

ESCOLA DE HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
MESTRADO EM LETRAS

VITOR ROMÁRIO MONTICELLI GARCIA

**O IMPACTO DA ESCOLARIDADE E DA DOENÇA DE ALZHEIMER NA
PRODUÇÃO ORAL DE NARRATIVA NO ENVELHECIMENTO: ESTUDO DA
MICRO E MACROESTRUTURA COM ANÁLISE AUTOMÁTICA DE TEXTOS**

Porto Alegre, RS
2019

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica
do Rio Grande do Sul

VITOR ROMÁRIO MONTICELLI GARCIA

**O IMPACTO DA ESCOLARIDADE E DA DOENÇA DE ALZHEIMER NA
PRODUÇÃO ORAL DE NARRATIVA NO ENVELHECIMENTO: ESTUDO DA
MICRO E MACROESTRUTURA COM ANÁLISE AUTOMÁTICA DE TEXTOS**

Dissertação apresentada como requisito para obtenção do grau de Mestre em Letras pelo Programa de Pós-Graduação em Letras da Escola de Humanidades da Pontifícia Universidade Católica do Rio Grande do Sul. Área de concentração: Linguística.

Orientadora: Prof^a Dr^a Lílian Cristine Hübner

Porto Alegre, RS

2019

VITOR ROMÁRIO MONTICELLI GARCIA

**O IMPACTO DA ESCOLARIDADE E DA DOENÇA DE ALZHEIMER NA
PRODUÇÃO ORAL DE NARRATIVA NO ENVELHECIMENTO: ESTUDO DA
MICRO E MACROESTRUTURA COM ANÁLISE AUTOMÁTICA DE TEXTOS**

Dissertação apresentada como requisito para obtenção do grau de Mestre em Letras pelo Programa de Pós-Graduação em Letras da Escola de Humanidades da Pontifícia Universidade Católica do Rio Grande do Sul. Área de concentração: Linguística.

Apresentada em: 09 de janeiro de 2019.

BANCA EXAMINADORA:

Profª Drª Cláudia Finger-Kratochvil - UFFS

Prof. Dr. Augusto Buchweitz - PUCRS

Profª Drª LÍlian Cristine Hübner (Presidente) - PUCRS

Porto Alegre, RS
2019

Ficha Catalográfica

G216i Garcia, Vítor Romário Monticelli

O impacto da escolaridade e da Doença de Alzheimer na produção oral de narrativa no envelhecimento : estudo da micro e macroestrutura com análise automática de textos / Vítor Romário Monticelli Garcia . – 2019.

117 p.

Dissertação (Mestrado) – Programa de Pós-Graduação em Letras, PUCRS.

Orientadora: Profa. Dra. Lilian Cristine Hübner.

1. Envelhecimento. 2. Doença de Alzheimer. 3. Escolaridade. 4. Narrativa. 5. Análise automática de textos. I. Hübner, Lilian Cristine. II. Título.

À minha família, suporte para todas as horas.

AGRADECIMENTOS

- À minha família (pais e irmãos) pelo apoio, amor, respeito e paciência incondicionais.
- Ao meu companheiro Robson por todas as vezes que me incentivou a concluir este trabalho.
- Aos meus avós pelo carinho e por me ensinarem a amar e respeitar a sabedoria de quem trilhou este caminho antes de mim.
- Aos meus queridos amigos pelos encontros regados a risadas e energias positivas ao longo destes últimos dois anos. Em especial para Gabriela, minha companheira de todas as horas.
- À minha querida orientadora Prof^a Dra. Lilian Cristine Hübner por seu apoio e orientação na conclusão dos estudos desta dissertação, e por ser um exemplo de carinho e afetividade no meio acadêmico.
- Aos colegas do Grupo de Estudos em Neurolinguística e Psicolinguística (GENP) pelo apoio com as coletas, organização de banco de dados, e incentivo à pesquisa. Ah, e também pelas risadas.
- À Coordenação do Programa de Pós-Graduação em Letras da Pontifícia Universidade Católica do Rio Grande do Sul, pelo suporte e excelência acadêmica.
- Aos professores Dr. Augusto Buchweitz, Dr. Márcio Leitão e Dr^a Cláudia Finger-Kratochvil pela participação nas bancas de qualificação e final e suas contribuições para este estudo.
- Ao Prof. Sérgio Kato pela assessoria estatística na análise dos dados aqui apresentados.
- A Leandro Borges, Marcos Treviso e Sandra Aluísio pelo apoio com as análises de métricas e segmentação de sentenças.
- Aos participantes desta pesquisa, que gentilmente cederam seu tempo, seu interesse e suas histórias para que este trabalho fosse possível.
- Ao Universo, pela oportunidade de estar vivo, estudando, e buscando aprimorar ainda mais minhas habilidades como profissional.
- O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal nível Superior – Brasil (CAPES) – Código de Financiamento 001, à qual também dedico meus sinceros agradecimentos.

“Viver é envelhecer. Nada mais.”

Simone de Beauvoir

RESUMO

A narrativa é um tipo textual bastante recorrente (JERÔNIMO; HÜBNER, 2014). Produzir um texto é uma tarefa que demanda planejamento cognitivo estratégico, na construção de um modelo cognitivo inicialmente recrutado da memória episódica (ME) (VAN DIJK, 2010). Na Doença de Alzheimer (DA), as habilidades linguísticas discursivas são afetadas desde o início do desenvolvimento da demência (SOARES; BRANDÃO; LACERDA, 2012). A análise automática do discurso tem potencial para distinguir grupos clínicos de grupos controles saudáveis, auxiliando no diagnóstico precoce (TOLEDO, et. al., 2017). A escolaridade é uma variável que também pode interferir na qualidade da produção de textos (JERÔNIMO, 2016; STERN, 2012). Sendo assim, esta dissertação investiga se há diferenças na produção oral de narrativas na comparação entre idosos saudáveis segundo a escolaridade, bem como entre saudáveis e acometidos por DA nos níveis micro- e macroestruturais, correlacionando-se esses dados a construtos de memória de trabalho (MT) e episódica (ME). Ainda, utiliza-se da análise automática de textos para uma avaliação complementar dos dados. Para tanto, aborda dois estudos: o Estudo 1 compara grupos de idosos saudáveis de alta e baixa escolaridade na produção dos níveis da micro- e macroestrutura das narrativas, correlacionando-os com o desempenho dos construtos da ME e da MT, junto com uma análise automática de dados relativos à microestrutura. Compõem o estudo um grupo de 28 idosos saudáveis (média de idade = 69,46) de alta escolaridade (≥ 9 anos de escolaridade) (SAE), e outro grupo de 35 idosos saudáveis (média de idade = 72,20) de baixa escolaridade (entre 2 e 8 anos de escolaridade) (SBE); o Estudo 2 investiga as mesmas relações do Estudo 1, porém pelo viés da variável demência. Compõem o estudo um grupo controle (GC) de baixa escolaridade (SBE no Estudo 1), e outro grupo de 12 idosos acometidos pela DA (média de idade = 74,58) de baixa escolaridade (DA). A tarefa para coleta de dados linguísticos é uma produção de narrativa a partir da sequência de figuras *The Dog Story* (LE BOEUF, 1976). As tarefas neuropsicológicas que medem a capacidade de ME e MT são, respectivamente, Tarefa de Aprendizagem Verbal da BALE (Bateria de Avaliação da Linguagem no Envelhecimento) (HUBNER et al, no prelo, adaptado de BUSCHKE, GROBER, 2012) e o Span de Dígitos (subteste do NEUPSILIN, de FONSECA et al., 2009). Para análise de microestrutura, 23 métricas linguísticas foram extraídas do software Coh-Matrix Dementia (ALUÍSIO; CUNHA; SCARTON, 2016). Para análise de macroestrutura utiliza-se o protocolo da BALE. São realizadas análises quantitativas e qualitativas. Os resultados principais indicam influência da escolaridade no desempenho da tarefa linguística, especialmente na construção da macroestrutura, sendo que idosos menos escolarizados demonstram falhas na construção da sequência narrativa, inclusão de informações extra-textuais, entre outras. Não houve diferenças significativas nos testes neuropsicológicos na distinção entre grupos saudáveis, porém correlacionaram positivamente com aspectos da construção da micro e macroestrutura das narrativas. Na comparação entre o grupo DA e GC, houve diferença no desempenho de ME, apontando declínio neste construto relacionado à demência. As principais diferenças linguísticas entre grupos GC e DA ficam por conta da produção de fragmentos de sentenças, maior utilização de palavras de conteúdo do que funcionais, maior número de pausas longas e hesitações no texto de DA. Espera-se que estes dados fomentem o uso da avaliação do discurso e o aprimoramento de ferramentas para sua análise para um diagnóstico precoce de quadros demenciais, auxiliando na melhora da qualidade de vida dos idosos.

Palavras-chave: Envelhecimento; Doença de Alzheimer; Produção oral de narrativa; Análise Automática de Texto; Escolaridade; Memória.

