andar de bicicleta, ao que Daniel retrucava que essa comparação não ajudava em nada quem nunca aprendeu a andar de bicicleta. Depois de alguns minutos, Marcos finalmente conseguiu se soltar das laterais, embora ainda com movimentos desajeitados. Sofia sugeriu a ele que tentassem dar uma volta juntos. Alice, já exausta, decidiu se sentar e observar. Daniel se aproximou dela e comentou que era engraçado como algumas pessoas pareciam nascer sabendo fazer certas coisas. Ela, com um meio sorriso, disse que talvez algumas apenas soubessem cair do jeito certo.



H

No centro de Curitiba, entre prédios modernos e cafés movimentados, havia uma pista de patinação improvisada em uma praça. O gelo refletia as luzes coloridas dos postes, e um vento frio soprava entre as árvores. Marcos tentava se equilibrar nos patins, enquanto Sofia já deslizava com naturalidade. Alice, rindo, dizia que patinar era como andar de bicicleta, ao que Daniel retrucava que essa comparação não ajudava em nada quem nunca aprendeu a andar de bicicleta. Depois de alguns minutos, Marcos finalmente conseguiu se soltar das laterais, embora ainda com movimentos desajeitados. Sofia sugeriu a ele que tentassem dar uma volta juntos. Alice, já exausta, decidiu se sentar e observar. Daniel se aproximou dela e comentou que era engraçado como algumas pessoas pareciam nascer sabendo fazer certas coisas. Ela, com um meio sorriso, disse que talvez algumas apenas soubessem cair do jeito certo.



I

No centro de Curitiba, entre prédios modernos e cafés movimentados, havia uma pista de patinação improvisada em uma praça. O gelo refletia as luzes coloridas dos postes, e um vento frio soprava entre as árvores. Marcos tentava se equilibrar nos patins, enquanto Sofia já deslizava com naturalidade. Alice, rindo, dizia que patinar era como andar de bicicleta, ao que Daniel retrucava que essa comparação não ajudava em nada quem nunca aprendeu a andar de bicicleta. Depois de alguns minutos, Marcos finalmente conseguiu se soltar das laterais, embora ainda com movimentos desajeitados. Sofia sugeriu a ele que tentassem dar uma volta juntos. Alice, já exausta, decidiu se sentar e observar. Daniel se aproximou dela e comentou que era engraçado como algumas pessoas pareciam nascer sabendo fazer certas coisas. Ela, com um meio sorriso, disse que talvez algumas apenas soubessem cair do jeito certo.



A Figura 7¹¹ esquematiza os conteúdos em vídeo e áudio incluídos em cada condição experimental e serve como material de consulta rápida para pontos futuros desta dissertação, já que a identificação por letras será mencionada com muita frequência. No intuito de ajudar com a memorização dessas categorias – fazendo, inclusive, um uso prático do processo de *chunking* que foi discutido na subseção 2.2.5 –, pode-se pensar em quatro grupos: o primeiro é o dos básicos (A, B e C), com leitura silenciosa, fala comum e a soma dos dois; o segundo é o dos vídeos (D e E), um com legenda e o outro com fala; o terceiro é o da música (F e G), com música cantada em um e apenas música instrumental no outro; e o quarto é o das notificações (H e I): sonoras no primeiro e visuais no segundo.

Foram empregadas outras ferramentas de inteligência artificial generativa para a produção das diferentes modalidades de informação, mantendo a ideia de homogeneidade dos conteúdos, excluindo, tanto quanto possível, discrepâncias que pudessem ser introduzidas pela criatividade e pelos diferentes pontos de vista humanos. Os textos narrados em áudio foram produzidos pela ferramenta “Narakeet”¹², cuja versão grátis pudemos usar porque conseguimos nos manter abaixo dos limites de tempo e de número de arquivos que exigiriam uma conta paga. O narrador escolhido dentro da plataforma foi o “Murilo”, no grupo “Portuguese - Brazilian”. Apesar de a fala gerada ser muito satisfatória até mesmo em termos de sotaque, algumas palavras tiveram que ser adaptadas e reescritas no *prompt* do aplicativo para que soassem corretamente: “entendêra”, “côco”, “desaparecêssem”, “cêstos”, “éle-pês”, “ouvíra” e “só-zinhos”.

Já os textos em música cantada foram produzidos pelo programa “Suno”¹³. Utilizamos a versão mais atualizada disponível que não requeria pagamento, tendo como limite o número de 20 músicas ao todo, sendo que cada *prompt*, por padrão, gerava duas canções diferentes com o mesmo texto. No geral, a plataforma produziu ótimos resultados em termos da adequação fonética ao português brasileiro, mas, provavelmente por particularidades do material usado para treinar o algoritmo, algumas frases ou palavras pareciam favorecer a saída no sotaque de Portugal (“parapeito” foi um desses termos). Uma vez baixada toda a coletânea produzida

¹¹ As figuras, os quadros e as tabelas apresentados ao longo desta dissertação – a não ser quando expressamente informado de maneira diferente – foram elaborados pelo autor a partir de dados coletados na pesquisa experimental.

¹² Disponível em: <https://www.narakeet.com/>; versão 2.43.66.

¹³ Disponível em: <https://suno.com/home>; versão 4.

pelo “Suno”, o autor desta dissertação precisou fazer algumas edições de áudio em um programa próprio para a tarefa. Os cortes feitos atacaram os seguintes problemas: inícios e finais de canções muito alongados (não queríamos desperdiçar o tempo dos participantes e nem lidar com um maior decaimento de suas memórias), longas seções instrumentais no meio das músicas, repetições indevidas do texto original, e palavras com pronúncia incorreta (como “estrava” e “capriturava”).

As montagens das imagens redundantes – compostas por pequenos trechos de vídeo e fotografias congruentes com o texto básico – foram produzidas pelo software “Invideo AI”¹⁴, e foi o processo generativo que mais exigiu reiterações de *prompts* e ajustes feitos manualmente. Como a versão gratuita dessa ferramenta apenas permitiu que a montagem de vídeo fosse feita a partir de bancos de imagens existentes (sem contar que incluiu “marcas d’água”, alusivas ao aplicativo e aos acervos de imagens, que não puderam ser removidas, tornando-se elementos visuais adicionais indesejados), foi preciso complementar com ilustrações geradas pela inteligência artificial do *ChatGPT*, o mesmo sistema utilizado para elaborar os textos-base. Esses desenhos foram inseridos para tornar possível que cada personagem da história fosse apresentado com coerência – as cenas selecionadas automaticamente pela aplicação “Invideo AI”, apesar de, no geral, serem concordantes com as ações do texto, apresentam novos atores/modelos em praticamente cada um dos segmentos. Sendo assim, as ilustrações criadas pelo *ChatGPT* foram agregadas manualmente, em software para edição de vídeo e pelo autor desta dissertação, aos vídeos-base gerados pelo “Invideo AI”. Esses esqueletos principais dos vídeos foram construídos partindo das configurações mais neutras que a aplicação nos ofereceu (“clean”, “bright”, “travel enthusiasts”, entre outros), sempre determinando que usasse exclusivamente o texto das narrativas, o que, por motivos que não conhecemos, muitas vezes não foi atendido e nos exigiu reiterações. Aproveitando que o sistema do “Invideo AI” permitia que se alterassem as imagens/vídeos inseridos por ele no corte geral, substituímos alguns trechos cuja ideia o algoritmo não foi capaz de interpretar adequadamente (p. ex., em momentos nos quais incluiu cenas de discussões acaloradas, quando o que o texto requisitava era apenas a representação de uma conversa amigável).

A música instrumental usada como fundo para a condição G foi a mesma para todos os nove textos. A obra escolhida foi “Inside Out”, de Sleepermane e Sling Dilly

¹⁴ Disponível em: <https://invideo.io/>; versão 3.0.

(Inside Out, 2021), uma das muitas músicas que compõem a popular *playlist* voltada para o relaxamento e para o acompanhamento de atividades de estudo “lofi hip hop radio – beats to relax/study to” (Lofi Girl, 2022), uma lista de reprodução ininterrupta do canal do YouTube “Lofi Girl”, que tem 14,9 milhões de inscritos¹⁵.

A versão de texto H contém uma trilha de áudio composta por um simples *bip* sonoro (usado para notificar a chegada de uma mensagem nova ao celular; extraído de um vídeo do YouTube que apresentava uma coleção de sons reproduzidos por múltiplos modelos de telefone) repetido em intervalos irregulares de tempo: há períodos em que os sons estão mais agrupados, mas, no geral, existe um espaçamento suficiente para evitar que o ouvinte se habituasse com muita facilidade a essa interrupção. O autor desta dissertação montou essa faixa de áudio em um programa apto para essa manipulação, buscando emular intervalos realísticos de notificações de mensagens de celular recebidas. Seguindo a idêntica ideia de padrão de distribuição ao longo do tempo que acabamos de relatar, a condição I exibe uma janela *pop-up* básica, de cor cinza, com alguma transparência (para permitir a leitura do trecho de texto sobre o qual ela surge) e sem qualquer elemento verbal ou icônico, no canto inferior direito da tela em que é feita a leitura (conforme mostrado na tela I da Figura 7).

Os modos de apresentação que contêm texto estático escrito (A, C, G, H e I) ocupam uma única tela inteira, com fonte *Times New Roman*, no tamanho 40 pixels, na cor preta sobre fundo branco. Cada linha de texto foi mantida dentro do número de caracteres e espaços comum de ser encontrado em livros impressos (por volta de 70, ao todo). A condição D, que é montada com legendas sobre vídeo, apresenta o texto em padrões similares aos que podem ser vistos em filmes legendados, com fonte *Open Sans*, tamanho 60 pixels, amarela com contorno preto. Seguindo o mesmo padrão da mídia citada, essa legenda mais convencional é exibida em blocos maiores e mais duradouros, em oposição ao estilo extremamente veloz e fragmentado que se vê em textos incluídos em conteúdos comumente visualizados em redes sociais como *Instagram* e *TikTok*. É importante destacar que sequenciamos essas legendas de modo que elas acompanhassem perfeitamente as narrações em áudio usadas nas condições B, C e E, mesmo que essas duas representações dos textos não coexistam em qualquer um dos estímulos que confeccionamos.

¹⁵ Números verificados na data de 30 de abril de 2025.

3.4.3 Questionários de compreensão leitora

Os questionários de compreensão leitora foram desenhados especificamente para cada uma das nove narrativas (textos 1 a 9) e não variam de acordo com a modalidade de apresentação: ou seja, o texto 1, fosse lido e/ou ouvido, foi sucedido pelo exato mesmo grupo de perguntas, que são distintas das feitas para o texto 2. Esse instrumento foi montado tomando por base a metodologia de Daneman e Carpenter (1980), portanto, fazendo quatro questões abertas (conforme classificação de Sousa; Hübner, 2015) ao término de cada passagem narrativa. A primeira delas se refere à frase final do texto, buscando que o sujeito respondesse qual o referente correto é recuperado por um pronome ali presente, exigindo um nível um pouco mais elevado de inferência do leitor. Os dois itens seguintes endereçam formas mais simples de memória, referindo-se a fatos e personagens explicitamente apresentados na história. A quarta pergunta procura verificar a compreensão global do texto por meio da solicitação da escrita de um título que resuma a ideia principal daquilo que foi lido. Adicionalmente, optamos por incluir uma quinta questão aberta, que simplesmente pede para o sujeito relatar qualquer parte da narrativa que tenha lhe chamado a atenção. Esse acréscimo teve o propósito de apresentar uma pergunta fácil (pelo nosso entendimento, uma resposta totalmente errada ou não aproveitável para esse item só poderia vir de uma leitura extremamente equivocada ou desinteressada) que pudesse coletar dados informativos mesmo de sujeitos que tivessem dificuldades com as outras questões e que, ainda, pudesse manter a motivação do participante em um nível mais alto. No final das contas, como os participantes sustentaram um bom nível de acertos em todos os questionários, a inclusão dessa última pergunta se revelou não tão necessária. Independentemente disso, considerando que ela agregou mais dados à pesquisa, foi uma adição benéfica ao estudo.

3.4.4 Dispositivo de rastreamento ocular (*eye tracker*)

A coleta de dados de rastreamento ocular foi realizada com o *eye tracker* “Tobii Pro Fusion (250 Hz)”, acoplado, por meio de um ímã próprio para a sua fixação correta, a um notebook “HP ProBook 470 G5” (com sistema operacional “Windows 11 Pro” e tela com 17.3 polegadas de diagonal e operando na resolução 1920 x 1080

pixels)¹⁶. O computador possuía instalado o software de análise “Tobii Pro Lab”, no qual as baterias de teste foram desenhadas e configuradas, os experimentos foram executados e gravados, e a partir do qual as múltiplas métricas referentes aos movimentos oculares dos participantes foram extraídas para análise posterior.

O equipamento de *eye tracker* foi cedido pela Université Libre de Bruxelles, Bélgica, graças à colaboração existente entre os grupos de pesquisa liderados pelas professoras Régine Kolinsky e Rosângela Gabriel. O teclado presente no notebook seguia o padrão “AZERTY”, um *layout* muito usado em países francófonos e, por isso, necessitou uma adaptação; caso contrário, causaria muita confusão para os participantes (brasileiros natos e completamente habituados ao padrão “QWERTY”) digitarem as suas respostas, o que afetaria, em maior ou menor medida, os resultados, fosse irritando os sujeitos ou os desconcentrando e afetando sua memória. A solução encontrada foi acoplar ao kit utilizado em todos os testes um teclado externo, que ficou posicionado de forma segura sobre a base do notebook com a ajuda de alguns recortes de placas de EVA (Etileno Acetato de Vinila; um material emborrachado). Outro cuidado relacionado que tivemos que tomar foi a instalação no sistema Windows do pacote de idioma Português-BR e o ajuste do fuso horário, para garantir que as gravações feitas pelo *eye tracker* estivessem com seus dados cronológicos corretos, facilitando futuras consultas.