ABSTRACT

Narrative is a frequent type of text (JERÔNIMO; HÜBNER, 2014). Producing a text is a strategic task of cognitive planning, which is based on the construction of a cognitive model of the text in episodic memory (EM) (VAN DIJK, 2010). In Alzheimer's disease (AD), discursive linguistic abilities are impaired since the very beginning of the continuum of the dementia (SOARES; BRANDÃO; LACERDA, 2012). In this sense, automated speech analysis has the potential to distinguish between healthy and cognitively impaired groups, helping in the establishment of early diagnosis (TOLEDO, et. al., 2017). The level of education can also interfere in the quality of text production (JERÔNIMO, 2016; STERN, 2012). Hence, this master's thesis investigates whether there are differences in oral production of micro and macrostructure of narratives in the comparison between groups of healthy and AD diagnosed elders. It also correlates working (WM) and episodic memory (EM) and linguistic data for the groups. To complement linguistic data, automatic text analysis is conducted in this study. To do so, this thesis is divided in two studies: Study 1 compares the production of micro and macrostructural levels of a narrative made by two groups of healthy elders (divided into high and low educational levels), correlating linguistic data with scores from WM and EM tasks. The performances of 28 high educational level (≥ 9 years of education) healthy elders (mean age = 69,46) (HHE), and 35 low educational level (between 2 and 8 years of education) healthy elders (mean age = 72,20) (HLE) are analyzed; Study 2 investigates the same comparisons reported in Study 1, but focusing on dementia. A low educational level control group (the same one described in Study 1) (HLE, now CG) is compared to a group of 12 low educational level AD-diagnosed elders (mean age = 74,58) (DA). The linguistic task proposed is an oral production of a narrative based on a sequence of pictures, *The Dog Story* (LE BOEUF, 1976). The neuropsychological tasks that measure EM and WM are, respectively, Verbal Learning Task, from BALE (*Bateria de Avaliação da Linguagem no Envelhecimento*) (HÜBNER, et. al., accepted for publication, adapted from BUSCHKE, GROBER, 2012) and *Digit Span* (subtest from NEUPSILIN, by FONSECA et al., 2009). To the analysis of the microstructure, 23 linguistic features are extracted from the software Coh-Metrix Dementia (ALUÍSIO; CUNHA; SCARTON, 2016). To the analysis of macrostructure the protocol from BALE is adopted. Both quantitative and qualitative analyses are conducted. Main results indicate influence of educational level on performance in the linguistic task, especially in the construction of the macrostructure, which means that low educational level elders had problems with the construction of the different steps of the narrative, included extra information, among other aspects. There were no significant differences between healthy participants in the neuropsychological tasks, even though positive correlations were identified between them and the construction of micro and macrostructure. Differences were also identified in the comparison between CG and AD in the EM task, which may indicate impairment in EM due to dementia. The main linguistic differences between CG and AD are related to the production of fragments of a sentence, higher incidence of content words rather than function words, higher incidence of pauses and hesitations in the narratives produced by ADs. It is expected that the data presented here contribute to foster the use of discourse in AD assessment, as well as to enhance automated speech analysis tools, making it possible to early diagnose dementias and improve life quality to the elderly.

Keywords: Aging; Alzheimer's Disease; Narrative Oral Production; Automated Speech Analysis; Educational level; Memory.

LISTA DE ILUSTRAÇÕES

Figura 1: The Dog Story.....	59
Figura 2: Modelo de transcrição e segmentação de texto.....	62
Figura 3: Exemplo de transcrição do grupo SAE.....	71
Figura 4: Exemplo de transcrição do grupo SBE.....	72
Figura 5: Exemplo de transcrição do grupo SBE (pausas e emissões vazias).....	76
Figura 6: Exemplo de transcrição do grupo SAE (microestrutura).....	77
Figura 7: Exemplo de transcrição do grupo DA (macroestrutura).....	78
Figura 8: Exemplo de transcrição do grupo DA (descrição de cena).....	79
Figura 9: Exemplo de transcrição do grupo DA (palavras de conteúdo).....	82
Figura 10: Exemplo de transcrição do grupo DA (pausas).....	82
Figura 11: Diferença transcrição X segmentação.....	90
Quadro 1: Descrição das 73 métricas extraídas pelo software Coh-Metrix Dementia.....	39
Quadro 2: Resumo de estudos que abordam a análise automática de textos.....	49
Quadro 3: Critérios de avaliação da tarefa de produção oral de narrativa.....	60
Quadro 4: Resumo da discussão das hipóteses.....	100

LISTA DE TABELAS

Tabela 1: Dados sociodemográficos e cognitivos dos participantes do Estudo 1	65
Tabela 2: Dados sociodemográficos e cognitivos dos participantes do Estudo 2	66
Tabela 3: Dados neuropsicológicos dos participantes do Estudo 1	67
Tabela 4: Dados neuropsicológicos dos participantes do Estudo 2	69
Tabela 5: Dados de macroestrutura e desempenho linguístico dos participantes do Estudo 1	71
Tabela 6: Dados de microestrutura dos participantes do Estudo 1	73
Tabela 7: Dados de macroestrutura e desempenho linguístico dos participantes do Estudo 2	78
Tabela 8: Dados de microestrutura dos participantes do Estudo 2	80
Tabela 9: Resultados de correlação de dados neuropsicológicos x macroestrutura	83
Tabela 10: Resultados de correlação de dados neuropsicológicos x microestrutura	84

LISTA DE ABREVIATURAS E SIGLAS

- ABEP** – Associação Brasileira de Empresas e Pesquisa
- APP** – Afasia Progressiva Primária
- APP-vnf** – Afasia Progressiva Primária – variante não-fluente
- APP-vs** – Afasia Progressiva Primária – variante semântica
- BALE** – Bateria de Avaliação da Linguagem no Envelhecimento
- BDAE** – *Boston Diagnostic Aphasia Examination*
- CCL** – Comprometimento Cognitivo Leve
- CDR** – *Clinical Dementia Rating*
- DA** – Doença de Alzheimer
- DeepBonDD** – *a Deep neural approach to Boundary and Disfluency Detection*
- DFT** – Demência Frontotemporal
- DR** – Difusividade Radial
- DS** – Demência Semântica
- DTI** – *Diffusion Tensor Imaging*
- FCSRT** – *Free and Cued Selective Reminding Test*
- FCSRT+IR** – *Free and Cued Selective Reminding Test with Immediate Recall*
- FES** – Funções executivas
- FLI** – Fascículo longitudinal inferior
- FLS** – Fascículo longitudinal superior
- GC** – Grupo controle
- GDS** – *Geriatric Depression Scale*
- HAROLD** – *Hemispheric Asymmetry Reduction in Older Adults*
- ME** – Memória episódica
- MEEM** – Mini-Exame do Estado Mental
- MRI** – Imagem por Ressonância Magnética
- MS** – Memória semântica
- MT** – Memória de trabalho
- NEUPSILIN** – Instrumento de Avaliação Neuropsicológica Breve
- PLN** – Programação de Linguagem Natural
- SAE** – Grupo Saudável de Alta Escolaridade
- SBE** – Grupo Saudável de Baixa Escolaridade
- SPSS** – *Statistical Package for the Social Sciences*

STAC-r – *Revised Model of the Scaffolding Theory of Aging and Cognition*

TCLE – Termo de Consentimento Livre e Esclarecido

VBM – *Voxel-Based Morphometry*

SUMÁRIO

1 INTRODUÇÃO	14
2 FUNDAMENTAÇÃO TEÓRICA.....	18
2.1 PRODUÇÃO DE NARRATIVAS NO ENVELHECIMENTO SAUDÁVEL E SUA RELAÇÃO COM A ESCOLARIDADE	18
2.1.1 Modelo situacional de processamento e produção textual	21
2.1.2 A produção de narrativas no envelhecimento saudável e a influência da escolaridade	25
2.2 PRODUÇÃO E ANÁLISE DE NARRATIVA NA DOENÇA DE ALZHEIMER E SUA RELAÇÃO COM CONSTRUTOS COGNITIVOS	29
2.2.1 Doença de Alzheimer e memória: aspectos das memórias episódica e de trabalho.....	32
2.2.2 Análise automática como ferramenta para avaliação da produção de narrativas	38
3 ESTUDO EXPERIMENTAL	52
3.1 OBJETIVOS	52
3.1.1 Objetivo geral da dissertação	52
3.1.2 Objetivos específicos – Estudo 1	52
3.1.3 Objetivos específicos - Estudo 2	52
3.2 HIPÓTESES	53
3.2.1 Hipóteses - Estudo 1.....	53
3.2.2 Hipóteses – Estudo 2	53
3.3 MÉTODO	54
3.3.1 Aspectos éticos.....	54
3.3.2 Participantes	54
3.3.2.1 Critérios de inclusão e exclusão de participantes.....	55
3.3.3 Instrumentos de seleção e caracterização dos participantes dos Estudos 1 e 2.....	55
3.3.4 Instrumentos para a coleta de dados neuropsicológicos.....	57
3.3.5 Instrumento para a coleta de dados linguísticos.....	58
3.3.6 Procedimentos para análise de dados	62
4 APRESENTAÇÃO DOS RESULTADOS.....	64