O programa “Tobii Pro Lab” e o computador em que estava instalado foram o ponto central de todas as baterias de teste. Todos os estímulos de imagem e vídeo tiveram que ser gerados em formatos de arquivos compatíveis com o software (“.JPG”, “.PNG” e “.MP4”) e sequenciados dentro das *timelines* (uma para cada participante), estipulando quais materiais avançariam ao comando (clique do mouse ou pressionamento de qualquer tecla do teclado) do participante. Os questionários, todos virtuais, gerados na plataforma Formulários Google, foram completamente integrados ao sistema em forma de “estímulos on-line”, que permitiam a inserção de um link a ser acessado dentro da progressão normal da bateria. Todo o material visual foi exibido pelo monitor padrão do computador, sempre em tela cheia, sem quaisquer distrações não controladas. O brilho do *display* foi sempre ajustado para o valor 90, de acordo com a escala usada pelo Windows 11. Para a parte de áudio, foi conectado, com fio, um fone de ouvido circum-auricular fechado Edifier W600BT,

¹⁶ Agradecemos muito à Profa. Dra. Régine Kolinsky por ter viabilizado o empréstimo desse equipamento, patrimônio da Université Libre de Bruxelles (ULB).

mantido no volume 26 do sistema operacional do notebook. Esse item assegurou uma boa e padronizada qualidade para a audição dos materiais sonoros e, pelo seu desenho, ofereceu um razoável nível de isolamento acústico ao sujeito.

3.4.5 Questionário de percepção sobre o teste de compreensão leitora

Imediatamente após o participante finalizar suas respostas ao nono, e, portanto, último questionário de compreensão leitora, abria-se um novo instrumento, no mesmo sistema em que foram visualizados os textos e respondidos os testes de compreensão, para coletar as percepções subjetivas do leitor (Apêndice D). Foi composto por duas telas: a primeira continha campos para o indivíduo escrever texto livre e a segunda foi configurada para a marcação de múltiplas escolhas. Começou-se perguntando qual foi a história preferida, de modo que o participante pudesse redigir quaisquer referências de que lembrasse, que pudessem indicar sua passagem narrativa favorita e, logo abaixo, seguia-se para uma pergunta quase idêntica, mas agora solicitando que se informasse qual foi a história de que menos gostou. A segunda seção pediu que o sujeito marcasse as duas versões de texto que considerou mais fáceis de acompanhar e compreender (p. ex., escrito, falado, ou vídeo com texto legendado) e, na sequência, que fossem marcados os dois formatos de texto mais difíceis.

3.5 Procedimentos de produção dos dados

Antes do início das coletas de dados que vieram a compor os resultados desta pesquisa, realizamos três tipos de estudo-piloto para testar os instrumentos, procurando falhas e pontos que poderiam ser aprimorados. O primeiro e mais importante foi conduzido no Laboratório de Pesquisa em Leitura do PPGL UNISC em que posteriormente foram realizados os experimentos regulares, contando com a colaboração de dois bolsistas de Iniciação Científica e uma doutoranda, todos membros do Grupo de Pesquisa Linguagem e Cognição. Cada um deles passou por uma bateria de testes completa e relatou as dificuldades encontradas. A partir de suas contribuições, foram feitos os seguintes ajustes: a inclusão de intervalos de tempo menores no Questionário de perfil (em vez de prever as opções “mais de 1h-6h-12h”, ajustamos para “mais de 1h-3h-6h-9h”, pois pareceram mais

proveitosas) e a revisão dos volumes das mídias feitas no formato F, que estavam muito altas na comparação com os outros estímulos auditivos.

Uma segunda variedade de estudo-piloto foi realizada mais no sentido de divulgação da pesquisa e demonstração do dispositivo de rastreamento ocular. Foram visitadas duas aulas de estudantes de Letras, uma de Licenciatura e outra de Pós-Graduação, nas quais um voluntário ficou posicionado como os participantes usuais, em frente à tela do computador e do *eye tracker*, e o restante da turma pôde acompanhar o teste em uma tela grande, cuja projeção duplicava o conteúdo que rodava no computador. As respostas aos questionários foram deliberadas em grupo, algo que permitiu observar o quanto cada indivíduo difere no que diz respeito a itens a que presta mais atenção e que tende a armazenar mais facilmente na memória.

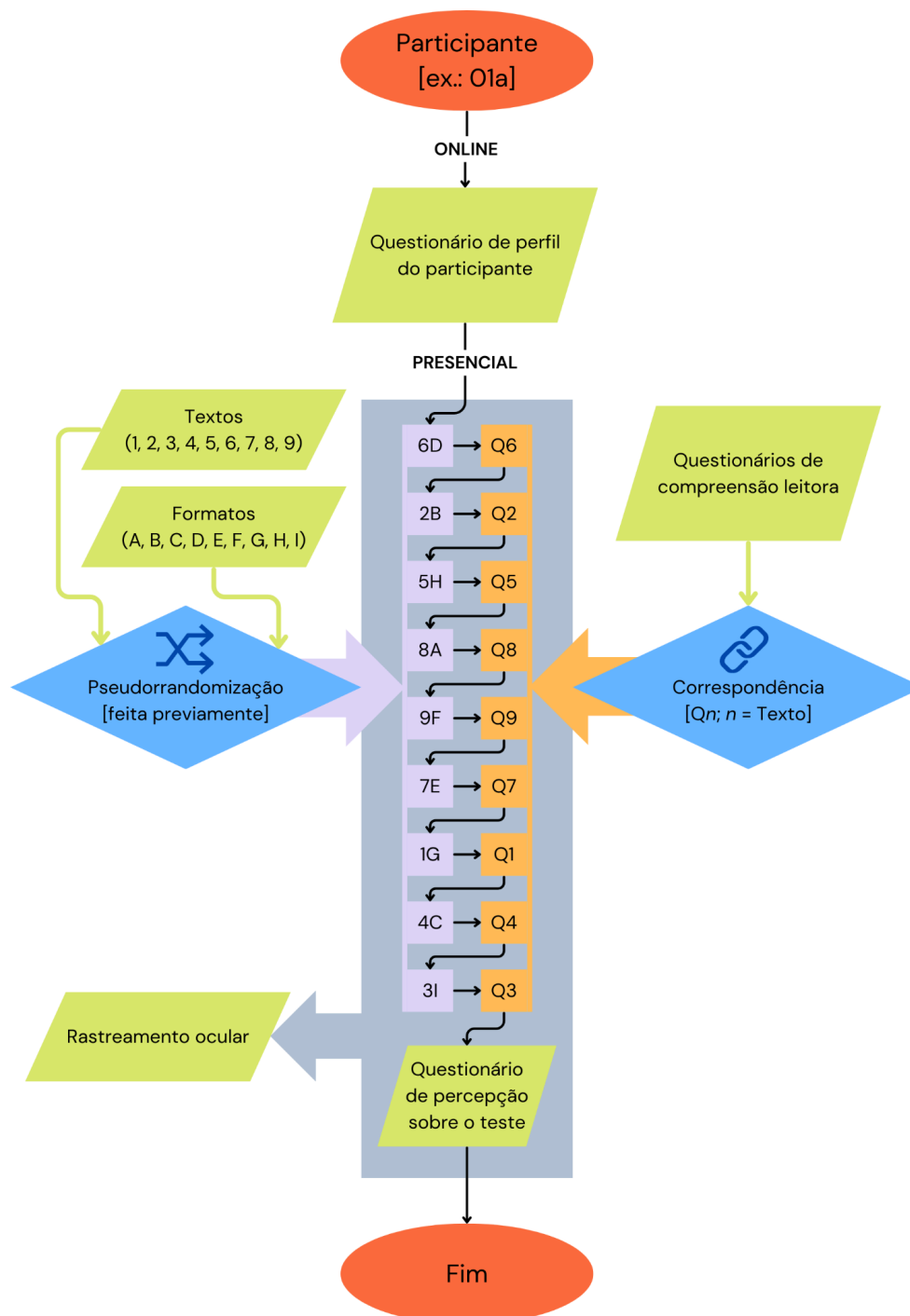
O terceiro piloto consistiu na elaboração de um formulário na plataforma Google, que permitia que o voluntário pudesse responder aos mesmos nove questionários de compreensão leitora mas, dessa vez, com a possibilidade de consultar os textos (todos apenas escritos, análogos à condição experimental A) a qualquer momento durante a fase da escrita das respostas. Houve três participantes para esse ensaio e, de modo geral, verificou-se um desempenho quase perfeito deles. As perguntas factuais (R2 e R3), com exceção de uma resposta incompleta (“cerveja”, quando se esperava “cerveja e frituras”), teriam todas a nota máxima. Na pergunta de resumo em forma de título (R4), ainda que o aproveitamento tenha sido próximo do máximo, ocorreram algumas respostas aparentemente apressadas (“No bar”), o que talvez se explique por um menor nível de engajamento do voluntário, também sugerido pela resposta “Sem fatos” na questão de reconto de qualquer fato da história (R5). A questão referencial (R1) registrou dois erros que valem ser comentados. Um deles aconteceu no texto 2, quando foi colocado o nome da outra personagem feminina que também era mencionada próximo do final da história – destaca-se que o voluntário que deu essa resposta também achou necessário justificar expressamente a sua escolha, sugerindo uma ambiguidade. O outro erro, bastante inusitado, foi no texto 3, para o qual houve a resposta “Clara” onde se esperava “Caio”, sendo que a frase final do texto especifica o pronome no gênero masculino (“[...] *e*le foi ajudá-la com o ajuste da câmera”). Os textos não foram alterados para os testes oficiais, mas nos mantivemos mais atentos para verificar se existiria a tendência à repetição desses padrões de erros, o que não se confirmou.

Avançando para os procedimentos do experimento formal, o convite a potenciais participantes foi feito com a ajuda de três professoras da universidade selecionada. Duas delas abriram um espaço na parte inicial de uma de suas aulas para que o autor desta dissertação explicasse a pesquisa e todo o seu contexto aos alunos. A terceira professora fez o recrutamento diretamente entre bolsistas ligados ao seu curso, já agendando horários para a coleta de dados. Para todos aqueles que tiveram interesse em fazer parte do estudo, foi entregue o TCLE para que o assinassem e informassem seu endereço de e-mail, para o qual foi enviado, em seguida, o Questionário de perfil do participante, a ser respondido pelos sujeitos em um momento e um local que julgassem oportunos (a Figura 8 apresenta a sequência da pesquisa experimental em um esquema gráfico).

Nosso plano inicial de conseguir 20 participantes foi, felizmente, superado, de modo que chegamos a 25. Houve, inclusive, alunos que assinaram o TCLE e completaram o Questionário de perfil, mas, como não tivemos tempo hábil¹⁷ para realizar o teste completo com todos os interessados, apenas foram aproveitados os dados daqueles que cumpriram todas as etapas. Como mostraremos na Tabela 1, presente na seção 4.1, apesar de não termos atingido um equilíbrio perfeito entre as três grandes áreas de estudo dos graduandos (Humanas, Exatas e Biológicas/Saúde), conseguimos uma boa distribuição nesse sentido. Quanto às exclusões, apenas tivemos que desconsiderar os dados de um aluno de doutorado, já que estava além do grau de educação formal que estipulamos para a nossa amostra. Participantes que informaram ter TDAH (transtorno do déficit de atenção com hiperatividade) sendo controlado com medicação ou que possuíam baixa visão em apenas um de seus olhos foram mantidos normalmente na pesquisa.

¹⁷ Entre a aprovação do protocolo de pesquisa pelo CEP-Unisc e a data em que foi feita a devolução do equipamento de rastreamento ocular à universidade belga (ULB) que nos concedeu o empréstimo, tivemos pouco mais de um mês. O período também coincidiu com o final do semestre de aulas dos nossos participantes, tornando suas agendas um pouco mais ocupadas, com provas e trabalhos finais.