4.1 DADOS SOCIODEMOGRÁFICOS E COGNITIVOS	64
4.1.1 Dados sociodemográficos e de funcionalidade dos participantes do Estudo 1	64
4.1.2 Dados sociodemográficos e de funcionalidade dos participantes do Estudo 2	66
4.2 DADOS NEUROPSICOLÓGICOS DOS PARTICIPANTES.....	67
4.2.1 Dados neuropsicológicos dos participantes do Estudo 1	67
4.2.2 Dados neuropsicológicos dos participantes do Estudo 2	69
4.3 DADOS LINGUÍSTICOS DA TAREFA DE NARRATIVA.....	70
4.3.1 Dados linguísticos dos participantes do Estudo 1	70
4.3.2 Dados linguísticos dos participantes do Estudo 2	77
4.3.3 Dados das correlações de tarefas neuropsicológicas e linguísticas.....	83
5 DISCUSSÃO DOS RESULTADOS	88
5.1 DISCUSSÃO DAS HIPÓTESES SOBRE O ESTUDO 1.....	88
5.2 DISCUSSÃO DAS HIPÓTESES SOBRE O ESTUDO 2.....	94
6 CONSIDERAÇÕES FINAIS.....	103
REFERÊNCIAS	108
ANEXOS	115
ANEXO A: Termo de Consentimento Livre e Esclarecido (TCLE)	115

1 INTRODUÇÃO

Diversas áreas têm procurado compreender os processos sociais, psicológicos e biológicos relacionados ao envelhecimento. Foi assim que a filosofia, as ciências sociais, a psicologia e a geriatria, para nomear algumas, passaram a incluir as questões relacionadas ao idoso em suas agendas (RIVERO, et. al. 2013). No Brasil, as estatísticas parecem apontar para a necessidade de intensificarem-se os estudos nessa área, devido ao crescimento no número de pessoas idosas no país (IBGE, 2015). Com o aumento da população idosa, redimensiona-se também a preocupação com os desafios relacionados ao envelhecimento. Dentre eles, destacamos o que é central a este estudo: as demências.

Demências são quadros clínicos resultantes da perda de neurônios necessários para a realização apropriada de determinadas funções cerebrais (como as memórias, por exemplo) (IZQUIERDO, 2011). Caracterizam-se principalmente por apresentar pelo menos dois déficits cognitivos e um comprometimento da vida cotidiana. Geralmente, um dos primeiros sintomas de processos demenciais é o esquecimento, seguido de falhas em alguma outra função cognitiva (que pode ser diminuição de funções executivas, linguísticas, motoras, entre outras). As demências, em geral, têm caráter neurodegenerativo, o que faz com que os comprometimentos sejam progressivos. Logo, quanto antes diagnosticados os sintomas, mais cedo é possível iniciar uma intervenção de tratamento terapêutico comportamental ou farmacológico. Alguns exemplos de demências neurodegenerativas são a Demência Vascular, a Demência com Corpos de Lewy e a mais comum delas, a Doença de Alzheimer (DA) (DE PAULA; FORLENZA, 2012).

Dentre os principais déficits cognitivos associados à DA estão os prejuízos de memória (episódica¹, principalmente), de atenção, de orientação espacial, de funções executivas frontais (planejamento, sequenciamento, abstração, tomada de decisão) e de linguagem (SOARES; SOARES; CAIXETA, 2012). Sobre esta última, o “déficit discursivo consiste no maior problema de comunicação das pessoas com DA” (SOARES; BRANDÃO; LACERDA, 2012, p. 168).

A neuropsicolinguística tem contribuído para a construção de instrumentos de avaliação de linguagem, que podem servir para o diagnóstico de afasias, demências e outras

¹ Os conceitos de memória episódica, assim como de funções executivas, serão apresentados detalhadamente no capítulo que trata da fundamentação teórica.

alterações no âmbito linguístico-cognitivo (MANSUR, 2010). Desta forma, baterias de testes neuropsicológicos e de tarefas linguísticas (desenvolvidas em projetos interdisciplinares com áreas da psicologia e fonoaudiologia, por exemplo) têm se mostrado úteis neste sentido (SOARES; BRANDÃO; LACERDA, 2012). Lefèvre e Caixeta (2012) apontam para a relevância da avaliação neuropsicológica na identificação e diagnóstico dos tipos de demência, o que acaba influenciando em um posterior tratamento mais direcionado às necessidades específicas dos avaliados.

Na área da linguística, a grande maioria dos estudos da linguagem têm enfatizado análises no nível da palavra, sendo que poucos descrevem como se dá o processamento do discurso, como no caso da produção de narrativas orais, abordando o estrato idoso de nossa população, tanto saudável, quanto com comprometimento cognitivo (BRANDÃO; PARENTE, 2010). Sabe-se que a narrativa é um tipo textual (MARCUSCHI, 2008) bastante utilizado na comunicação humana, em diferentes faixas etárias (JERÔNIMO; HÜBNER, 2014).

A escolaridade também é um fator que tem sido levado em consideração em estudos sobre linguagem e envelhecimento (ARDILA et. al., 2000; PARENTE, FONSECA, SCHERER, 2008; PARENTE et. al., 2009; PASSOS et al., 2015). Dentre eles, Ardila et. al. (2000) destacam uma vantagem de pessoas com maior escolaridade sobre pessoas com menor escolaridade em testes neuropsicológicos. Além disso, como postulam Passos et. al. (2015), a baixa escolaridade, em alguns casos, dificulta o diagnóstico de demências. Logo, estima-se que, observando as produções de narrativas em idosos com DA em comparação com idosos saudáveis, será possível analisar fatores indicativos da demência nesta produção. Além disso, a partir da comparação entre idosos saudáveis, tendo como variável a escolaridade, será possível identificar qual é o papel que a educação pode desempenhar na diferenciação entre os grupos e se ela pode influenciar no desempenho de idosos em tarefas de avaliação neuropsicológica e/ou linguística.

Uma ferramenta de análise para este tipo de texto que tem se mostrado eficaz em estudos ao redor do mundo é a da análise automática. A análise automática consiste em apontar métricas linguísticas como complexidade sintática, densidade semântica, entre outras, como marcadores de demências no discurso. Pensando nisto, softwares têm sido desenvolvidos para auxiliar na avaliação de textos. No contexto brasileiro, exemplo disso é o Coh-Matrix Dementia (ALUÍSIO; CUNHA; SCARTON, 2016). O software conta com mais de 70 métricas linguísticas para avaliar a produção de narrativas orais (transcritas) de idosos

com declínio cognitivo, como no Comprometimento Cognitivo Leve (CCL)² e na DA, visando a elencar padrões que podem auxiliar no diagnóstico precoce das doenças (CUNHA, 2015).

Pensando nestas características, destacamos o estudo de Jerônimo (2016), que aborda aspectos da produção oral de narrativas por pessoas acometidas pela DA ou pelo CCL, em comparação com idosos saudáveis. Este estudo, porém, não aborda a análise automática das produções orais de narrativas, ferramenta que tem demonstrado relevância na distinção entre grupos clínicos e saudáveis. Nesta área, destacamos os estudos de Kavé e Dassa (2018), Toledo et. al. (2017), Orimaye et. al. (2017), Marcotte et. al. (2017) e Fraser et. al. (2012), que aplicaram as técnicas da análise automática aos discursos produzidos por pessoas acometidas por diferentes tipos de demências de maneira a identificar marcadores discursivos de tais condições clínicas.

Analisando as lacunas deixadas por pesquisas realizadas anteriormente, e pensando na replicação de estudos semelhantes, esta dissertação tem como objetivo principal apresentar uma comparação entre a produção oral de narrativas de pessoas acometidas pela Doença de Alzheimer e saudáveis, bem como a comparação do desempenho entre grupos idosos saudáveis com escolaridade diferente. Entende-se que a partir dessa análise comparativa possam-se apontar fatores linguísticos potenciais marcadores de demências no discurso do idoso, assim como analisar se a escolaridade impacta no desempenho entre grupos saudáveis. Para tanto, serão analisados ambos os níveis micro e macroestrutural das narrativas, conforme proposto por Kintsch e Van Dijk (1978) e Van Dijk e Kintsch (1983). Marcadores linguísticos como esses podem contribuir tanto para os estudos na área da neurolinguística, quanto para áreas da saúde que se beneficiariam do uso dessas ferramentas de análise computacional para fins de diagnóstico e tratamento das doenças.

Acredita-se que idosos saudáveis de alta escolaridade obtenham desempenho superior a idosos saudáveis de baixa escolaridade na produção oral de narrativa nos níveis de micro e macroestrutura, ao passo que, na comparação entre idosos saudáveis e acometidos pela DA, ambos de baixa escolaridade, acredita-se que ambos os grupos tenham melhor desempenho no nível macroestrutural em detrimento do microestrutural. Espera-se também que a análise

² O Comprometimento Cognitivo Leve (CCL) refere-se a alterações cognitivas leves, principalmente associadas a falhas de memória episódica. Ainda que, para alguns pesquisadores, o CCL seja um possível indicador do desenvolvimento da DA, não é correto afirmar que todas as pessoas acometidas por CCL necessariamente progredirão para um quadro de DA – isto se evidencia mais fortemente em pessoas com CCL com prejuízo na memória. Por esta razão, o CCL não é considerado uma demência (STELLA, 2012).

automática evidencie marcadores da DA no texto, isto é, acredita-se que, a partir da análise das métricas linguísticas, será possível diferenciar o texto produzido pelos idosos acometidos pela DA do texto produzido pelos idosos saudáveis.

Mais especificamente, visando dar conta dos objetivos aqui descritos, esta dissertação apresenta dois estudos. O primeiro, com foco na escolaridade, visa a comparar dois grupos (um de idosos saudáveis de alta escolaridade e outro, de idosos saudáveis de baixa escolaridade) a fim de investigar as diferenças que brotam do desempenho dos grupos em tarefa de produção oral de narrativa no nível de macroestrutura (a partir do uso de proposições), microestrutura (com base na análise de métricas extraídas pela análise automática de textos) e construtos de memória de trabalho (MT) e memória episódica (ME) (de acordo com escores oriundos da administração de testes neuropsicológicos). No segundo estudo também será realizada uma comparação entre dois grupos, sendo estes divididos entre idosos saudáveis de baixa escolaridade e idosos acometidos pela DA de baixa escolaridade. Neste estudo, a variável de observação é a demência. Para tanto, serão investigados os mesmos construtos descritos para o primeiro estudo.

De maneira a melhor organizar e apresentar os dados e discussões propostas, esta dissertação está dividida em seis capítulos. O primeiro é a introdução, na qual apresentam-se a justificativa da pesquisa, os objetivos e hipóteses que a norteiam. O capítulo seguinte versa sobre as bases teóricas fundamentais para o desenvolvimento e a discussão dos estudos propostos. O capítulo está dividido em duas sessões principais, que visam dar conta de elencar pressupostos teóricos que embasem as noções de narrativa, os níveis micro e macroestrutural de produção de narrativas orais, o envelhecimento saudável e demencial, as influências da escolaridade no discurso, assim como uma revisão de trabalhos desenvolvidos na área da análise automática de textos. O capítulo 3 detalha o experimento realizado nesta pesquisa, apresentando os objetivos, hipóteses e método adotados, assim como a divisão entre os dois estudos. A apresentação dos resultados será feita no capítulo 4 e a discussão dos resultados, por hipóteses, no capítulo 5. O capítulo 6 encerra o presente trabalho tratando das considerações finais.

2 FUNDAMENTAÇÃO TEÓRICA

Visando a elencar fundamentos necessários às discussões e análises apresentadas nesta dissertação, este capítulo tem por objetivo apresentar as referências teóricas consultadas para a elaboração deste trabalho. Sendo assim, o capítulo divide-se em duas grandes seções: a primeira, com foco na narrativa e na produção textual, versa sobre os principais aspectos teóricos que dão conta da caracterização e da produção deste tipo textual, assim como características específicas do processo de produção de narrativas no envelhecimento saudável e a influência da escolaridade; a segunda seção, por sua vez, trata dos reassentamentos biopsicossociais do idoso acometido pela Doença de Alzheimer, abordando aspectos da linguagem (como a produção de narrativas) e construtos cognitivos (como memória episódica e memória de trabalho). Ainda nesta seção, apresenta-se uma revisão da literatura da área da análise automática de produção textual, ferramenta que tem demonstrado potencial para apontar marcadores de demências no discurso.

2.1 PRODUÇÃO DE NARRATIVAS NO ENVELHECIMENTO SAUDÁVEL E SUA RELAÇÃO COM A ESCOLARIDADE

O texto narrativo, seja ele oral ou escrito, faz parte da vida cotidiana, desde a vida da criança que é posta para dormir ao som dos apreciados contos de fadas, até a do idoso que compartilha suas experiências de vida por meio de narrativas. Entender os processos subjacentes à produção narrativa, neste contexto, se faz necessário principalmente no envelhecimento, quando quadros clínicos de demências parecem afetar esta habilidade comunicativa. Mas antes de especificar nossos estudos à análise das narrativas, julgamos necessário aqui abordar a questão um tanto complexa da noção de texto. É comum na literatura uma associação entre os termos texto e discurso, seja ela convergente ou diferenciadora. Em uma perspectiva sociointerativa, conforme Marcuschi (2008), compreendemos que texto e discurso emergem da produção linguística como um todo, ainda que cada um ocupe a sua posição. O texto é uma “entidade comunicativa que forma uma unidade de sentido” (MARCUSCHI, 2008, p. 72). O discurso, por outro lado, estabelece-se com a noção de prática, situada em contexto, referida pelo autor como “objeto do dizer” (MARCUSCHI, 2008, p. 84). Ainda assim, o autor frisa que uma distinção muito delimitada

entre os termos não é necessária, visto que ambos se completam na ação enunciativa. Sendo assim, utilizaremos os termos *texto* e *discurso* sem distinção, tomando por base o que sugere Marcuschi (2008), segundo o qual tanto texto quanto discurso se inserem num contexto sociocultural e histórico.

Ao falar de materialização do texto é necessário que nos apropriemos, também, do conceito de gênero textual. O gênero, para Marcuschi (2008, p.84), é “prática social e prática textual-discursiva”. Isto porque, para o autor, o gênero é a forma através da qual o texto se instala como um discurso. Em um exemplo simples, um discurso poderia ser o anúncio da venda de uma casa, enquanto que o texto se constituiria, vamos dizer, da palavra VENDE-SE, junto com as informações de contato do vendedor e/ou imobiliária, e o gênero seria o veículo através do qual esta venda seria divulgada, como uma placa afixada às grades do imóvel.

Nessa esteira, emerge também a questão do tipo textual, conforme apresentada por Marcuschi (2008). Para o autor, os tipos textuais se diferem dos gêneros textuais de várias maneiras. Por exemplo, enquanto os gêneros textuais são ilimitados e se recriam e reinventam socialmente com o passar do tempo, os tipos são mais estipulados e sem tendência à alteração. Além disso, os gêneros estão ligados à materialidade do texto, enquanto o tipo se relaciona com o modo pelo qual o texto vai se materializar. Nesta perspectiva, temos como exemplos de tipos textuais a narração, a argumentação, a exposição, a descrição, entre outros poucos (MARCUSCHI, 2008). Logo, o texto narrativo pode se manifestar através de diferentes gêneros textuais, em diferentes situações discursivas.

Para Labov e Valetzky (1967), a narrativa tem como característica principal a retomada de fatos passados, analisando este tipo textual por uma perspectiva mais sociolinguística, isto é, tendo como objeto de análise narrativas produzidas em contextos naturais, em detrimento de narrativas literárias, por exemplo. Van Dijk (1983) também enxerga a narrativa como um relato de fatos passados, salientando dois aspectos linguísticos principais do texto narrativo: este tipo textual se caracteriza pela descrição de ações de pessoas (aspecto semântico), e a necessidade de contá-las se dá pela sua relevância (aspecto pragmático). Neste sentido, Van Dijk (1983) coloca a narrativa em um nível de análise a que chama de superestrutura narrativa. A superestrutura de um texto é um nível de análise global, mais abrangente, a que o autor relaciona à noção de tipo, que se assemelha ao discutido previamente, a partir das ideias de Marcuschi (2008).

Labov e Walesky (1967) propõem um modelo de estrutura narrativa que se divide em cinco estágios principais: são eles *a orientação, a complicação, a avaliação, a resolução, e a conclusão*. Destes, o que parece ser mais relevante ao sentido da narrativa é o estágio da avaliação, em que acontecem os julgamentos das ações por parte dos personagens. Para os autores, é a partir da avaliação que se altera o estágio inicial da narrativa, processo que encaminha o texto para a conclusão. Van Dijk (1983) simplifica o modelo, propondo três estágios fundamentais da superestrutura narrativa. São eles *o cenário, a complicação e a resolução*. Para o autor, a avaliação não é parte integrante do esquema pois ela não faz parte da trama em si, mas surge do julgamento dos fatos realizado pelo narrador. Neste sentido, Van Dijk (1983) difere-se de Labov e Valesky (1967), dando ênfase à complicação, estágio em que acontecem as alterações do estágio inicial da história.

Outro modelo que visa a explicar a estrutura da narrativa é o proposto por Adam (2008). O autor apresenta um esquema de estrutura narrativa que se divide em cinco macroproposições (Pn):

1. Situação inicial (Orientação) (Pn1): estágio que se dá antes do início do processo, em que se situam os personagens, o local, o tempo, etc;
2. Nó (Desencadeador) (Pn2): estágio em que se inicia o processo narrativo, quando se apresentam os fatos;
3. Re-ação ou Avaliação (Pn3): estágio em que se dá o momento de reação ou avaliação dos fatos por parte dos personagens, isto é, o desenvolvimento do processo;
4. Desenlace (Resolução) (Pn4): estágio em que se modifica o estado inicial, se encaminhando para o fim do processo;
5. Situação final (Pn5): conclusão do processo.

Para Adam (2008), o mais importante no texto narrativo não é o enfoque em um estágio especificamente, mas sim a transição do estágio inicial (Pn1) para a conclusão (Pn5). Assim, destacam-se no modelo do autor as macroproposições intermediárias (Pn2, Pn3, e Pn4) que, de acordo com ele, são cruciais para a identificação de um núcleo da narrativa. Adam (2008, p.228) explica ainda que “a aplicação do esquema de enredo é um processo interpretativo e de construção de sentido”, logo, a avaliação de componentes linguísticos é importante para a definição dos estágios da narrativa. Neste trabalho, adotamos os pressupostos teóricos de produção narrativa conforme discutidos por Adam (2008). A divisão

de macroproposições descrita pelo autor foi utilizada na construção da tarefa linguística adotada nos estudos desta dissertação, sendo apresentada mais adiante, na seção que trata do Método.