Figura 8 – Esquema da pesquisa experimental

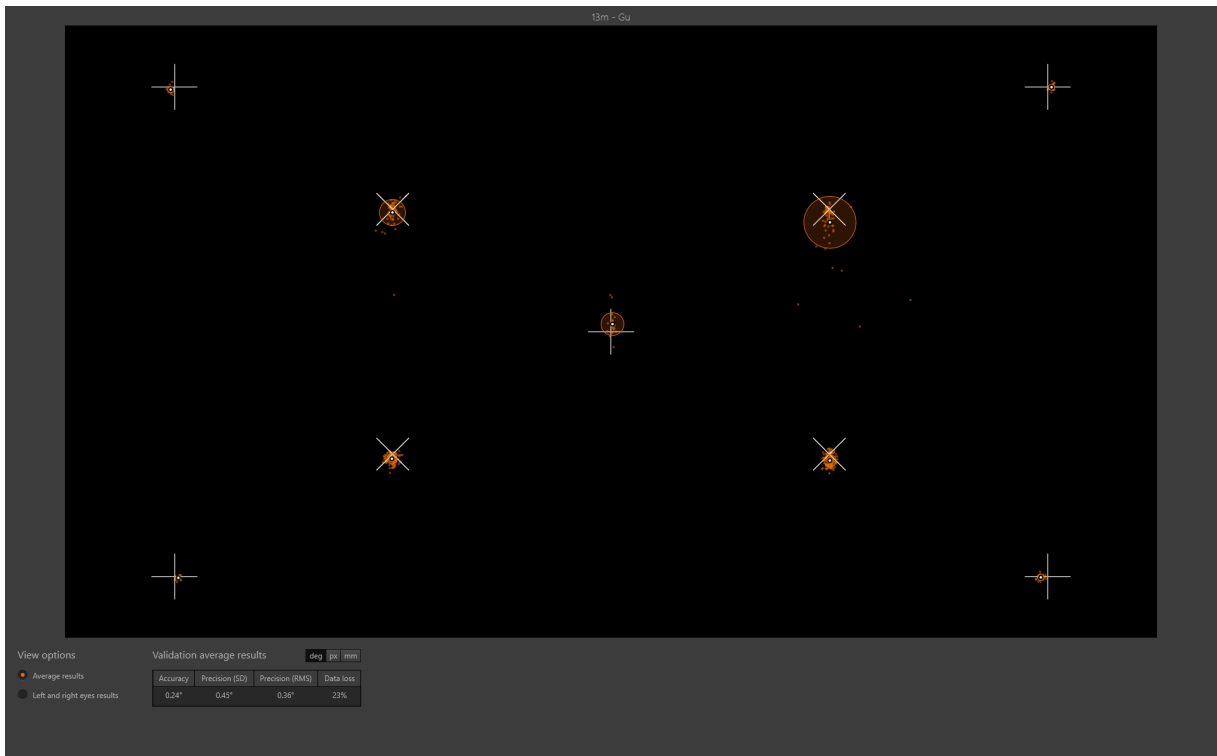


A porção principal do experimento, que englobou os textos e questionários de compreensão leitora, bem como o questionário de percepção sobre estes, foi conduzida em três locais da Universidade de Santa Cruz do Sul: o Laboratório de Pesquisa em Leitura do Programa de Pós-Graduação em Letras (sala 1024; nove participantes), uma sala de aula (5317; seis participantes) e um escritório de um

professor (sem numeração, vizinho à sala 5314A; dez participantes). Pretendia-se, originalmente, ocupar um único lugar padronizado, mas se tornou mais viável proceder da forma como fizemos, porque garantiu um processo mais conveniente e ágil para os voluntários, assegurando um número maior de coletas. Os experimentos foram realizados à tarde ou à noite e, embora nenhuma das salas usadas tenha estado em completo silêncio durante a totalidade dos testes, foi possível manter os ambientes suficientemente livres de distrações.

No momento em que o participante chegava à sala onde seria aplicado o seu teste, o autor desta dissertação o encaminhava até a cadeira em que deveria ficar sentado, diante do notebook e, por conseguinte, do dispositivo de rastreamento ocular. Em seguida, eram dadas as instruções básicas do teste, apresentadas em um tablet (conforme pode ser visualizado no Apêndice E) ao mesmo tempo em que eram faladas, de forma livre, pelo pesquisador – como a bateria de testes já era muito densa em conteúdo, achou-se mais sensato tornar mais leve essa introdução, favorecendo, além disso, o diálogo com o sujeito para tirar suas dúvidas. Passada essa etapa, o participante era solicitado a colocar os fones de ouvido e era iniciado o processo de calibração do *eye tracker*, o que consistiu em o indivíduo seguir com o olhar uma série de pontos que se moviam sistematicamente pela tela. Nossa meta era manter as medidas de acurácia e precisão, calculadas pelo “Tobii Pro Lab”, abaixo de 0,5 graus, o que nem sempre foi possível, independentemente do número de repetições da calibragem. Então, se já tivéssemos tentado demasiadas vezes e se o gráfico gerado (exemplo de um participante na Figura 9) não parecesse apresentar nenhuma anomalia grave entre os pontos previstos e os pontos medidos, optamos por seguir e conduzir o teste, considerando que seria mais benéfico aproveitar a coleta do que descartá-la, ainda que imperfeita.

Figura 9 – Resultado da calibração do rastreador ocular efetuada em um participante



Fonte: Tobii Pro Lab (2025), a partir de dados da pesquisa

Em seguida, o participante era informado de que, idealmente, seria capaz de prosseguir completamente sozinho com o teste, mas que poderia pedir auxílios pontuais (obviamente, não sobre o conteúdo dos textos) e reportar problemas para o pesquisador, que se manteve como a única outra pessoa presente no ambiente durante cada bateria. Vale destacar que houve alguma pequena interação entre os participantes e o autor desta dissertação em praticamente todas as coletas de dados: muitos precisaram de ajuda para localizar a tecla “F10” do teclado (não é utilizada com muita frequência para aplicações gerais, mas teve que ser usada no nosso experimento porque era o único atalho do sistema que encerrava a exibição dos questionários na internet, dada a sua condição de “estímulo on-line”) e alguns apenas manifestaram em voz alta a sua surpresa por “não lembrar de nada”, geralmente no momento em que conheciam as perguntas do seu primeiro questionário recebido.

Após a apresentação de uma tela introdutória contendo as seguintes instruções: “Você pode concluir a exibição de textos escritos pressionando a BARRA

DE ESPAÇO do teclado quando tiver terminado a leitura. Os áudios e vídeos avançarão apenas ao final do conteúdo. Quando estiver pronto/a para iniciar, pressione a BARRA DE ESPAÇO do teclado”, iniciava-se o teste propriamente dito. A sessão de testes com cada sujeito foi composta pela exibição, ao todo, de nove textos narrativos diferentes (textos 1 a 9; descritos na subseção 3.4.2), cada um em apenas uma versão, sem repetições (de A até I; descritos na subseção 3.4.2.1). Cada bateria teve variadas, de forma pseudoaleatória, as combinações entre texto e versão de apresentação, além da ordem em que elas foram dispostas para cada sujeito, contemplando o maior número de permutações possível e evitando prováveis efeitos de habituação nos participantes. O sequenciamento das sessões foi feito com o auxílio do *ChatGPT* e está exposto no Quadro 1. As células dos pares texto-condição que estão coloridas em amarelo indicam combinações que somente se repetiram uma vez ao todo, e as que estão na cor laranja são as que não ocorreram em mais de um teste – todas as células brancas apareceram em três ou mais momentos.

Quadro 1 – Sequenciamento dos estímulos para cada participante

Participante	Ordem do estímulo								
	1º	2º	3º	4º	5º	6º	7º	8ª	9º
01a	6D	2B	5H	8A	9F	7E	1G	4C	3I
02b	7I	3C	6G	1H	4F	9B	8E	5D	2A
03c	2G	1B	7E	9A	3F	4D	5H	6I	8C
04d	3E	4C	9G	6I	1A	5D	7H	2F	8B
05e	8B	6E	1I	3H	7F	2D	9C	4G	5A
06f	2E	7G	5C	1F	4A	8I	6D	3B	9H
07g	9I	1G	5F	3A	2D	6H	4C	7B	8E
08h	6C	3H	8F	5E	1D	9A	7G	2B	4I
09i	3I	6F	4H	2C	8G	1E	7A	5B	9D
10j	5G	4E	7C	9B	3H	2F	6I	1A	8D
11k	2C	7A	4D	8I	5B	3F	6H	1E	9G
12l	6B	9C	3G	2F	8H	4E	1I	5A	7D
13m	1F	2A	9E	6D	3G	4B	5C	7I	8H
14n	8C	5A	2H	3B	6I	1G	9D	4F	7E
15o	7F	1C	6E	2I	3A	8G	4H	9B	5D

16p	9B	4A	5H	1G	2C	8D	7I	3F	6E
17q	4G	3B	1D	5I	6H	2E	9F	8A	7C
18r	2F	7C	6A	4I	9D	5E	1B	3H	8G
19s	4B	6H	7E	3D	2I	9A	5F	1G	8C
20t	1D	8I	4B	6G	5H	3A	7F	9C	2E
21u	5C	1H	3E	7B	2G	8F	6D	4A	9I
22v	8F	3A	7G	9I	5E	2H	6C	4B	1D
23w	3C	6H	8E	2D	5F	1A	9G	7B	4I
24x	9D	2G	5I	6B	4C	1H	8E	7A	3F
25y	1C	8A	7H	6F	3G	9E	5B	4D	2I

Superados os nove textos e seus respectivos questionários de compreensão leitora, o participante era comunicado de que, se preferisse, poderia retirar os fones de ouvido e era solicitado a responder, com calma e sem preocupação com o cronômetro, o Questionário de percepção sobre o teste recém-concluído, apresentado também como um formulário do Google e dentro do mesmo ambiente virtual do sistema do *eye tracker*. Todas as respostas dadas pelos participantes nos múltiplos testes, as informações relativas aos tempos decorridos e os dados de movimentação ocular foram armazenados nos respectivos sistemas, diretamente associados ao código do sujeito correspondente. Terminada a última etapa, agradecemos pela contribuição do voluntário e o liberávamos.

Uma vez concluída a descrição de como ocorreu a coleta com os 25 participantes, é interessante destacar alguns pontos. Houve comentários espontâneos consideravelmente diversos, dos quais salientamos a reclamação de muitos sobre existirem nomes demais para memorizar nas histórias. E duas ocorrências são importantes reportar: uma foi a identificação, por parte de um participante, de um erro que existia no questionário de percepção (caso detalhado na parte final da seção 4.4), e outra foi a falha na exibição de um estímulo no formato E (vídeo com áudio) que teve a parte visual substituída por uma tela preta. Esse foi o único problema técnico desse tipo que pudemos perceber, ocorrido no material “8E” do participante “24x”; como a história, ainda assim, foi reproduzida normalmente via áudio e como a pontuação atingida teve valores dentro da normalidade, os dados decorrentes dessa observação foram utilizados nas análises.

3.6 Computação dos dados coletados

As informações centrais para esta investigação – as respostas de todos os participantes para os Questionários de compreensão leitora – foram avaliadas pelo autor desta dissertação. A classificação para cada item foi inspirada nos mesmos critérios de correção e a escala decrescente usados por Daneman e Carpenter (1980): resposta exata, substituição razoável (p. ex., generalizações), resposta incompleta (parcialmente correta), engano (resposta incorreta, mas que envolve algum elemento presente na passagem), fabricação (tem semelhanças com a resposta correta, mas não existe no texto), confusão (incompreensão de aspectos fundamentais) e esquecimento (ausência de resposta apreciável). Para o nosso estudo, adaptamos as categorias usadas pelas autoras a valores numéricos inteiros (de 6 a 0), incluindo uma segunda classificação de engano, para quando a resposta é incorreta e envolve um elemento de uma passagem anterior, indicando que houve um processo de memorização, algo suficientemente diferente de uma confusão total (o esquema de avaliação que usamos pode ser visto no Quadro 2). Além disso, apesar de Daneman e Carpenter (1980) optarem por avaliar de forma distinta a questão que pede um título para a passagem narrativa, atribuindo uma nota de 0 a 2, consideramos que sua escala de correção mais completa foi perfeitamente adequada para qualificar e atribuir valor numérico a essa resposta, bem como para avaliar a quinta questão que incluímos, que solicitava o reconto de qualquer ponto da narrativa que tivesse chamado a atenção do leitor.

Quadro 2 – Esquema de avaliação das respostas

Pontuação	Classificação básica	Descrição complementar
6	resposta exata	--
5	substituição razoável (p. ex., generalizações)	completa, mas não perfeita
4	resposta incompleta (parcialmente correta)	correta, mas falta algo
3	engano (resposta incorreta, mas que envolve algum elemento presente na passagem)	mal colocada
2	engano trazendo elementos de outra passagem anterior	mal colocada (com desvio maior)
1	fabricação (tem semelhanças com a resposta correta, mas	tem algo que pode ser

	não existe no texto)	aproveitado
0	confusão (incompreensão de aspectos fundamentais)	errada
0	esquecimento (ausência de resposta apreciável)	--

O processo de correção foi conduzido apenas pelo autor desta dissertação, de acordo com o seguinte fluxo. Todas as respostas registradas na plataforma Formulários Google foram organizadas e colocadas em planilhas simplificadas (com código do participante e identificação dos números dos textos e das perguntas), a fim de minimizar quaisquer chances de cometer enganos na atribuição de notas. Os textos que os sujeitos digitaram foram apreciados com a ajuda de um gabarito simplificado (Apêndice F). Uma primeira rodada de correções foi feita acompanhando os textos/narrativas como ordenamento primário e as perguntas como secundário, o que significa que se iniciou pela primeira questão (R1) do texto 1 (“A Pista de Patinação”), seguiu-se até a última pergunta (R5) do mesmo texto e, na sequência, recomeçou-se da R1 do texto 2, avançando assim até a R5 do texto 9. A opção por fazer a correção em blocos por narrativa – e não em qualquer outro recorte possível, como o formato – teve como intenção aproximar e facilitar a comparação entre respostas que tratavam sobre um mesmo conteúdo. Essa primeira definição das notas foi realizada em uma sequência de três dias, corrigindo três textos em cada um deles.

Uma semana depois, foi executada uma segunda rodada de correções, essa feita em um único dia, tentando evitar variações espúrias que pudessem aparecer nas notas. Ela foi realizada a partir de uma planilha em branco, sem qualquer referência às notas dadas anteriormente, e organizada seguindo um outro ciclo: começando do texto 5, retrocedendo até o 1, seguindo para o 9 e voltando até encerrar no 6, de maneira que se evitasse qualquer favorecimento pela ordem das correções (p. ex., as notas do texto 1 serem mais confiáveis porque receberiam sempre a primeira análise, menos cansada). Como houve alguma variação entre as pontuações da primeira e da segunda rodada, ainda que o contorno geral tenha ficado muito próximo, foi feita uma espécie de acareação, cerca de uma semana depois, entre os dois conjuntos, comparando onde houve discrepâncias e definindo um número final que representasse de maneira mais fidedigna a resposta do participante (muitas vezes, sendo um meio-termo entre as duas notas originais).