Como já mencionado, a comunicação verbal, mais especificamente o ato de contar histórias, é fator corriqueiro na vida em sociedade. Novamente, retomamos a importância do texto de tipo narrativo para o desenvolvimento de nossas habilidades linguísticas, sendo este o primeiro tipo textual que adquirimos, ainda na infância, e o último que perdemos, no envelhecimento (JERÔNIMO; HÜBNER, 2014). Pensando nisso, estudos vêm sendo desenvolvidos ao longo dos últimos anos nas áreas da psicologia cognitiva e da neuropsicolinguística para dar conta de compreender e/ou explicar os processos cognitivos subjacentes a esta prática. Desta forma, surgiram na literatura modelos de processamento (principalmente) e de produção textual, como o proposto por Kintsch e Van Dijk (1978) e Van Dijk e Kintsch (1983), do qual trataremos com mais detalhes na subseção a seguir.

2.1.1 Modelo situacional de processamento e produção textual

Até aqui discutimos sobre a narrativa, elencando definições e características fundamentais ao estudo da estrutura deste tipo textual. Conhecer o processo de estruturação do texto narrativo se faz pertinente para a análise do produto final, isto é, do texto em si. Assim, a literatura da área permite identificar dois níveis principais de análise da narrativa: a compreensão e a produção. Os estudos nas áreas da psicologia cognitiva e da neuropsicolinguística têm se dedicado, majoritariamente, ao processo de compreensão do texto narrativo. Neste trabalho, todavia, temos como objeto de análise o processo de produção destes textos e quais são as diferenças na comparação entre textos produzidos por idosos saudáveis e idosos acometidos pela DA. Para fundamentar as discussões, nesta subseção apresentaremos o modelo situacional de processamento e produção textual proposto por Kintsch e Van Dijk (1978) e Van Dijk e Kintsch (1983).

O estudo de Kintsch e Van Dijk (1978) foi elaborado, primeiramente, focando na análise do processamento do texto em dois níveis principais, a micro e a macroestrutura, a serem descritos com mais detalhes em seguida. Na atualização do modelo, Van Dijk e Kintsch (1983) reforçaram o uso de estratégias comunicativas para o processamento do texto e

apresentaram mais um nível a ser incluído neste esquema: a elaboração de um modelo situacional. Conforme Van Dijk (2010), um modelo situacional de compreensão textual se dá quando uma representação cognitiva do texto é formada e armazenada na memória episódica (ME). Esta representação mental se alia a uma representação semântica do texto, construída a partir de unidades de sentido denominadas proposições. A aliança entre ambas permite elaborar uma interpretação situada do texto, levando em consideração aspectos pragmáticos da linguagem, como o contexto em que o texto se insere. Por conta disto, o modelo é sensível à incorporação de conhecimentos prévios e à ação de processamento textual, tornando-o um processo único e individual. Na análise a ser feita nesta dissertação, concentrar-nos-emos nos dois primeiros níveis, micro e macroestrutura.

No que diz respeito à produção textual, a atualização do modelo, em 1983, também foi de grande valia. O trabalho de Kintsch e Van Dijk (1978) se baseava, principalmente, na compreensão do texto (oral e escrito), reservando à produção aspectos associados às práticas de reconto ou de elaboração de texto a partir de uma imagem e/ou sequência de imagens. Na atualização do modelo, por outro lado, a atenção dada à produção textual foi estendida à análise do planejamento e da organização textual (VAN DIJK; KINTSCH, 1983). Sendo assim, entende-se que o modelo apresentado pelos autores é sensível aos aspectos de produção textual, não somente de compreensão, justificando a escolha teórica deste trabalho nos pressupostos elaborados por eles. Trataremos, então, da manifestação oral da produção do texto de tipo narrativo, por entender que esta está fortemente relacionada com a prática da comunicação humana, especialmente na velhice, tema central deste trabalho.

Mas antes de passar aos processos mais elaborados da produção textual, como o planejamento e a organização, é necessário que analisemos as unidades elementares de sentido do texto: as proposições. Proposições são unidades semânticas derivadas de estruturas linguísticas que auxiliam na construção mental do sentido de um texto (VAN DIJK; KINTSCH, 1983). A união das proposições vai ser responsável por uma das etapas principais do modelo: a formação de um texto-base (*textbase*). Este texto-base é a representação semântica do texto registrada na ME, a qual só se tornará realmente eficaz se atender às demandas básicas das noções de coerência local e global. (VAN DIJK; KINTSCH, 1983). A coerência local se dá no nível da sentença, isto é, na microestrutura do texto. A microestrutura se caracteriza pelos elementos sintáticos, semânticos e discursivos, que compõem as proposições. Sendo assim, a coerência local é responsável por monitorar as relações entre os elementos que compõem as proposições de maneira que façam sentido e elaborem uma

sentença possível em determinada língua. Por outro lado, a coerência global está relacionada às ligações de sentido realizadas em um nível superior do texto, a macroestrutura. A macroestrutura é composta a partir da relação entre as proposições, elaborando, assim, um significado mais geral do texto (KINTSCH; VAN DIJK, 1978). Em linhas gerais, a microestrutura é responsável pelos elementos que compõem as proposições, ao passo que a macroestrutura é a união das proposições com foco na elaboração do sentido global do texto.

Desta forma, a coerência é fundamental para ambos os processos, tanto de compreensão quanto de produção textual. Marcuschi (2008) aponta para o papel fundamental da coerência na construção textual. “A coerência é, sobretudo, uma relação de sentido que se manifesta entre os enunciados, em geral de maneira global e não localizada” (MARCUSCHI, 2008, p. 121). Kintsch e Van Dijk (1978) e Van Dijk e Kintsch (1983) também consideram a coerência (em ambos os níveis, local e global) como fundamental para o processo de construção do texto.

Além das necessárias relações de coerência, Van Dijk e Kintsch (1983) elencam uma série de estratégias que contribuem para o processo de produção textual. Dentre elas destaca-se, principalmente, a estratégia de macroplanejamento. Para os autores, a primeira etapa da produção textual é a elaboração de um plano que se adapte à situação de comunicação desejada, “e este plano deve ser a representação global de um macro ato de fala” (VAN DIJK; KINTSCH, 1983, p. 265). De acordo com Brandão e Parente (2010), é no macroplanejamento que se cria a macroestrutura. Nesta etapa inicial ocorre “a elaboração de objetivos comunicativos para criar objetivos subordinados, recuperando informação apropriada para a produção do discurso” (BRANDÃO; PARENTE, 2010, p. 345). Este planejamento oferece determinado controle sobre a produção, ao passo que atua de maneira hierárquica (*top-down*) nas decisões da produção, isto é, é a partir deste macropiano situacional que o texto será produzido. Esta característica faz com que o macropiano seja visto como o *input* da produção do discurso (VAN DIJK; KINTSCH, 1983).

Van Dijk e Kintsch (1983) postulam que não é possível considerar o macroplanejamento de atos de fala como uma informação pragmática somente, visto que a decisão por um ato de fala em detrimento de outro não surge de maneira independente, mas sim, baseia-se em aspectos semânticos da situação comunicativa em que o falante está inserido. Sendo assim, os autores definem que atuam na produção do discurso um plano pragmático e um plano semântico. Para tanto, o fator sociocultural também desempenha um papel importante no processo de produção textual: é a partir da união dos fatores

socioculturais, comunicativos e pragmáticos que será possível “selecionar informações relevantes, interessantes, novas e importantes” (VAN DIJK; KINTSCH, 1983, p. 288).

Todas essas escolhas de primeira ordem levam à seleção das palavras ou expressões que irão compor as proposições, no nível da microestrutura. Nesta etapa entram os planejamentos fonológico, morfológico, lexical, sintático e semântico, que auxiliam no processo de especificação da ideia elaborada no macro plano. Assim, “informações serão adicionadas, particularizadas e analisadas” (VAN DIJK; KINTSCH, 1983, p. 274). Para a elaboração de proposições, os autores destacam o importante papel da memória episódica no resgate de informações referentes às pessoas, lugares e informações. Em geral, o tipo de texto produzido interfere no tipo de informação mais relevante à construção das proposições, ainda que, em todos os casos, seja necessário que uma determinada ordem de apresentação de informações seja respeitada. No caso de um texto produzido sobre uma pessoa, por exemplo, é importante que se situe, primeiramente, quem é a pessoa, depois a insira em um mundo, isto é, um local possível, e então se especifique as ações relevantes ao que se quer dizer sobre ela. (VAN DIJK; KINTSCH, 1983).

Nesta produção surge novamente a questão da coerência local, que será responsável pela construção de sentido realizada a partir das relações entre as palavras que compõem as proposições. Sendo assim, os autores afirmam que as sentenças vão sendo produzidas à medida que a representação semântica do texto vai sendo construída. Isto se reflete no que Van Dijk (1997) chama de estrutura de superfície, uma estrutura superficial do texto, composta de uma sequência de sentenças (que podemos associar à noção de microestrutura). Esta estrutura de superfície depende dos planos pragmático e semântico vistos anteriormente, vindo a se manifestar a partir da elaboração deles. Algumas estratégias para a elaboração da estrutura de superfície são a entonação, a ordenação das palavras, os sintagmas nominais definidos e/ou indefinidos, os pronomes, os predicados, os tempos verbais, as partículas de ênfase (como *ah*, *hmm*), os conectivos, os advérbios e a estruturação de sentenças, entre outros (VAN DIJK; KINTSCH, 1983).

Em resumo, a produção textual se apropria de ambos os níveis da micro- e macroestrutura, conforme propostos por Kintsch e Van Dijk (1978) e Van Dijk e Kintsch (1983). As principais estratégias de produção são (i) a elaboração de um macroplanejamento (no nível de macroestrutura), responsável pela união de aspectos pragmáticos e semânticos no que se denomina o *input* do discurso; e (ii) a construção de uma estrutura de superfície (no nível de microestrutura), em que os fatores fonológico, morfológico, lexical, sintático e

semântico serão agrupados na elaboração de proposições que serão determinantes para especificar as ideias elaboradas na etapa do macroplanejamento.

Por entender que o texto narrativo é bastante presente na comunicação em sociedade, apresentaremos, na subseção a seguir, aspectos mais específicos da produção deste tipo de texto especialmente no envelhecimento saudável, buscando compreender quais são as características da produção de texto narrativo nesta fase da vida. Além disso, de maneira a compreender melhor quais fatores podem estar associados à qualidade da produção de texto narrativo, damos ênfase à análise da influência do grau de escolaridade neste processo.

2.1.2 A produção de narrativas no envelhecimento saudável e a influência da escolaridade

Após abordar os principais pressupostos teóricos subjacentes à estrutura do texto narrativo, assim como aos aspectos cognitivos de produção textual, nesta subseção direcionamos o foco da discussão à análise de estudos que trazem evidências sobre a produção de narrativas no envelhecimento saudável e a influência da escolaridade nessas produções.

É de conhecimento geral que o envelhecimento acarreta uma série de mudanças biopsicossociais que podem impactar diretamente na qualidade de vida. Problemas de memória em qualquer faixa etária têm estreita relação com falhas na linguagem (IZQUIERDO, 2011), o que acaba trazendo dificuldades de comunicação e afetando as relações interpessoais, visto que a linguagem se caracteriza pelo aspecto social, de interação (MARCUSCHI, 2008). Logo, o idoso precisa buscar a superação de possíveis dificuldades para atribuir sentido às atividades que desempenha (RIVERO et. al., 2013).

É importante ressaltar, porém, que o envelhecimento é um processo biológico, logo, envelhecer não significa adoecer. Inclusive, os hábitos de vida pré-envelhecimento são determinantes para um processo de envelhecimento ainda mais saudável. Pensando nisso, diferentes áreas do conhecimento vêm elencando como prioridade estudos que auxiliem na quebra dos estereótipos associados à velhice como um estágio “não-produtivo” da vida, ou ainda do “velho” como sinônimo de doença, conforme apontam Rivero, et al, (2013). Alguns destes estudos abordam os termos reserva cognitiva e reserva cerebral como construtos responsáveis por apresentar benefícios cognitivos ao processo de envelhecimento. Cosenza e

Malloy-Diniz (2013, p. 339) diferenciam ambos os termos, definindo reserva cerebral como “substrato neural estrutural”, isto é, o tamanho de estruturas cerebrais e a quantidade de neurônios; e reserva cognitiva como a habilidade do cérebro de compensar falhas cognitivas através da “implementação de processos cognitivos alternativos”.

A literatura apresenta modelos de reorganização cerebral decorrente do envelhecimento. Alguns desses modelos inserem em suas discussões o papel das reservas cognitivas nessa reorganização. O modelo *Hemispheric Asymmetry Reduction in Older Adults* (HAROLD) (CABEZA et. al. (2002) indica haver um padrão de ativação hemisférica bilateral para dar conta de processos cognitivos no envelhecimento, preservando o nível de performance cognitiva de idosos. Já a teoria da Reserva Cognitiva (STERN, et. al., 2003) postula que o cérebro idoso se utiliza de reservas cognitivas associadas à escolaridade, estilo de vida, dentre outros fatores, para compensar os danos cognitivos, em um modelo que pode ver a reserva como sendo passiva (relacionada à quantidade de sinapses) ou ativa (relativa ao recrutamento de diferentes redes neurais).

Outro modelo, o *Revised Model of the Scaffolding Theory of Aging and Cognition* (STAC-r) (REUTER-LORENZ; PARK, 2014), uma revisão do modelo de *scaffolding* (andamento) (PARK; REUTER-LORENZ 2009), apresenta uma abordagem holística, ou seja, leva em consideração fatores biológicos, psicológicos, cognitivos e sociais para fortalecimento dos processos compensatórios que ocorrem ao longo da vida e no período do envelhecimento. Destacamos os hábitos de lazer, o bilinguismo e a escolaridade como fatores relevantes ao modelo dos autores.

Dentre as vantagens da reserva cognitiva, Stern (2012) enfatiza mais uma vez o papel do grau de escolaridade para os índices de reserva cognitiva. Ainda que a reserva cognitiva não possa impedir o desenvolvimento de doenças como demências, a partir dos dados obtidos em um estudo com 593 participantes saudáveis, Stern (2012) constatou que pessoas mais escolarizadas têm menos risco de desenvolver sintomas associados a demências na comparação com pessoas menos escolarizadas.

A análise dos estudos permite concluir que os hábitos que adotamos ao longo da vida têm relação direta com o processo de envelhecimento mais saudável. Fatores socioculturais parecem apresentar relação direta com uma melhor qualidade de vida, sendo possível identificar a importância da escolaridade, por exemplo, como tendo papel de destaque no desenvolvimento da cognição humana. Logo, compreender sua influência no desempenho de

idosos em tarefas de linguagem pode ser fator determinante na caracterização das diferenças individuais obtidas em resultados desses tipos de tarefas.

Aspectos gerais do discurso do idoso são apresentados por Salles e Brandão (2013), que fazem um apanhado de características da produção textual no envelhecimento saudável. De acordo com as autoras, em geral, idosos saudáveis apresentam declínio na produção de frases complexas (por exemplo, orações subordinadas), o que poderia estar associado ao declínio na memória operacional (ou de trabalho), que fica sobrecarregada ao dar conta de sentenças complexas. De acordo com as autoras, também é comum no discurso do idoso um acréscimo de informações pessoais e uma diminuição do grau de coerência. Estes fatores poderiam estar associados, de acordo com elas, a duas hipóteses distintas: a hipótese pragmática e a do déficit de inibição. De acordo com a hipótese pragmática, o aumento de informações pessoais e irrelevantes ao assunto poderia estar associado ao contexto social ou à identidade do falante, sendo a intenção comunicativa o aspecto mais relevante. Por outro lado, a hipótese do déficit de inibição está associada a uma falha na função executiva que inibe as interferências durante o processamento na memória de trabalho no momento da produção do discurso (SALLES; BRANDÃO, 2013).

Uma forma de compreender o papel da escolaridade no discurso (seja ele oral ou escrito) do idoso se dá através da avaliação da linguagem a partir de tarefas linguísticas em conjunto com a análise de fatores socioculturais e/ou neuropsicológicos. Na avaliação da linguagem no envelhecimento, diferentes abordagens metodológicas vêm sendo adotadas na análise linguística, movidas pelo interesse em compreender o desempenho de idosos em todos os níveis - da palavra ao texto. Conforme visto anteriormente, a grande maioria dos estudos realizados no âmbito do texto acaba por abordar a questão da compreensão, em detrimento da produção textual. Assim, em seguida apresentaremos alguns estudos que abordaram em especial a produção textual, por estar mais relacionada ao tópico desta pesquisa.

Destacamos o estudo de Jerônimo (2016), que aborda aspectos da produção oral de narrativas por pessoas acometidas pela DA ou pelo CCL e a sua relação com construtos cognitivos. Além dos grupos clínicos, a autora também avaliou os mesmos construtos em idosos saudáveis, levando em consideração a variável escolaridade. A tarefa linguística utilizada pela autora consiste em elaborar uma produção oral de narrativa a partir de uma sequência de figuras. De maneira geral, a autora analisa a construção da superestrutura narrativa, a coerência (local e global) e o uso de estratégias comunicativas. Além disso, também analisa resultados advindos da administração de testes neuropsicológicos visando a

identificar marcadores de comprometimento cognitivo em construtos como a memória episódica e memória de trabalho, os quais vamos abordar mais adiante. A autora encontrou diferenças significativas na comparação entre grupo de idosos saudáveis de alta e baixa escolaridade. De acordo com ela, idosos mais escolarizados tiveram melhor desempenho na construção de todas as etapas da superestrutura narrativa, isto é, a união das macroproposições, conforme o modelo de Adam (2008). Além disso, também foram observadas diferenças significativas no nível da coerência local e global, com vantagem para o grupo mais escolarizado em ambas.

O estudo de Vargas (2015) também abordou a produção de narrativas e sua relação com a escolaridade, assim como na DA. Novamente, o estudo apresenta uma comparação entre idosos saudáveis de alta e baixa escolaridade em tarefas de produção textual. Neste estudo, a autora avaliou duas tarefas linguísticas de produção de narrativa: uma que solicitava ao participante contar uma notícia, e outra que contasse uma história a partir da observância de uma sequência de figuras. A análise se deu de forma quantitativa e qualitativa, a partir dos pressupostos teóricos da referenciação. Seus resultados principais indicaram que o grupo saudável com menor grau de instrução demonstrou pior desempenho em ambas as tarefas na comparação com o grupo mais escolarizado. Em geral, as narrativas produzidas pelo grupo de alta escolaridade eram mais coerentes e sem problemas de referenciação, ao passo que as narrativas produzidas por idosos de baixa escolaridade eram menos elaboradas e continham mais expressões de sentido vago e/ou ambíguo.