Esse procedimento foi feito em ordem decrescente, começando pelo texto 9 e terminando no 1.

O longo e minucioso processo de correção fez com que alguns critérios para a atribuição de pontuação fossem aperfeiçoados conforme se tornavam mais claros. De modo geral, as notas foram maiores para respostas que apresentavam indícios de melhor compreensão, deixando de priorizar uma abordagem demasiadamente objetiva, que valorizasse de forma desproporcional qualquer demonstração de memória, por mais descontextualizada que fosse. Um caso específico que ilustra bem essa situação: para a pergunta “Onde estava o toca-discos?”, houve um participante que retornou “Vinil”. Uma interpretação muito estreita do nosso esquema de avaliação poderia dar, facilmente, a nota 3 para essa resposta, já que “vinil” realmente é um outro elemento da passagem (texto 9), mas, no fim, pareceu-nos mais adequado classificá-la com a nota 0, porque está muito deslocada, a ponto de não poder ser aproveitável como comprovação de algum nível de compreensão leitora naquele item. Outra decisão tomada no sentido de favorecer demonstrações de entendimento mais global do leitor foi aplicar a pontuação 5 para respostas genéricas como “uma moça/um rapaz”, ainda que a declaração de nomes específicos errados, caso ocorressem na narrativa, receberiam, no máximo, 3. Essa pareceu a opção mais metodologicamente estável, coerente com outras respostas do mesmo tipo, e contou a favor dela o fato de não termos identificado qualquer participante que tenha adotado desproporcionalmente a generalização como uma estratégia de maximizar a sua pontuação – o que seria improvável, já que nenhum deles conhecia os critérios de avaliação.

As questões que pediam um título que capturasse a ideia central do texto (R4), conforme os alertas feitos por Sousa e Hübner (2015) para as tarefas de avaliação produtiva do tipo resumo, apresentaram desafios adicionais, tanto ao participante quanto à correção. Assim como colocado pelas autoras mencionadas, percebemos alguns indivíduos com maior dificuldade para entender o gênero textual a que deveriam se submeter, e o próprio conceito de “ideia central” pode ter sido desafiador para vários dos leitores. Além disso, alguns títulos gerados pareceram muito mais preocupados com questões estéticas e com a demonstração de criatividade do que qualquer outro parâmetro. Junto a isso, houve o problema inerente da subjetividade superior que a avaliação desse tipo de questão trouxe. Adotou-se partir da nota 4 para títulos que contivessem ideias marginais do texto,

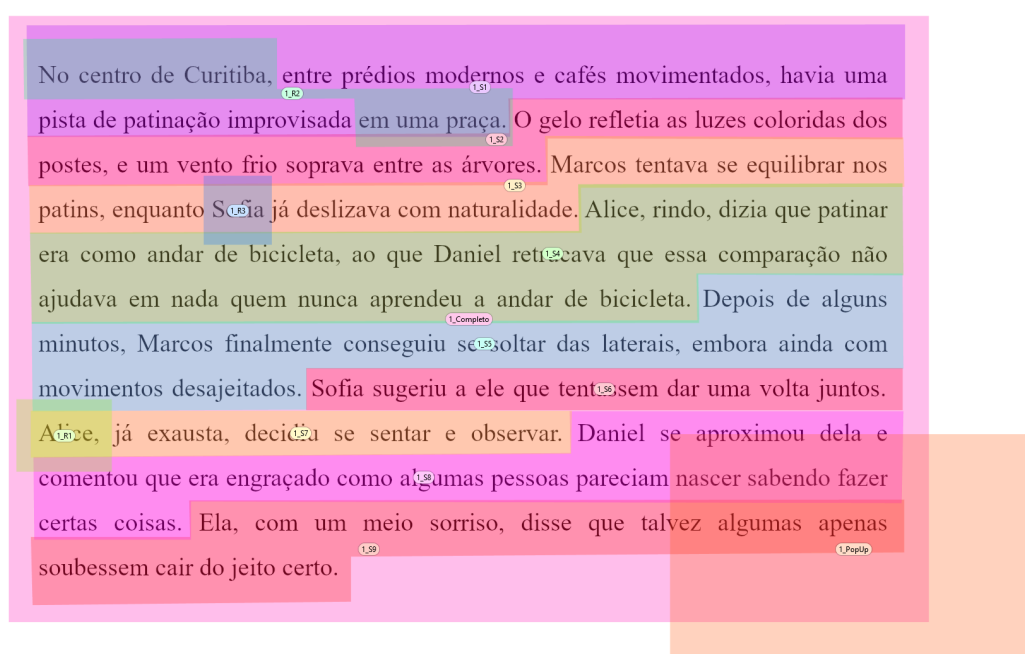
avançando até 6 conforme a aproximação que tivessem do tema central; outros pontos de referência foram a pontuação 3, dada a respostas demasiadamente vagas (p. ex., “Paraty” para o texto 2), e a 1, reservada a itens com relações minúsculas com a história (p. ex., “reflexos da vida” para o texto 1). De forma análoga, a R5, que solicitava o reconto de qualquer fato da história que tivesse chamado a atenção do participante, foi avaliada em função de sua proximidade com o que havia efetivamente na narrativa. A principal distinção trazida por essa pergunta foram as respostas vagas como “o ambiente” ou “descrição dos prédios”, que foram pontuadas com 3 por considerarmos que não entregaram elementos suficientes para comprovar que algo específico do respectivo texto havia sido retido.

Outro dado coletado durante o teste principal, esse diretamente quantificável, diferentemente da pontuação em compreensão leitora, foi o tempo gasto pelo participante na leitura (onde aplicável; as versões de apresentação baseadas em áudio ou vídeo tinham duração fixa), e também na fase de resposta ao questionário de compreensão leitora correspondente. As medições cronológicas de todas as etapas dos testes, com precisão de microssegundos, foram extraídas das informações gravadas pelo sistema de rastreamento ocular.

Desse mesmo equipamento, foram colhidos outros elementos muito mais específicos, que descrevem o comportamento dos olhos dos participantes durante o experimento. Dentre o universo de métricas que o dispositivo “Tobii Pro Fusion” captou e que o software “Tobii Pro Lab” registrou, selecionamos aquelas que teriam o potencial de descrever as diferenças nos processos das variadas leituras que preparamos para os sujeitos da pesquisa. Análises quantitativas foram efetuadas sobre todas as condições experimentais, mas são especialmente significativas para os formatos de apresentação de texto escrito simples (condições A, C, G, H e I), que podem ser diretamente comparadas em praticamente quaisquer aspectos (a exceção mais notável é a duração padronizada que existe nos estímulos C). Para os textos de classificação B e F, os únicos sem estímulos visuais (a não ser por uma tela cinza estática), foram computadas medidas mais gerais, com destaque para a dilatação média da pupila, um aspecto que pode ser uma representação da carga de trabalho mental (Holmqvist *et al.*, 2015). Os formatos D e E, baseados em vídeo, apesar de precisarem de um tratamento diferenciado, por causa de sua natureza específica, já puderam ser examinados a partir de métricas mais específicas – quantias e durações de fixações e sacadas, bem como as características de

velocidade e amplitude destas –, especialmente a condição D, que contém linguagem verbal escrita, ainda que em apresentação dinâmica, em forma de legendas. A atenção orientacional (Dehaene, 2020) nesses estímulos com imagens e vídeos e nos compostos por texto estático pôde ser aferida a partir da delimitação de segmentos relevantes na tela, as AOIs (*areas of interest*; áreas de interesse). Os principais locais que demarcamos foram os que possuíam o conteúdo necessário para as respostas às três perguntas “curtas” (R1, R2 e R3) do questionário de compreensão leitora.

Figura 10 – AOIs delimitadas no estímulo 1A



Fonte: Tobii Pro Lab (2025) e o autor, a partir de dados da pesquisa.

Figura 11 – AOIs delimitadas em trecho do estímulo 1D



Fonte: Tobii Pro Lab (2025) e o autor, a partir de dados da pesquisa.

As Figuras 10 e 11 demonstram exemplos de como foram desenhadas as AOIs em duas das condições experimentais com as maiores distinções. No formato A, além da divisão do texto em frases (que não receberá atenção mais profunda no presente estudo), há os já mencionados destaques para trechos que contêm respostas, uma área que engloba todo o texto, e um retângulo no canto inferior direito, planejado para fazer comparações com a condição experimental I, na qual há uma janela *pop-up* intermitente. A visualização das AOIs no formato D é mais complexa, porque envolve áreas que são ativadas e desativadas nos devidos tempos, então todas as formas geométricas que englobam elementos que ocorrem durante a reprodução do vídeo estão aglomeradas nessa visualização estática que trouxemos. O importante é apresentar que, para um único conteúdo de resposta ao questionário (p. ex., “Alice”), pode haver até três áreas de interesse: o termo na legenda, a ilustração do personagem e o ator retratando, em vídeo, outra representação desse mesmo personagem. Na Figura 11, também é possível ver dois pequenos círculos vermelhos sobre a legenda, que são a maneira como o sistema “Tobii Pro Lab” simboliza a fixação no tempo atual (com contorno branco) e a que a precedeu, com uma linha entre as duas para indicar o traçado da sacada.

A análise qualitativa das gravações registradas pelo *eye tracker*, mostrando todo o percurso realizado pelos olhos, foi um expediente que usamos para sanar dúvidas a respeito dos números informados nas planilhas geradas pelo sistema. Em estudos futuros, poderão ser feitas observações mais detalhadas desse material, sendo, aí, possível computar movimentos oculares mais intrincados, como sacadas para a frente, regressões, refixações e varreduras de retorno (Rayner et al., 2016; Riess; Gabriel, 2020).

As respostas coletadas no Questionário de perfil do participante e no Questionário de percepção sobre o teste de compreensão leitora foram inseridas em planilhas que caracterizaram os sujeitos de maneira global, não vinculadas às subdivisões por estímulo, essenciais para as outras métricas (pontuação, tempo e dados oculares). As informações categóricas (nominais e ordinais) foram relacionadas às medidas coletadas a fim de encontrar explicações complementares ou alternativas para as variações existentes nos padrões de leitura verificados (ver seção 4.9).

A análise estatística dos dados foi conduzida no software Jamovi¹⁸, fazendo uso, também, do módulo/extensão GAMLj3¹⁹ (*General Analyses for Linear Models*), que amplia as capacidades nativas do programa para manejar modelos lineares. O nível de significância de $p < 0,05$ (alfa (α) = 5%) foi usado para todos os testes. As ferramentas específicas aplicadas em cada situação do nosso estudo estarão detalhadas nos momentos correspondentes do capítulo seguinte.

¹⁸ The jamovi project (2024). *jamovi*. (Version 2.6) [Computer Software]. Retrieved from <https://www.jamovi.org>.

¹⁹ Gallucci, M. (2019). *GAMLj: General analyses for linear models (Version 3.6.4)*. [jamovi module]. Retrieved from <https://gamlj.github.io/>.

As páginas 124 a 238, referentes ao Capítulo 4 “Apresentação e análise dos dados”, foram omitidas desta versão da dissertação. Para ter acesso a elas, entre em contato com o autor pelo e-mail jordanjunges@mx2.unisc.br.

5 CONSIDERAÇÕES FINAIS

Ao chegarmos a este capítulo de desfecho da dissertação, voltamos à dúvida essencial, que inspirou e norteou esta pesquisa: há benefícios para a compreensão linguística quando o texto é recebido simultaneamente em duas modalidades, visual e auditiva? Conforme foi apresentado e discutido nas últimas seções, o trabalho que conduzimos torna consideravelmente claro que responder a essa pergunta envolve uma rede complexa de associação de variáveis que não são facilmente isoladas e cuja dominância não é evidente, reforçando o padrão encontrado até mesmo em um estudo de metanálise concentrado de forma mais específica sobre a distração auditiva durante a leitura (Vasilev; Kirkby; Angele, 2018), no qual os estudos individuais compilados apontam para efeitos com uma multiplicidade de direções e magnitudes, tendo como resultado geral uma influência pequena (negativa, no caso). Sendo assim, nenhuma de nossas hipóteses restaram confirmadas – ou sequer negadas com rigor –, embora tenhamos percebido tendências que ajudam a compreender o panorama e a provisionar futuros estudos sobre o mesmo tema.

Nosso primeiro objetivo específico e sua hipótese correspondente, os quais esperavam encontrar vantagens à compreensão linguística em condições experimentais onde houvesse bimodalidade sensorial redundante, ou seja, áudio e vídeo apresentando conteúdos equivalentes, quando em comparação com os estímulos simples unimodais, apenas são suportados por evidências frágeis: pontuações marginalmente superiores (longe da significância estatística) da classe bimodal sobre a unimodal e, de forma individualizada, da condição E²⁰ (vídeo com áudio falado) sobre B (apenas áudio falado). Considerando-se o tempo necessário para responder aos questionários de compreensão leitora, o formato E, mais uma vez em posição superior, revelou a tendência de ser mais eficiente do que B; entretanto, essa mesma métrica é menos favorável ao formato bimodal C (texto estático com áudio falado) em relação a ambos B e A (apenas texto estático). Em termos de dados de rastreamento ocular, maior dilatação de pupilas – um indicador de aumento de carga de trabalho mental (Holmqvist *et al.*, 2015) – foi verificada em E, com diferença estatisticamente significativa na comparação com B. Contudo, não

²⁰ Recomendamos a consulta ao texto e à figura no início da subseção 3.4.2.1 para que o leitor desta dissertação possa, rapidamente, recordar o código de letras que usamos para identificar as condições experimentais.

se pode ignorar a existência plausível de uma explicação alternativa para esse contraste, pela qual a abertura das pupilas variaria unicamente em resposta aos diferentes brilhos de tela dos estímulos visuais distintos.