Conforme observado na descrição dos estudos apresentados até aqui, a avaliação da linguagem costuma aparecer associada à avaliação neuropsicológica. Dessa forma, Parente et. al. (2009), em sua revisão de literatura, concluem que a variável escolaridade desempenha papel significativo também na avaliação de construtos neuropsicológicos no envelhecimento. De acordo com as autoras, ainda que não haja um consenso entre os estudos disponíveis na literatura, parece haver uma tendência de pessoas mais escolarizadas obterem melhores escores em avaliações desta natureza. Ainda de acordo com elas, a baixa escolaridade pode atuar de maneira negativa no diagnóstico de demências, por exemplo, visto que pessoas saudáveis de baixa escolaridade podem apresentar desempenho semelhante ao de pessoas acometidas por demências. Esta advertência também é apontada pela *American Psychiatric Association* (2014), que aponta o padrão sociocultural e educacional mais baixo como um desafio para a avaliação cognitiva. Sendo assim, parece pertinente investigar a influência da

escolaridade não somente em tarefas linguísticas, mas também em avaliações neuropsicológicas no envelhecimento.

Retomando alguns pontos importantes desta subseção, mais uma vez mencionamos o processo de envelhecimento como natural e biológico, variando muito de indivíduo para indivíduo, principalmente a julgar pelos diferentes estilos de vida adotados e pelo estado de saúde em geral. No que diz respeito às características da linguagem, buscamos elencar exemplos de alterações cognitivas que podem interferir na competência linguística do idoso, mais uma vez, sendo um processo bastante individual. Mas como identificar as alterações de linguagem decorrentes de processos demenciais? É possível apontar marcadores linguísticos da DA no discurso do idoso? Esperamos discutir essas respostas na seção seguinte, que trata mais especificamente dos processos linguísticos e cognitivos afetados na Doença de Alzheimer.

2.2 PRODUÇÃO E ANÁLISE DE NARRATIVA NA DOENÇA DE ALZHEIMER E SUA RELAÇÃO COM CONSTRUTOS COGNITIVOS

Esta seção tem como objetivo discutir o processo de produção de narrativas na Doença de Alzheimer. Para tanto, serão apresentados aspectos característicos da linguagem de pessoas acometidas pela DA. Na primeira seção, discutiremos as mudanças cognitivas, mais especificamente relacionadas às memórias do tipo episódica (ME) e de trabalho (MT), no quadro demencial. Em seguida, é apresentada uma revisão da literatura na área da análise automática de texto, ferramenta utilizada para a análise de construtos linguísticos, a qual tem potencial para diagnóstico de demências por meio da comparação com dados linguísticos de pessoas idosas saudáveis, buscando a identificação de marcadores linguísticos da DA no discurso.

Antes de tudo, porém, faz-se necessário compreender o que são demências. As demências, como visto anteriormente, se caracterizam por ser um processo degenerativo de perda de neurônios essenciais para funções cognitivas importantes, como a criação de memórias, por exemplo (IZQUIERDO, 2011). Sendo assim, um quadro demencial é um quadro clínico, não sendo possível associar os sintomas de demências ao processo de envelhecimento natural. Nem toda a pessoa que envelhece vai, necessariamente, desenvolver algum tipo de demência.

O número de pessoas com diagnóstico de demência no Brasil e no mundo vem crescendo nos últimos anos. Dentre os tipos de demência, a mais comum é a demência do tipo Alzheimer, ou Doença de Alzheimer (DA). Dentre os critérios para definição do diagnóstico de provável doença de Alzheimer apresentados pela *American Psychiatric Association* (2014) estão: a. perda de memória e dificuldade de aprendizagem associadas a pelo menos mais domínio cognitivo; b. declínio progressivo na cognição; c. ausência de outras doenças neurodegenerativas possíveis causadoras dos declínios descritos anteriormente. Evidências genéticas também são fortes indicativos para diagnóstico da DA. Dubois et. al. (2014) são ainda mais específicos, associando as primeiras falhas de memória a déficits na memória episódica (construto do qual trataremos na subseção a seguir). Os autores ainda apontam para a importância de biomarcadores (como a presença da proteína beta-amiloide no cérebro) que podem auxiliar no diagnóstico precoce da doença.

No que diz respeito a falhas na linguagem, os primeiros sintomas expressados através da fala em pessoas acometidas pela DA estão associados a anomias (dificuldade ou incapacidade de encontrar palavras), parafasias semânticas (troca de uma palavra por outra, do mesmo grupo semântico), substituição de palavras, uso de hiperônimos e empobrecimento do vocabulário em geral (SOARES; BRANDÃO; LACERDA, 2012). Brandão e Parente (2010) também fizeram importantes contribuições para a compreensão de aspectos da produção de discurso oral por pessoas acometidas pela DA. De uma maneira geral, as autoras destacam a questão da coerência, ou melhor, da *incoerência* global. O termo escolhido por elas visa a explicitar a característica de um discurso tido como de baixo valor significativo, em que se nota pouca relação entre as ideias complementares e as ideias principais (o macroplano, conforme visto anteriormente). Por outro lado, tanto a fluência quanto a sintaxe parecem estar preservadas nos estágios iniciais da doença. Entretanto, de acordo com as autoras, as construções sintáticas são, em geral, mais simplificadas, demonstrando que a complexidade sintática pode estar afetada. No que diz respeito à fluência, as pausas, hesitações e revisões parecem ser comuns no discurso oral destas pessoas (BRANDÃO; PARENTE, 2010).

Soares, Brandão e Lacerda (2012) também avaliam que, na DA, a produção oral de narrativa pode estar comprometida. De acordo com elas, para tarefas de produção oral de narrativa em que o participante com DA deve criar uma história a partir de uma única figura, há uma tendência à descrição da imagem em detrimento da construção de uma sequência narrativa. Já quando é exposto a uma sequência de figuras, percebe a continuidade da história

mas tem dificuldade em narrá-la, pois sofre com uma demanda alta de processos cognitivos oriundos de funções cognitivas que são comprometidas na doença.

Kavé e Dassa (2018) também apresentam uma série de aspectos característicos da linguagem de pessoas acometidas pela DA em sua revisão de literatura. Em geral, em tarefas de produção discursiva a partir de figuras, pessoas com DA costumam utilizar mais informações irrelevantes, apresentar dificuldade na construção da ideia geral (macroestrutura) do texto, incluir pronomes em demasia (inclusive sem referentes) e ter dificuldade para encontrar palavras. Em linhas gerais, declínio em aspectos semânticos e gramaticais parecem estar evidentes no discurso de pessoas com este tipo de demência (KAVÉ E DASSA, 2018).

Novamente destacamos os trabalhos de Jerônimo (2016) e Vargas (2015), desta vez focando em seus achados relacionados às populações clínicas. Jerônimo (2016) aborda aspectos da produção oral de narrativas por pessoas acometidas pela DA ou pelo CCL utilizando uma tarefa linguística que consiste na produção oral de uma narrativa a partir de uma sequência de figuras. Os dados linguísticos apontam para comprometimentos na produção de narrativas em idosos acometidos pela DA e pelo CCL, especialmente nas áreas de superestrutura narrativa e coerência, não demonstrando diferenças significativas entre DA, CCL e controles no que diz respeito às estratégias comunicativas, apontando para preservação de competência pragmática nos grupos com comprometimento.

O estudo de Vargas (2015) avaliou duas tarefas linguísticas de produção de narrativa, com foco no processo de referenciação: uma notícia e uma história elaborada a partir de uma sequência de figuras. Seus resultados principais indicam, para o grupo com DA, maiores dificuldades no processo de referenciação na construção da notícia e dificuldade de contar fatos recentes; maior número de pronomes demonstrativos (dêiticos) vagos na tarefa de construção de narrativa a partir da figura, o que denota ambiguidade nos referentes, além de predominância na descrição de cenas em detrimento à estrutura narrativa.

Lima et. al. (2014) avaliaram a produção de narrativa a partir de sequência de figuras em pessoas acometidas pela DA. As autoras conduziram uma análise qualitativa dos textos narrativos, a partir dos quais puderam observar falhas na coerência local e global, assim como falta de informações na construção do texto e repetições de palavras. De acordo com as autoras, este tipo de tarefa, com sequência de figuras, exige dos participantes com DA habilidades que são afetadas pela demência, como a percepção visual. Esse fato aponta para o

cuidado de se observarem o nível de acometimento das capacidades de audição e visão ao se administrarem tarefas a esse grupo clínico.

Estes exemplos parecem corroborar os achados de Salles e Brandão (2013) e Brandão e Lacerda (2012), visto que (i) os dados de Lima et. al. (2014) e Kavé e Dassa (2018) indicaram dificuldades em relação à tarefa com sequência de figuras, possivelmente afetada também por comprometimentos na percepção visual dos participantes com DA; (ii) os de Jerônimo (2016) apontaram para preservação das competências pragmáticas na DA; e (iii) os de Vargas (2015) apresentaram a característica da predominância na descrição de cenas. O estudo de Jerônimo (2016) ainda apontou para declínios em tarefas neuropsicológicas, atentando para a importância da avaliação de construtos cognitivos enquanto possíveis marcadores de demências. Por esta razão, na subseção seguinte abordamos aspectos cognitivos afetados pelos quadros demenciais, mais especificamente as memórias episódica e de trabalho.