Desde a análise desse objetivo específico, já aparece de forma notória a ideia de que a compreensão linguística está menos sujeita à *quantidade* de modalidades do que à *qualidade*. É dentro da mesma classe (unimodal) que se pode verificar uma tendência acentuada de o formato visual (A) ter índices de acurácia mais altos do que os do formato auditivo (B). Esta mesma apresentação simples em áudio, de fato, pareceu se beneficiar com a adição de elementos no canal visual, fosse texto estático (C) ou vídeo com imagens redundantes (E). O que entendemos a partir desses dados, agregando, também, o conhecimento gerado por pesquisas que fizeram uso de materiais infográficos (Mousavi; Low; Sweller, 1995; Mayer; Moreno, 1998) é que o conteúdo base que se quer comunicar tem uma “vocalização”, uma predisposição a ser codificado mais aptamente de uma maneira ou outra, ainda mais quando se espera que o processo de compreensão atinja uma meta objetiva, p. ex., chegar à resposta exata de um problema matemático. Da mesma forma que parece muito mais eficiente apresentar uma forma geométrica desenhada em vez de descrevê-la estritamente pela fala ou pelo texto escrito (detalhando, passo a passo, a organização de elementos para que o sujeito represente o todo mentalmente), as passagens narrativas curtas que utilizamos em nossos testes e a exigência de que os participantes lembrassem de nomes e fatos específicos pareceram mais bem ajustadas à estrutura do texto estático e às possibilidades de revisão que essa mídia proporciona. De maneira menos óbvia, nossos resultados sugerem que a adição de vídeo (um conteúdo não verbal, menos específico do que as palavras, e, além do mais, fora da possibilidade de controle temporal ou sequencial pelo sujeito) ao texto falado, rendeu índices mais altos de acurácia do que quando este foi ouvido de forma pura. Esse resultado sugere que as informações podem ser reforçadas articulando-se mais de uma modalidade, desde que, como entendido por Brasel e Gips (2014), a densidade total de dados transmitidos não exceda a capacidade cognitiva do leitor/espectador.

O segundo objetivo específico desta dissertação, que procurou medir se a multimodalidade poderia sobrecarregar a memória de trabalho do leitor, não recebeu resposta robusta, fosse para o sentido positivo ou negativo, a partir do método que desenvolvemos e executamos. Com o nível de significância estatística definido e

com a tomada de cuidados para se evitar a ocorrência de falsos positivos, pode-se afirmar que todas as nove condições experimentais são idênticas tanto do ponto de vista da pontuação de compreensão leitora quanto do tempo de resposta. Essa constatação, ainda mais quando analisada em conjunto, por exemplo, com a diferença significativa entre as médias de dilatação pupilar das condições experimentais D e E (ambas muito similares, baseadas nos mesmos vídeos, esta última sendo a condição bimodal e que registrou os valores maiores de diâmetro das pupilas), aponta para uma conclusão intrigante: embora nenhum dos parâmetros medidos acuse uma *sobrecarga* da memória de trabalho, já que se encontram bastante nivelados, toda a *carga* variável com a qual a cognição do participante teve que lidar (afinal, são formatos de leitura bastante diversos) foi absorvida de alguma maneira. No caso das condições experimentais derivadas do mesmo texto estático (A, C, G, H e I), a ausência, inclusive, de disparidades nos pormenorizados dados de rastreamento ocular motiva-nos a refletir sobre quais construtos mentais poderiam ser responsáveis por atenuar diferenças entre os estímulos recebidos, e de que forma isso se processaria.

Pensando em termos das tendências que conseguimos observar, a audição de música instrumental durante a leitura de texto escrito (condição G) gerou pontuações de compreensão inferiores às da leitura silenciosa, o que está em linha com os achados de Vasilev, Kirkby e Angele (2018), e também abaixo de formatos contendo itens com o único propósito de causar perturbações (H e I), chamando a atenção para o efeito negativo que um elemento aparentemente inócuo – até mesmo agradável para muitos indivíduos – pode exercer sobre a concentração do leitor. De todo modo, o principal desafio que nosso teste propôs à memória de trabalho dos participantes aparenta ter sido comum às nove condições experimentais: a quantidade de nomes de personagens apresentados nas narrativas e a exigência de retenção e evocação de parte desses para responder aos questionários. Essa constatação encontra amparo em relatos espontâneos dados pelos sujeitos e no fato de existirem déficits homogêneos nos índices de acurácia em geral, ou seja, independentes dos atributos que dizem respeito às modalidades sensoriais.

A terceira hipótese desta dissertação pretendia encontrar a submodalidade (múltiplos fluxos de informação no mesmo canal sensorial, conforme Gardner e Gardner, 2023) como um ponto médio em termos de desempenho, acima da unimodalidade e abaixo da bimodalidade. Como nem mesmo a comparação entre

essas duas últimas classes indicou contrastes relevantes, a submodalidade apresentou nuances ainda mais complexas, que parecem reforçar a ideia mencionada anteriormente das particularidades, potencialidades e fraquezas, inerentes a cada modo sensorial e formatação de texto. A condição experimental D (vídeo silencioso com legendas) aparentemente representa o efeito mais óbvio, desfavorável, nesse caso, quando feita a comparação com um formato de apresentação similar (E, vídeo sem texto escrito mas com narração, pertencente à classe bimodal). As legendas, além de excluïrem os benefícios do controle do tempo de leitura por parte do usuáριο, da possibilidade de revisão de trechos e da perspectiva de visão global do texto, também capturaram mais de 80% do olhar do sujeito, reduzindo possíveis vantagens que viriam da visualização dos elementos não verbais presentes no vídeo. Apesar de esses efeitos não aparecerem senão sutilmente nas diferenças nos escores de acurácia de D e E, as preferências dos sujeitos sugerem claramente um apreço mais baixo por aquela condição submodal. Esse padrão de predileção se inverte quando consideramos a apresentação de texto F (sem vídeo, apenas com o texto da narrativa codificado em forma de música) comparada com a B (texto falado, unimodal). Nesse caso, a exemplo do que foi percebido por Schön e colaboradores (2008), os elementos musicais parecem agregar valor às informações, estruturando-os de forma mais saliente para o tratamento da memória de trabalho e de sistemas de armazenamento de mais longa duração. A submodalidade, portanto, aqui, é vista como superior à unimodalidade, já que B obteve pontuações ligeiramente menores e apareceu mais vezes classificada como difícil. Ressaltamos, no entanto, que não existe um padrão que coloque a classe submodal no patamar mediano que inicialmente hipotetizamos: a condição D, por exemplo, está ligada a menores desempenhos do que os medidos na A, unimodal, considerando-se que ambas são comparáveis porque apresentam a linguagem verbal na forma escrita.

O quarto e o quinto objetivos específicos podem ser analisados conjuntamente, já que são bastante aproximados e comparam a condição experimental A (texto escrito) diretamente com os formatos que contêm elementos não redundantes (G, H e I, todos constituídos a partir do mesmo texto presente em A). Enquanto uma dessas hipóteses avaliou impactos à compreensão leitora nos estímulos bimodais desse conjunto (exclui-se I, que é submodal, ocupando apenas o canal da visão), a seguinte se concentrou especificamente sobre as condições experimentais que

continham entradas de perturbações intermitentes (auditivas no H e visuais no I; exclui-se G desse grupo, já que a música instrumental que o compõe é uma perturbação contínua). De modo geral, o único indício de que as três condições com elementos não redundantes modificaram o desempenho dos participantes para pior são os tempos totais (somando leitura e resposta ao questionário) que foram ligeiramente mais altos do que no formato unimodal A. Conforme referimos, a leitura acompanhada por música instrumental (G) apresentou uma tendência de ter menor pontuação de compreensão leitora quando comparada à leitura silenciosa de A; porém, nem o formato H (o outro formato bimodal sem redundância) e nem o I apresentaram escores de acurácia que fossem notavelmente inferiores. A condição I dá uma ideia bastante concreta do quanto as perturbações intermitentes que desenvolvemos foram inofensivas: a janela *pop-up* que surgia em intervalos irregulares não alterou significativamente nem mesmo o tempo e o número de fixações sobre a área da tela que ocupava. A julgar pelas preferências relatadas pelos participantes, os *bips* sonoros presentes em H foram mais irritantes – até mesmo porque as naturezas do som e da nossa audição não permitem uma estratégia de direcionamento de atenção equivalente a desviar o olhar – mas, como reiteramos ao longo de grande parte do capítulo de apresentação e análise dos dados, o fundamental é ter em mente que as diferenças em parâmetros secundários não foram respaldadas por variações consistentes em tempo de resposta e, sobretudo, em pontuação de compreensão linguística.

Para além dos cinco objetivos específicos e levando em conta a constatação de poucas diferenças estatisticamente significativas ou tendências sólidas pela óptica da quantidade de modalidades sensoriais acionadas, conduzimos análises adicionais para entender a influência de possíveis fatores alternativos sobre nossos resultados. A partir de um primeiro grupo de testes, mais aproximado de nossas investigações principais, não identificamos contrastes que estivessem relacionados ao fato de o estímulo ser completamente silencioso ou não, bem como não houve diferenças relevantes quando as condições experimentais foram agrupadas em função de qual canal sensorial carregava a informação verbal (visão, audição ou a união de ambas). Em um segundo grupo, os fatores analisados foram aqueles cujas possíveis distorções nós tentamos controlar por meio do desenho experimental que adotamos. A ordem de apresentação dos estímulos exibiu uma curva de desempenho que sugere uma queda por volta dos 15 minutos de teste, bem como

evidencia um comportamento consideravelmente diferente no primeiro bloco de texto e questionário, o que justificaria a execução de baterias preliminares (*trials*) para que o sujeito se acostumasse previamente à dinâmica, com o cuidado simultâneo de não cansá-lo antes de iniciar a testagem propriamente dita. Já as narrativas que elaboramos com o uso de inteligência artificial generativa, e seus respectivos questionários, esses feitos manualmente – lembrando que produzimos esses materiais dessa maneira porque encontramos muita dificuldade para encontrar textos e testes padronizados que servissem ao nosso propósito –, indicaram variações que precisam ser controladas em instrumentos futuros: os textos 7 e 8 tiveram, nitidamente, as melhores pontuações, principalmente nas perguntas referenciais e factuais. Finalmente, um terceiro conjunto de análises que agrupou os participantes de acordo com determinados atributos apontou diferenças curiosas que merecem futuro aprofundamento. O gênero masculino e a área de estudo das Ciências Exatas apresentaram diferenças significativas nas suas respectivas comparações, mas é importante destacar que ambas as delimitações representam quase o mesmo subconjunto de indivíduos, que também é composto por um número comparativamente menor de sujeitos.

Considerando o objetivo geral desta dissertação – analisar possíveis interferências, desde a promoção de auxílio até a geração de obstáculo, provocadas a partir da associação, em um único texto, das modalidades sensoriais visual e auditiva à compreensão leitora de estudantes universitários – não foi possível identificar se a multimodalidade atua, de maneira consistente, a favor ou contra a compreensão do leitor/ouvinte/espectador. Como mencionado ao longo das últimas seções deste trabalho, a quantidade de modalidades, aparentemente, é um dos fatores com influência mais limitada sobre o desempenho em compreensão linguística, pelo menos adotando o método que elaboramos, considerando, também, a amostra investigada. Como defende Dehaene (2014), a união dos *inputs* da visão, audição e outros sentidos tipicamente ocorre antes de o bloco de informação (ou *chunk*) chegar até a consciência, o que sugere que níveis cognitivos mais elevados, como os relacionados à compreensão, já lidam com dados abstraídos de sua forma material – que deixaram de ser algo que, necessariamente, leu-se ou ouviu-se para se tornarem uma ideia em geral. Esse entendimento tem consonância com os dados neurológicos observados por Deniz e colaboradores (2019), que identificaram que a

informação semântica é representada no córtex cerebral de uma forma que não varia pela modalidade de entrada do estímulo, tenha ele sido lido ou ouvido.

No entanto, é possível perceber que a qualidade das modalidades (e isso inclui a forma específica como as submodalidades são utilizadas) tem influência considerável nos processos mais elaborados de memorização e integração de informações. A estruturação dos dados, ou seja, a concisão do texto e a sua organização para facilitar a compreensão, atua em conjunto com as potencialidades e limitações intrínsecas da mídia e do formato em que esse conteúdo está codificado. Então, da mesma maneira que uma imagem ou fotografia é muito mais aptamente representada graficamente do que verbalmente, o texto escrito serve melhor ao intuito de reproduzir os detalhes de uma passagem narrativa, ainda mais quando o propósito de sua leitura é responder a questões majoritariamente pontuais, envolvendo nomes e fatos descritos. Sendo assim, no contexto do teste de compreensão que desenvolvemos, não surpreende que a condição experimental A (texto estático sem qualquer adição de elementos) tenha apresentado a tendência de ser o formato superior de maneira geral, algo que pode ser constatado a partir de seu maior escore médio de acurácia, seu menor tempo de reação para resposta aos questionários e sua alta votação na predileção dos participantes por esse formato de texto. Sublinhamos que isso não significa dizer que a unimodalidade é superior em todos os casos: nossos dados apontam para um desempenho consideravelmente inferior na condição B (a única outra unimodal, que contém apenas texto falado). As vantagens dos estímulos do formato A para a tarefa investigada são claras, pois o leitor, além de poder controlar seu tempo de exibição, não teve que inibir a resposta a quaisquer distrações programadas, pôde reler à vontade os trechos que julgou precisar de melhor compreensão ou memorização e teve a possibilidade de examinar a estrutura geral da narrativa a fim de facilitar a construção de seu modelo mental da história.

É claro que esses benefícios do texto estático (A) foram mais bem aproveitados porque os sujeitos da pesquisa contavam com um alto nível de educação formal, todos frequentando aulas de nível universitário, o que significa que já passaram por inúmeras leituras e testes em formatos similares, estando devidamente treinados para essa situação. Levando-se em conta, também, a faixa etária jovem dos 25 sujeitos pesquisados e o contexto em que vivem, é lícito deduzir que a maioria deles é familiarizada, em maior ou menor grau, com todas as nove condições

experimentais às quais foi exposta. Esse ponto é especialmente crítico quando consideramos que esses indivíduos, presumivelmente, encontram-se fortemente habituados a uma variedade de distrações enquanto conduzem suas atividades de estudo e trabalho. Dessa forma, as perturbações demasiadamente controladas que desenhamos (formatos G, H e I) podem não ter sido páreo para as situações reais que costumam enfrentar, e, se nem esses piores cenários que propusemos tiveram impacto significativo, não é lógico esperar que informações redundantes, apresentadas em bimodalidade ou submodalidade, tivessem qualquer influência mais perceptível. Todo esse quadro, inegavelmente, poderia exibir alterações dramáticas caso a população estudada fosse, por exemplo, mais idosa, muito mais jovem (com pouca experiência de leitura), com educação formal inferior, ou sem acesso a tecnologias de informação e comunicação.

Além dessas questões que dizem respeito ao grupo de pessoas cuja compreensão linguística em condições de leitura multimodal decidimos investigar, há outras limitações que identificamos em nossa pesquisa, que podem vir a inspirar estudos futuros. Uma parte delas está no nível da análise de dados e pode tanto ser aprimorada a partir do que já foi coletado quanto orientar futuros experimentos. No nível mais simples, seria possível computar e interpretar os movimentos oculares realizados pelo sujeito durante a fase de resposta aos questionários, e não apenas da recepção dos textos. Para se atingir um aprofundamento ainda maior, análises quantitativas mais detalhadas podem vir a esclarecer os motivos por trás do fato de haver maior dilatação média das pupilas quando o vídeo continha conteúdo verbal em som, não em legenda. Nessa mesma linha, seria interessante investigar por que foram medidas mais fixações no estímulo com música cantada e mais movimentos oculares classificados como “outros” naquele com texto falado, já que esses estímulos não continham qualquer diferença em termos de elementos visuais (ambos exibiram a mesma tela estática cinza). Um levantamento manual e consideravelmente trabalhoso poderia computar dados de *eye tracking* mais granulares, no nível das sentenças ou até das palavras, em busca de origens pontuais de falhas de compreensão leitora. Esse mesmo procedimento possivelmente tenha como ser automatizado com técnicas de análise de dados mais elaboradas, ou, certamente, pode ser facilitado por meio do uso de softwares mais sofisticados de gravação e computação de dados oculares, algo que vale, igualmente, para a extração de dados de piscadas do sujeito, que são,

potencialmente, reveladoras sobre sua carga de processamento mental (Holmqvist *et al.*, 2015).

Quanto aos estímulos e às condições experimentais, alguns aprimoramentos ou alternativas poderiam ser testados. Como o formato I que elaboramos (texto estático com entradas irregulares de uma janela *pop-up* cinza transparente sobre o canto inferior direito da tela) mostrou pouco poder de distração, pelo menos três abordagens diferentes poderiam ser implementadas: a primeira, seguindo a mesma ideia restritiva (evitando a inserção de quaisquer elementos novos), seria fazer com que a mesma janela cinza transparente surgisse em pontos aleatórios da tela; a segunda, admitindo um pouco mais de variação, seria fazer o *pop-up* trocar de cor e/ou formato de maneira randômica; e a terceira, partindo em direção a uma maior validade ecológica (em detrimento do controle rígido das variáveis), seria adicionar outros textos às notificações visuais que aparecem, talvez, simulando conversas típicas em aplicativos de mensagem. Adotando modificações similares, os *bips* sonoros do formato H poderiam ser substituídos por sons que variassem aleatoriamente (não apenas tocando a mesma campainha monótona) ou até mesmo por trechos de voz falada. Pensando sobre os estímulos cujo conteúdo é exclusivamente sonoro (B e F), uma alteração possível seria o desenho de áreas de interesse (AOIs) que cobrissem toda a tela e estivessem ativas apenas quando um elemento de resposta aos questionários (nome, fato, objeto, entre outros) fosse mencionado no decorrer do áudio, o que poderia permitir uma análise qualitativa da atenção (especialmente da dilatação pupilar) em pontos-chave da narrativa. Outra possível modificação em termos de formato de apresentação de texto poderia ser a inclusão de uma versão C (texto estático com áudio falado) que contasse com legendas iguais às da condição experimental D, mantendo-se sem elementos imagéticos não verbais. O principal motivo para elaborar esse formato seria garantir que o olhar do leitor seguisse o mesmo ritmo da fala, algo que não foi tão previsível, já que o estímulo C que desenhamos mantinha na tela o texto escrito completo desde o princípio até o fim da reprodução do áudio.

Sobre a elaboração das narrativas e, principalmente, dos materiais de áudio e vídeo construídos com o uso de ferramentas de inteligência artificial generativa, é seguro afirmar que há um grande potencial para se atingir resultados mais apropriados e em menos tempo, envolvendo menos intervenções manuais. Esses aplicativos passam por constantes aprimoramentos e, para se ter acesso ao sistema

de ponta disponível no momento, uma alternativa é contratar as versões pagas desses serviços. Uma alteração que seria benéfica aos textos-base que utilizamos e que pode ser implementada a despeito da tecnologia é o cuidado com os nomes usados em cada história: garantir que nenhum deles se repita em mais de uma narrativa é a forma de assegurar que se possa mapear de onde exatamente o participante recuperou memórias dessa natureza. Para os questionários, além de ser recomendável ter maior cuidado para que as perguntas factuais sejam equivalentes entre narrativas, uma ideia interessante seria pensar em incluir questões que fossem exclusivamente sobre aspectos não verbais presentes nos vídeos (cores, aparências de personagens e cenários), propiciando uma maneira de identificar lembranças específicas capturadas por cada modalidade sensorial.

Por último, é importante voltar a atenção para um dos pontos mais elementares, discutido por pesquisadoras que referenciamos (Tomitch, 2008; Sousa; Hübner, 2015), e refletir sobre qual tipo de compreensão leitora/linguística queremos e podemos aferir, e também de que forma isso pode ser feito da melhor maneira. O questionário que empregamos, apesar de ter sido desenvolvido para evitar as tradicionais armadilhas dos instrumentos de múltipla escolha (sobretudo, a possibilidade de adivinhação de alternativas corretas) e de oferecer uma perspectiva não dicotômica de apreciação das respostas informadas (usando uma escala de pontuação inspirada pela classificação de erros de Daneman e Carpenter, 1980), ainda está muito ancorado em processos básicos de memorização, especialmente de nomes ou outros termos curtos. Seria propício aprimorar o instrumento que empregamos nesta pesquisa para que ele tenha a capacidade de lidar com processos inferenciais mais complexos, desde que não se perca o controle das demais variáveis que supomos ter restringido.

Em conclusão, reforçamos a ideia que se tornou mais clara conforme este trabalho foi desenvolvido: a quantidade de modalidades sensoriais envolvidas na comunicação de uma informação é um fator que parece ter impactos moderados sobre a compreensão linguística e a memorização de dados. A forma como esses conhecimentos são estruturados apresenta-se como o aspecto fundamental e, a partir daí, a multimodalidade pode ser capaz de potencializar a transmissão do conteúdo, formando *chunks* de memória multifacetados e com mais conexões disponíveis ao conhecimento prévio do indivíduo. A despeito disso, é fundamental sublinhar que, para boa parte das aplicações e tarefas, o uso de um simples texto

escrito ainda parece ser a ferramenta mais indicada. Destacamos, finalmente, que um achado positivo desta pesquisa é que alunos universitários, com idade média pouco superior aos 20 anos, aparentemente estão mais preparados do que o esperado para lidar com distrações em tarefas de estudo. Resta investigar e entender se essa habilidade se apresenta, também, em outros grupos, especialmente naqueles compostos por indivíduos mais jovens, que têm sido desafiados por mídias com potencial ainda mais distrator do que as da geração prévia, e qual o custo cognitivo dessa adaptação.

REFERÊNCIAS

- ALCARÁ, A. R.; SANTOS, A. A. A. dos. Avaliação e desenvolvimento da compreensão de leitura em universitários. *Estudos de Psicologia*, Campinas, v. 32, n. 1, p. 63–73, jan./mar. 2015. DOI 10.1590/0103-166x2015000100006. Disponível em: <https://www.scielo.br/j/estpsi/a/L6tnnsWqGWKkfV5wwQdPLXK/?lang=pt>. Acesso em: 1 abr. 2025.
- ARMSTRONG, G. B.; BOIARSKY, G. A.; MARES, M. Background television and reading performance. *Communication Monographs*, v. 58, n. 3, p. 235–253, Sept. 1991. DOI 10.1080/03637759109376228. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/03637759109376228>. Acesso em: 3 mar. 2025.
- AUMONT, J. *A imagem*. 16. ed. 4. reimp. Tradução: Estela dos Santos Abreu, Cláudio C. Santoro. Campinas: Papyrus, 2014.
- BACK, R. Compreensão, aprendizagem e linguagem: uma abordagem hermenêutica. *Educação e Pesquisa*, São Paulo, v. 49, p. e260638, 2023. DOI 10.1590/s1678-4634202349260638. Disponível em: <https://www.scielo.br/j/ep/a/CXLJZhG8mYcWg7zjYcNFNQm/?lang=pt>. Acesso em: 20 mar. 2025.
- BADDELEY, A. O que é a memória?. In: BADDELEY, A.; ANDERSON, M. C.; EYSENCK, M. W. *Memória*. Tradução: Cornélia Stolting. Porto Alegre: Artmed, 2011a. cap. 1, p. 13-30. *E-book* (471 p.). Disponível em: <https://biblioteca-a.read.garden/viewer/9788536325194/13>. Acesso em: 4 mar. 2025.
- BADDELEY, A. Memória de curta duração. In: BADDELEY, A.; ANDERSON, M. C.; EYSENCK, M. W. *Memória*. Tradução: Cornélia Stolting. Porto Alegre: Artmed, 2011b. cap. 2, p. 31-53. *E-book* (471 p.). Disponível em: <https://biblioteca-a.read.garden/viewer/9788536325194/31>. Acesso em: 4 mar. 2025.
- BADDELEY, A. Memória de trabalho. In: BADDELEY, A.; ANDERSON, M. C.; EYSENCK, M. W. *Memória*. Tradução: Cornélia Stolting. Porto Alegre: Artmed, 2011c. cap. 3, p. 54-82. *E-book* (471 p.). Disponível em: <https://biblioteca-a.read.garden/viewer/9788536325194/54>. Acesso em: 4 mar. 2025.
- BADDELEY, A. A aprendizagem. In: BADDELEY, A.; ANDERSON, M. C.; EYSENCK, M. W. *Memória*. Tradução: Cornélia Stolting. Porto Alegre: Artmed, 2011d. cap. 4, p. 83-106. *E-book* (471 p.). Disponível em: <https://biblioteca-a.read.garden/viewer/9788536325194/83>. Acesso em: 4 mar. 2025.
- BADDELEY, A. D.; HITCH, G. Working memory. *Psychology of Learning and Motivation*, [S.l.], v. 8. p. 47-89, 1974. DOI 10.1016/s0079-7421(08)60452-1. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0079742108604521?via%3Dihub>. Acesso em: 19 fev. 2025.

BALZAN, C. F. P. *Da leitura prática às práticas de leitura: o caso dos cursos superiores de tecnologia*. 2018. Tese (Doutorado) - Programa de Pós-Graduação em Letras, Universidade de Caxias do Sul em associação ampla UniRitter, Caxias do Sul, 2018. Disponível em: <https://repositorio.ucs.br/xmlui/handle/11338/3922>. Acesso em: 1 abr. 2025.

BARBOSA, P. A. Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 20, n. 1, p. 11–27, jun. 2012. DOI 10.17851/2237-2083.20.1.11-27. Disponível em: <https://periodicos.ufmg.br/index.php/relin/article/view/28640>. Acesso em: 5 mar. 2025.

BEILOCK, S. L.; CARR, T.; MACMAHON, C.; STARKES, J. When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied*, v. 8, n. 1, p. 6–16, 2002. DOI 10.1037/1076-898X.8.1.6. Disponível em: <https://psycnet.apa.org/doiLanding?doi=10.1037%2F1076-898X.8.1.6>. Acesso em: 4 mar. 2025.

BRASEL, S. A.; GIPS, J. Enhancing television advertising: same-language subtitles can improve brand recall, verbal memory, and behavioral intent. *Journal of the Academy of Marketing Science*, v. 42, n. 3, p. 322–336, May 2014. DOI 10.1007/s11747-013-0358-1. Disponível em: <https://link.springer.com/article/10.1007/s11747-013-0358-1>. Acesso em: 3 mar. 2025.

BRASIL. Ministério da Educação. Secretaria de Alfabetização. *PNA: Política Nacional de Alfabetização*. Brasília: MEC, SEALF, 2019.

BRASIL. Lei nº 15.100, de 13 de janeiro de 2025. Dispõe sobre a utilização, por estudantes, de aparelhos eletrônicos portáteis pessoais nos estabelecimentos públicos e privados de ensino da educação básica. *Diário Oficial da União*, Brasília, DF, 14 de janeiro de 2025. Disponível em: <https://www.in.gov.br/en/web/dou/-/lei-n-15.100-de-13-de-janeiro-de-2025-606772935>. Acesso em: 19 fev. 2025.

BROADBENT, D. E. *Perception and communication*. Elmsford: Pergamon Press, 1958. DOI 10.1037/10037-000. Disponível em: <https://psycnet.apa.org/PsycBOOKS/toc/10037>. Acesso em: 4 mar. 2025.

CARVALHO, M. G. M. de; SOUZA, A. C. de. A avaliação da leitura no Brasil entre os anos 2014-2020: instrumentos e habilidades. *Educação e pesquisa*, São Paulo, v. 49, p. e259865, 2023. DOI 10.1590/S1678-4634202349259865. Disponível em: <https://www.scielo.br/j/ep/a/WwBSFWPPzT6Pc9L8gMPQPm/>. Acesso em: 1 abr. 2025.

COLTHEART, M., CURTIS, B., ATKINS, P., HALLER, M. Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, v.

100, n. 4, p. 589–608, Oct. 1993. DOI 10.1037/0033-295X.100.4.589. Disponível em: <https://psycnet.apa.org/record/1994-04292-001>. Acesso em: 19 fev. 2026.

CORRÊA, L. M. S. O desencadeamento (*bootstrapping*) da sintaxe numa abordagem psicolinguística para a aquisição da linguagem. In: FINGER, I.; QUADROS, R. M. de. (org.). *Teorias de aquisição da linguagem*. Florianópolis: Ed. da UFSC, 2008. cap. 6, p. 169-220.

COWAN, N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, v. 24, n. 1, p. 87–185, 2000. DOI 10.1017/S0140525X01003922. Disponível em: https://www.researchgate.net/publication/11830840_The_Magical_Number_4_in_Short-Term_Memory_A_Reconsideration_of_Mental_Storage_Capacity. Acesso em: 18 fev. 2025.

DANEMAN, M.; CARPENTER, P. A. Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, [S.l.], v. 19, p. 450–466, 1980. DOI 10.1016/S0022-5371(80)90312-6. Disponível em: <https://psycnet.apa.org/record/1981-22775-001>. Acesso em: 1 abr. 2025.

DEHAENE, S. *Os neurônios da leitura: como a ciência explica a nossa capacidade de ler*. Tradução: Leonor Scliar-Cabral. Porto Alegre: Penso, 2012.

DEHAENE, S. *Consciousness and the brain: deciphering how the brain codes our thoughts*. New York: Viking Penguin, 2014.

DEHAENE, S. *How we learn: why brains learn better than any machine... for now*. New York: Viking Penguin, 2020.

DEHAENE, S.; KERSZBERG, M.; CHANGEUX, J.-P. A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences*, v. 95, n. 24, p. 14529–14534, Nov. 1998. DOI 10.1073/pnas.95.24.14529. Disponível em: <https://www.pnas.org/doi/full/10.1073/pnas.95.24.14529>. Acesso em: 4 mar. 2025.

DENIZ F., NUNEZ-ELIZALDE A. O., HUTH A. G, GALLANT J. L. The Representation of Semantic Information Across Human Cerebral Cortex During Listening Versus Reading Is Invariant to Stimulus Modality. *The Journal of Neuroscience*, v. 39, n. 39, p. 7722–7736, Sept. 2019. DOI 10.1523/JNEUROSCI.0675-19.2019. Disponível em: <https://www.jneurosci.org/content/39/39/7722.long>. Acesso em: 22 fev. 2026.

DIAS, N. M.; LEÓN, C. B. R.; PAZETO, T. de C. B.; MARTINS, G. L. L.; PEREIRA, A. P. P.; SEABRA, A. G. Avaliação da Leitura no Brasil: Revisão da Literatura no Recorte 2009-2013. *Psicologia - Teoria e Prática*, São Paulo, v. 18, n. 1, p. 113–128, abr. 2016. DOI 10.15348/1980-6906/psicologia.v18n1p113-128. Disponível em: <https://www.redalyc.org/articulo.oa?id=193846361009>. Acesso em: 1 abr. 2025.

DIONÍSIO, A. P. Análise da conversação. In: MUSSALIM, F.; BENTES, A. C. (org.) *Introdução à linguística: fundamentos epistemológicos*. v.2. 4. ed. São Paulo: Cortez, 2004. cap. 3, p. 69-100.

EBERHARD, D. M.; SIMONS, G. F.; FENNIG, C. D. (ed.). *Ethnologue: Languages of the World*. 28. ed. Dallas, Texas: SIL International, 2025. Disponível em: https://www.ethnologue.com/faq/#node_25551. Acesso em: 28 fev. 2025.

FARMER, T. A.; MATLIN, M. W. *Cognition*. 10th ed. Hoboken, New Jersey: Wiley, 2019. 432 p.

FEI-FEI, L.; IYER, A.; KOCH, C.; PERONA, P. What do we perceive in a glance of a real-world scene? *Journal of Vision*, v. 7, n. 1, p. 10, Jan. 2007. DOI 10.1167/7.1.10. Disponível em: <https://jov.arvojournals.org/article.aspx?articleid=2192891>. Acesso em: 4 mar. 2025.

FERRARI, L. *Introdução à linguística cognitiva*. 1. ed. São Paulo: Contexto, 2011. E-book (178 p.). Disponível em: <https://plataforma.bvirtual.com.br>. Acesso em: 25 fev. 2025.

FIGUEIREDO, A. A. F. de; MINERVINO, C. A. S. M.; PEREIRA, E. E. L. D.; SILVEIRA, N. J. D. da. Compreensão leitora e estratégias de estudo: estudo correlacional com universitários. *Psicologia Argumento*, Curitiba, v. 34, n. 86, jul./set. 2016. DOI 10.7213/psicol.argum.34.086.AO03. Disponível em: <https://periodicos.pucpr.br/psicologiaargumento/article/view/18270>. Acesso em: 1 abr. 2025.

FREITAG, R. M. Ko. A sociolinguística da leitura. *Letrônica*, Porto Alegre, v. 13, n. 4, p. e37508, out. 2020. DOI 10.15448/1984-4301.2020.4.37508. Disponível em: <https://revistaseletronicas.pucrs.br/letronica/article/view/37508>. Acesso em: 20 mar. 2025.

GABRIEL, R.; MORAIS, J.; KOLINSKY, R. A aprendizagem da leitura e suas implicações sobre a memória e a cognição. *Ilha do Desterro: A Journal of English Language, Literatures in English and Cultural Studies*, Florianópolis, v. 69, n. 1, p. 61–78, jan./abr. 2016. DOI 10.5007/2175-8026.2016v69n1p61. Disponível em: <https://www.redalyc.org/articulo.oa?id=478355241005>. Acesso em: 4 mar. 2025.

GARDNER, E. P.; GARDNER, D. Codificação sensorial. In: KANDEL, E. R.; KOESTER, J. D.; MACK, S. H.; SIEGELBAUM, S. A. *Princípios de neurociência*. Porto Alegre: Artmed, 2023. cap. 17, p. 1396-1472. E-book (5325 p.). Disponível em: <https://biblioteca-a.read.garden/viewer/9786558040255/1396>. Acesso em: 21 fev. 2025.

GOMES, F.; NASCIMENTO, G. P. do; LOPES, I. de A. Criação e uso de sinais compostos por surdos: um estudo com alunos do curso Letras-Libras da UFPI. *Signo*, Santa Cruz do Sul, v. 48, n. 93, p. 130–149, nov. 2023. DOI 10.17058/signo.v48i93.18661. Disponível em: <https://online.unisc.br/seer/index.php/signo/article/view/18661>. Acesso em: 20 mar. 2025.

HARVEY, H.; WALKER, R. Reading comprehension and its relationship with working memory capacity when reading horizontally scrolling text. *Quarterly Journal of Experimental Psychology*, v. 71, n. 9, p. 1887–1897, Sept. 2018. DOI

10.1080/17470218.2017.1363258. Disponível em:
<https://journals.sagepub.com/doi/10.1080/17470218.2017.1363258>. Acesso em: 1 abr. 2025.

HOCKETT, C. F. The Origin of Speech. *Scientific American*, v. 203, n. 3, p. 88–96, Sept. 1960. DOI 10.1038/scientificamerican0960-88. Disponível em:
<https://web.stanford.edu/class/linguist197a/hockett60sciam.pdf>. Acesso em: 5 mar. 2025.

HOLMQVIST, K.; NYSTRÖM, M.; ANDERSSON, R.; DEWHURST, R., JARODZKA, H.; VAN DE WEIJER, J. *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press, 2015.

HOOVER, W. A.; GOUGH, P. B. The simple view of reading. *Reading and Writing*, v. 2, n. 2, p. 127–160, June 1990. DOI 10.1007/BF00401799. Disponível em:
<https://psycnet.apa.org/record/1992-04979-001>. Acesso em: 19 fev. 2026.

ILARI, R. O estruturalismo linguístico: alguns caminhos. In: MUSSALIM, F.; BENTES, A. C. (org.) *Introdução à linguística: fundamentos epistemológicos*. v.3. 5. ed. São Paulo: Cortez, 2011. cap. 2, p. 53-92.

INSIDE OUT. Intérpretes: Sleepermane e Sling Dilly. Compositores: Tim Drullman e Wilco ter Horst. In: *Cozy winter*. Paris: Lofi Records, 2021. Streaming de música Spotify, faixa 19 (2 min 7 s). Disponível em:
<https://open.spotify.com/intl-pt/album/0DqDy8wMAYvjuCcSUs6TMF?highlight=spotify:track:6KvGSPNhfNyWN99MOJ1TC>. Acesso em: 30 abr. 2025.

IZQUIERDO, I. *Memória*. 3. ed. Porto Alegre: Artmed, 2018. *E-book* (181 p.). Disponível em: <https://biblioteca-a.read.garden/viewer/9788582714928/1>. Acesso em: 4 mar. 2025.

IZQUIERDO, I.; BARROS, D. M.; MELLO E SOUZA, T.; SOUZA, M. M. de; IZQUIERDO, L. A.; MEDINA, J. H. Mechanisms for memory types differ. *Nature*, v. 393, n. 6686, p. 635–636, June 1998. DOI 10.1038/31371. Disponível em:
<https://www.nature.com/articles/31371>. Acesso em: 4 mar. 2025.

JAMES, W. *The principles of psychology*. v.1. New York: Henry Holt and Company, 1931.

JIANG, J.; ZHOU, L.; RICKSON, D.; JIANG, C. The effects of sedative and stimulative music on stress reduction depend on music preference. *Arts in Psychotherapy*, v. 40, n. 2, p. 201-205, Apr. 2013. DOI 10.1016/j.aip.2013.02.002. Disponível em:
<https://www.sciencedirect.com/science/article/abs/pii/S0197455613000725?via%3Dihub>. Acesso em: 2 fev. 2025.

KAPLAN, G. Animal communication. *WIREs Cognitive Science*, v. 5, n. 6, p. 661–677, Nov. 2014. DOI 10.1002/wcs.1321. Disponível em:
<https://wires.onlinelibrary.wiley.com/doi/10.1002/wcs.1321>. Acesso em: 5 mar. 2025.

KINTSCH, W. *Comprehension: a paradigm for cognition*. New York: Cambridge University Press, 1998.

KINTSCH, W. An overview of top-down and bottom-up effects in comprehension: the CI perspective. *Discourse processes*, [S. l.], v. 39, n. 2&3, p. 125-128, 2005.

Disponível em:

<https://condor.depaul.edu/dallbrit/extra/hon207/readings/kintsch-2005-overview-of-top-down-and-bottom-up-effects-CI.pdf>. Acesso em: 18 mar. 2025.

KINTSCH, W.; RAWSON, K. A. Comprehension. In: SNOWLING, M. J.; HULME, C. (ed.). *The Science of Reading: a Handbook*. Oxford, UK: Blackwell Publishing, 2005. cap. 12, p. 211–226. DOI 10.1002/9780470757642.ch12. Disponível em:

<https://onlinelibrary.wiley.com/doi/10.1002/9780470757642.ch12>. Acesso em: 20 mar. 2025.

KOCH, I. V.; CUNHA-LIMA, M. L. Do cognitivismo ao sociocognitivismo. In: MUSSALIM, F.; BENTES, A. C. (org.) *Introdução à linguística: fundamentos epistemológicos*. v.3. 5. ed. São Paulo: Cortez, 2011. cap. 7, p. 251-300.

KOLINSKY, R.; MORAIS, J. The worries of wearing literate glasses. *L'Année psychologique*, v. Vol. 118, n. 4, p. 321–347, 22 Feb. 2018. DOI 10.3917/anpsy1.184.0321. Disponível em:

<https://shs.cairn.info/revue-l-annee-psychologique-2018-4-page-321?lang=en>. Acesso em: 19 fev. 2026.

LAMERS, J.; VAN DER MEER, T.; TESTERINK, C. How Plants Sense and Respond to Stressful Environments. *Plant Physiology*, v. 182, n. 4, p. 1624–1635, Apr. 2020. DOI 10.1104/pp.19.01464. Disponível em:

<https://academic.oup.com/plphys/article/182/4/1624/6116416?login=false>. Acesso em: 5 mar. 2025.

LAYER, J. *Principles of phonetics*. Cambridge: Cambridge University Press, 1994.

LENT, R. *Cem bilhões de neurônios: conceitos fundamentais de neurociência*. 2. ed. São Paulo: Atheneu, 2010.

LIVINGSTONE, S. Audiences and interpretations. *E-Compós*, Brasília, v. 10, jun. 2008. DOI 10.30962/ec.185. Disponível em:

<https://www.e-compos.org.br/e-compos/article/view/185>. Acesso em: 20 mar. 2025.

LOFI GIRL. *Lofi hip hop radio: beats to relax/study to*. Paris: Lofi Girl, 2022. 1 vídeo (12 h). Disponível em: <https://www.youtube.com/watch?v=jfKfPfyJRdk>. Acesso em: 30 abr. 2025.

LOPES, E. *Fundamentos da linguística contemporânea*. 19. ed. São Paulo: Cultrix, 2000.

LUSTOSA, S. S. *Análise dos níveis de letramento de ingressantes e concluintes do ensino superior: estudo de caso*. 2014. Tese (Doutorado) - Programa de Mestrado e Doutorado em Distúrbios da Comunicação Humana, Universidade Tuiuti do Paraná,

Curitiba, 2014. Disponível em: <https://tede.utp.br/jspui/handle/tede/1503>. Acesso em: 1 abr. 2025.

MAYER, R. E.; MORENO, R. A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, v. 90, n. 2, p. 312–320, June 1998. DOI 10.1037/0022-0663.90.2.312. Disponível em: <https://psycnet.apa.org/record/1998-02710-011>. Acesso em: 3 mar. 2025.

MENGARDA, E. J. *Fundamentos da linguística*. Indaial: Uniasselvi, 2012.

MERTZANI, M.; DE MONTE, M. T.; FERNANDES, S. de F. Apresentação. *Signo*, Santa Cruz do Sul, v. 48, n. 93, p. 1–8, nov. 2023. Disponível em: <https://online.unisc.br/seer/index.php/signo/article/view/18888>. Acesso em: 20 mar. 2025.

MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, v. 63, n. 2, p. 81–97, Mar. 1956. DOI 10.1037/h0043158. Disponível em: <https://psycnet.apa.org/record/1957-02914-001>. Acesso em: 4 mar. 2025.

MORAIS, J. O que faz a diferença entre a linguagem rica e a linguagem pobre? *Signo*, v. 44, n. 81, p. 2–21, set. 2019. DOI 10.17058/signo.v44i81.14574. Disponível em: <https://online.unisc.br/seer/index.php/signo/article/view/14574>. Acesso em: 19 fev. 2025.

MOUSAVI, S. Y.; LOW, R.; SWELLER, J. Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, v. 87, n. 2, p. 319–334, June 1995. DOI 10.1037/0022-0663.87.2.319. Disponível em: <https://psycnet.apa.org/record/1995-38914-001>. Acesso em: 3 mar. 2025.

NÖTH, W. *Panorama da Semiótica: de Platão a Peirce*. 4. ed. São Paulo: Annablume, 2003.

OLIVEIRA, R. M. O conceito de executivo central e suas origens. *Psicologia: Teoria e Pesquisa*, v. 23, n. 4, p. 399–406, dez. 2007. DOI 10.1590/S0102-37722007000400005. Disponível em: <https://www.scielo.br/j/ptp/a/zSHcVx5nrswZSmfYqr6564j/?lang=pt>. Acesso em: 4 mar. 2025.

OpenAI. Geração de textos narrativos curtos a partir de exemplos apresentados. ChatGPT-4o versão de 29 mar. 2025a. Inteligência Artificial. Disponível em: <https://chatgpt.com/>. Acesso em: 30 abr. 2025.

OpenAI. Geração de textos narrativos curtos a partir de exemplos apresentados. ChatGPT-4o versão de 30 mar. 2025b. Inteligência Artificial. Disponível em: <https://chatgpt.com/>. Acesso em: 30 abr. 2025.

PAGE, M. P. A.; NORRIS, D. G. The Irrelevant Sound Effect: What Needs Modelling, and a Tentative Model. *The Quarterly Journal of Experimental Psychology Section A*, v. 56, n. 8, p. 1289–1300, Nov. 2003. DOI 10.1080/02724980343000233. Disponível

em: <https://journals.sagepub.com/doi/10.1080/02724980343000233>. Acesso em: 5 mar. 2025.

PERHAM, N.; CURRIE, H. Does listening to preferred music improve reading comprehension performance? *Applied Cognitive Psychology*, v. 28, n. 2, p. 279-284, Jan. 2014. DOI 10.1002/acp.2994. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/acp.2994>. Acesso em: 19 fev. 2025.

PETERSEN, S. E.; POSNER, M. I. The Attention System of the Human Brain: 20 Years After. *Annual Review of Neuroscience*, v. 35, n. 1, p. 73–89, July 2012. DOI 10.1146/annurev-neuro-062111-150525. Disponível em: <https://www.annualreviews.org/content/journals/10.1146/annurev-neuro-062111-150525>. Acesso em: 4 mar. 2025.

POSNER, M. I.; PETERSEN, S. E. The Attention System of the Human Brain. *Annual Review of Neuroscience*, v. 13, n. 1, p. 25–42, Mar. 1990. DOI 10.1146/annurev.ne.13.030190.000325. Disponível em: <https://www.annualreviews.org/content/journals/10.1146/annurev.ne.13.030190.000325>. Acesso em: 4 mar. 2025.

RAYNER, K.; SCHOTTER, E. R.; MASSON, M. E. J.; POTTER, M. C.; TREIMAN, R. So Much to Read, So Little Time. *Psychological Science in the Public Interest*, [S.l.], v. 17, n. 1, p. 4–34, May 2016. DOI 10.1177/1529100615623267. Disponível em: <https://journals.sagepub.com/doi/10.1177/1529100615623267>. Acesso em: 20 mar. 2025.

RIBEIRO, A. E. Multimodalidade e produção de textos: questões para o letramento na atualidade. *Signo*, Santa Cruz do Sul, v. 38, n. 64, p. 21–34, jan./jun. 2013. DOI 10.17058/signo.v38i64.3714. Disponível em: <https://online.unisc.br/seer/index.php/signo/article/view/3714>. Acesso em: 27 nov. 2024.

RIBEIRO, S. *O oráculo da noite: a história e a ciência do sonho*. São Paulo: Companhia das Letras, 2019.

RIESS, A. B.; GABRIEL, R. Perspectivas de pesquisa com a metodologia de Eye tracking em leitura e cognição no Brasil. *Revista Odisseia*, Natal, v. 5, n. 1, p. 103-119, jan./jun. 2020. DOI 10.21680/1983-2435.2020v5n1ID19449. Disponível em: <https://periodicos.ufrn.br/odisseia/article/view/19449>. Acesso em: 18 mar. 2025.

SALAMÉ, P.; BADDELEY, A. Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology Section A*, v. 41, n. 1, p. 107-122, Feb. 1989. DOI 10.1080/14640748908402355. Disponível em: <https://journals.sagepub.com/doi/10.1080/14640748908402355>. Acesso em: 19 fev. 2025.

SALLES, J. F. de; PARENTE, M. A. de M. P. Processos cognitivos na leitura de palavras em crianças: relações com compreensão e tempo de leitura. *Psicologia: Reflexão e Crítica*, Porto Alegre, v. 15, n. 2, p. 321–331, 2002. DOI 10.1590/S0102-79722002000200010. Disponível em:

<https://www.scielo.br/j/prc/a/3FcxqfWxLMVq3c83Gs3q8mJ/?lang=pt>. Acesso em: 20 mar. 2025.

SANTAELLA, L. KAUFMAN, D. A inteligência artificial generativa como quarta ferida narcísica do humano. *Matrizes*, São Paulo, v. 18, n. 1, p. 37–53, jan./abr. 2024. DOI 10.11606/issn.1982-8160.v18i1p37-53. Disponível em: <https://www.revistas.usp.br/matrizes/article/view/210834/204171>. Acesso em: 1 abr. 2025.

SANTHOSH, J.; JAGTAP, S.; DENGEL, A.; ISHIMARU, S. Digital Distractions in Reading: Investigating Impact of Cognitive Control Training on Reading Behavior and Outcomes. In: UBICOMP '24: COMPANION OF THE 2024 ON ACM INTERNATIONAL JOINT CONFERENCE ON PERVASIVE AND UBIQUITOUS COMPUTING. *Anais[...]*. New York, NY, USA: ACM, Oct. 2024. DOI 10.1145/3675094.3677591. Disponível em: <https://dl.acm.org/doi/10.1145/3675094.3677591>. Acesso em: 3 mar. 2025.

SARDELICH, M. E. Leitura de imagens, cultura visual e prática educativa. *Cadernos de Pesquisa*, São Paulo, v. 36, n. 128, p. 451–472, maio/ago. 2006. DOI 10.1590/S0100-15742006000200009. Disponível em: <https://publicacoes.fcc.org.br/cp/article/view/405>. Acesso em: 20 mar. 2025.

SAUSSURE, F. de. *Curso de linguística geral*. Tradução: Antônio Chelini, José Paulo Paes e Izidoro Blikstein. 27. ed. São Paulo: Cultrix, 2006.

SCARBOROUGH, H. S. Connecting early language and literacy to later reading (dis)abilities: evidence, theory, and practice. In: NEUMAN, S.; DICKINSON, D. (ed.), *Handbook for research in early literacy*. New York: Guildford Press, 2001. cap. 8, p. 97-110.

SCHÖN, D.; BOYER, M.; MORENO, S.; BESSON, M.; PERETZ, I.; KOLINSKY, R. Song as an aid for language acquisition. *Cognition*, Amsterdam, v. 106, n. 2, p. 975–983, Feb. 2008. DOI 10.1016/j.cognition.2007.03.005. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0010027707000868?via%3Dihub>. Acesso em: 5 mar. 2025.

SIMOS, P. G.; BREIER, J. I.; FLETCHER, J. M.; FOORMAN, B. R.; CASTILLO, E. M.; PAPANICOLAOU, A. C. Brain Mechanisms for Reading Words and Pseudowords: an Integrated Approach. *Cerebral Cortex*, Oxford, v. 12, n. 3, p. 297–305, Mar. 2002. DOI 10.1093/cercor/12.3.297. Disponível em: <https://academic.oup.com/cercor/article-abstract/12/3/297/278132?redirectedFrom=fulltext&login=false>. Acesso em: 20 mar. 2025.

SMALL, D. M.; PRESCOTT, J. Odor/taste integration and the perception of flavor. *Experimental Brain Research*, v. 166, n. 3–4, p. 345–357, Oct. 2005. DOI 10.1007/s00221-005-2376-9. Disponível em: <https://link.springer.com/article/10.1007/s00221-005-2376-9>. Acesso em: 5 mar. 2025.

SOUSA, L. B.; HÜBNER, L. C. Desafios na avaliação da compreensão leitora: demanda cognitiva e leitura textual. *Neuropsicologia Latinoamericana*, Santiago, Chile, v. 7, n. 1, p. 34–46, 2015. DOI 10.5579/rnl.2013.0237. Disponível em: <https://pepsic.bvsalud.org/pdf/rnl/v7n1/v7n1a04.pdf>. Acesso em: 20 mar. 2025.

STEINBERG, M. *Os elementos não-verbais da conversação*. São Paulo: Atual, 1988.

TAYLOR, W. L. “Cloze Procedure”: A New Tool for Measuring Readability. *Journalism Quarterly*, v. 30, n. 4, p. 415–433, Sept. 1953. DOI 10.1177/107769905303000401. Disponível em: <https://journals.sagepub.com/doi/10.1177/107769905303000401>. Acesso em: 19 fev. 2026.

TOMITCH, L. M. B. A metodologia da pesquisa em leitura: das perguntas de compreensão à ressonância magnética funcional. In: TOMITCH, L. M. B. (org.). *Aspectos cognitivos e instrucionais da leitura*. Bauru: EDUSC, 2008. p. 37-56.

TSUCHIYA, N.; KOCH, C. The Relationship Between Consciousness and Attention. In: LAUREYS, S.; TONONI, G. (ed.). *The Neurology of Consciousness: Cognitive Neuroscience and Neuropathology*. San Diego, CA: Academic Press, 2009. p. 63–77. DOI 10.1016/B978-0-12-374168-4.00006-X. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/B978012374168400006X?via%3Dihub>. Acesso em: 4 mar. 2025.

UNESCO. *International Literacy Day 2024: concept note*. Paris: UNESCO, 2024. Disponível em: <https://www.unesco.org/sites/default/files/medias/fichiers/2024/09/ild-2024-cn-en.pdf>. Acesso em: 20 mar. 2025.

VAN GAAL, S. RIDDERINKHOF, K.; SCHOLTE, H.; LAMME, V. Unconscious Activation of the Prefrontal No-Go Network. *The Journal of Neuroscience*, v. 30, n. 11, p. 4143–4150, Mar. 2010. DOI 10.1523/JNEUROSCI.2992-09.2010. Disponível em: <https://www.jneurosci.org/content/30/11/4143>. Acesso em: 4 mar. 2025.

VASILEV, M. R.; KIRKBY, J. A.; ANGELE, B. Auditory distraction during reading: a Bayesian meta-analysis of a continuing controversy. *Perspectives on Psychological Science*, v. 13, n. 5, p. 567-597, Sept. 2018. DOI 10.1177/1745691617747398. Disponível em: <https://journals.sagepub.com/doi/10.1177/1745691617747398>. Acesso em: 19 fev. 2025.

WOLF, M. *O cérebro no mundo digital: os desafios da leitura na nossa era*. 1. ed. Tradução: Rodolfo Ilari, Mayumi Ilari. São Paulo: Contexto, 2019. E-book (258 p.). Disponível em: <https://plataforma.bvirtual.com.br/Acervo/Publicacao/173192>. Acesso em: 20 mar. 2025.

As páginas 260 a 278, referentes aos Apêndices A a G, foram omitidas desta versão da dissertação. Para ter acesso a elas, entre em contato com o autor pelo e-mail jordanjunges@mx2.unisc.br.