2.2.1 Doença de Alzheimer e memória: aspectos das memórias episódica e de trabalho

Uma das principais queixas no envelhecimento é em relação à memória. Para Izquierdo (2011), as memórias fazem parte da construção da nossa identidade, são parte integrante daquilo que nós somos. “O acervo de nossas memórias faz com que cada um de nós seja o que é: um indivíduo, um ser para o qual não existe outro idêntico” (IZQUIERDO, 2011, p. 11). De acordo com Abrisqueta-Gomes (2013), o declínio na velocidade de processamento pode ser o responsável pelas falhas de memória, e não problemas na memória em si. Nesta seção temos por objetivo discutir alguns aspectos da memória e sua relação com a DA. Para tanto, serão apresentados também alguns dados referentes aos construtos no envelhecimento saudável, de maneira a compará-los com as mudanças advindas da demência.

Em geral, as memórias são divididas entre dois grandes grupos: a memória de curta e de longa duração. A memória de curta duração é responsável pela retenção de poucas informações a curto prazo. A memória de longa duração faz exatamente o oposto: armazena grandes quantidades de informação por período indeterminado. Por ser mais complexa, divide-se entre memória implícita e explícita. A memória implícita está relacionada ao desempenho, isto é, à repetição e automatização de processos, não podendo facilmente ser acessada conscientemente; enquanto que a memória explícita é caracterizada pelo acesso

consciente às lembranças armazenadas e divide-se entre memória episódica (ME) e memória semântica (MS) (CAPUANO, 2010). Neste trabalho, daremos atenção maior à ME, por sua estrita relação com a DA, conforme entenderemos mais adiante.

A *memória episódica (ME)* é, grosso modo, responsável pela “capacidade de lembrar eventos específicos” (BADDELEY, 2011a, p. 107). É neste tipo de memória que se registram, por exemplo, o dia de nosso casamento, de nossa formatura, do nascimento dos filhos, do primeiro dia de aula. Para Izquierdo (2011), a ME é responsável por armazenar os momentos que vivemos e/ou a que assistimos. Por esta razão, muitas dessas memórias são autobiográficas, pois existem a partir das experiências que tivemos. Ainda assim, a ME não pode ser vista como um construto único, visto que os processos de armazenamento, codificação e resgate de informações têm papéis distintos no processo de recordação (GROBER; VEROFF, LIPTON, 2018).

No envelhecimento saudável, de acordo com Abrisqueta-Gomes (2013), a ME sofre declínios significativos no que diz respeito aos detalhes, já que, em geral, idosos conseguem lembrar de eventos passados. A autora utiliza como exemplo o fato de que, em geral, idosos têm mais dificuldade em lembrar as roupas que vestiram alguns dias atrás do que adultos mais jovens. Na DA, a ME é, de acordo com Baddeley (2011b), a principal deficiência observada.

Um tipo de avaliação neuropsicológica amplamente utilizado no mapeamento das funções da ME é o teste *Free and Cued Selective Reminding Test (FCSRT)* (GROBER, et. al., 1988). Esta tarefa tem como etapas o aprendizado e a recordação de uma lista de figuras. Para tanto, na etapa do aprendizado, o participante deve identificar e nomear uma figura, relacionando-a com uma determinada categoria semântica, por exemplo, *macaco – mamífero*. As categorias semânticas auxiliam na etapa de recordação de informações, que se divide em recordação livre (sem pista semântica) e recordação guiada (com pista semântica). A BALE (Bateria de Avaliação da Linguagem no Envelhecimento) (HÜBNER, et. al., no prelo) sugere que a etapa de recordação seja repetida com intervalos curtos (de 20 segundos) por duas vezes, seguidas de uma terceira sessão após 20 minutos de intervalo. Além disso, a BALE ainda sugere a recordação imediata na fase de aprendizado das figuras, uma adaptação de versão mais recente do teste (FCSRT+IR) (BUSCHKE, GROBER, 2012). A pontuação se dá pelo total de recordação sem pista, com pista, e total geral, obtido através da soma das etapas anteriores.

Grober, Veroff e Lipton (2018) salientam que os primeiros sinais de declínio da ME podem ser percebidos na etapa de recordação livre. Segundo os pesquisadores, falhas na recordação sem pista podem indicar déficits na capacidade de recuperação de informações, ao contrário de déficits no armazenamento, visto que, em geral, o participante consegue recordar com o auxílio da pista semântica. Isso também é indicativo de que escores mais baixos na soma do total de recordação livre e guiada refletem dificuldade de armazenamento, visto que, mesmo que com auxílio das pistas, o participante não consegue recordar as palavras estudadas. Portanto, o total de recordações atua, geralmente, como indicativo de DA (GROBER, VEROFF E LIPTON, 2018).

A literatura traz estudos que se utilizaram do FCSRT para avaliação e distinção entre grupos de idosos acometidos por mais de um tipo de demência. Em geral, o teste parece diferenciar a DA de outros quadros demenciais, além de apontar com alta probabilidade evolução de quadros pré-demenciais, como o CCL, para quadros de Alzheimer. Por essa e outras razões, o FCSRT é recomendado pelo *International Working Group*, instituição que busca propor biomarcadores para facilitar o diagnóstico da DA, para a avaliação da ME (DUBOIS, et. al., 2014). A seguir discutiremos algumas evidências trazidas por estudos recentes que debatem sobre o assunto.

Teichmann et. al. (2017) avaliaram a precisão do FCRST no diagnóstico de DA. Na comparação com outras demências (como Afasia Primária Progressiva (APP), Doença com Corpos de Lewy, Demência Frontotemporal (DFT), entre outras), a avaliação foi capaz de distinguir a DA de todas as outras demências, sendo que pessoas com DA obtiveram baixos escores em todas as etapas de recordação de palavras. De acordo com os autores, os resultados indicam que a avaliação da ME na DA pode ter papel importante para o diagnóstico da demência.

Lemos et. al. (2014) e Lemos et. al. (2017) também avaliaram a eficácia do teste na distinção entre quadros clínicos. Em um estudo comparativo, Lemos. et. al. (2014) tinham por objetivo observar a eficácia do FCSRT na diferenciação entre grupos de idosos acometidos pela DA ou pela DFT. Para tanto, foi utilizada uma versão em português europeu do teste adaptada por Lemos, et. al. (2012), além de outros testes neuropsicológicos e de funcionalidade para avaliação de construtos como memória de trabalho, atenção, fluência verbal, entre outros. Os principais resultados do FCSRT permitem constatar que, enquanto o grupo com DA teve piores escores em todas as etapas do teste (recordação livre, guiada, e escore total), os participantes com DFT pareceram se beneficiar das pistas semânticas na

recordação das palavras, não apresentando diferença significativa na comparação com grupo controle para esta variável. A acurácia na diferenciação dos grupos com demência chegou próxima de 80%, tornando possível uma identificação de 95% dos participantes com DA a partir de seus resultados na avaliação de ME.

Lemos et. al. (2017) avaliaram 100 participantes diagnosticados com CCL amnésico em um estudo longitudinal, com o objetivo de identificar a performance no FCSRT como indício do desenvolvimento do quadro de DA. Outros testes neuropsicológicos também foram administrados neste protocolo de coleta de dados, que acontecia de seis em seis meses. Os participantes eram acompanhados até apresentarem mudança no quadro demencial, conforme testagem clínica e neuropsicológica realizada. Dos 100 participantes inicialmente recrutados, apenas 88 realizaram as sessões de acompanhamento, sendo incluídos nos resultados finais. Durante este período, os participantes diagnosticados com CCL em risco de desenvolvimento de DA foram separados dos participantes diagnosticados com CCL, porém estáveis. Utilizando a mesma versão do teste descrito no estudo anterior, em português europeu, os pesquisadores conseguiram identificar um número de 29 participantes (33%) que evoluíram para um quadro demencial, sendo, em todos os casos, quadros de DA. Em geral, a média de tempo de conversão de CCL para DA foi de 23,82 meses. Este resultado sugere que, provavelmente, os participantes que evoluíram de quadro já deveriam estar em estágio pré-demencial desde o início da testagem. De todas as variáveis (biológicas e neuropsicológicas) observadas, os escores para o total de recordação pareceram estar mais diretamente correlacionados com a progressão da demência, isto é, pessoas com menores escores tinham mais riscos de converterem para DA (LEMOS, et. al., 2017).

Além disso, os resultados apontados pela administração do FCSRT e testes nele baseados parecem não ser afetados por distintos graus de escolaridade. Hübner, et. al. (2018), em estudo cujo objetivo foi observar o impacto do grau de escolaridade em tarefas neuropsicológicas (como a nomeação, a aprendizagem verbal com pistas, entre outros), apresentaram evidências de que não há diferenças significativas entre os escores de grupos de participantes saudáveis divididos pelo grau de escolaridade (muito baixo, baixo e alto) na tarefa de ME da BALE, a qual, como já apontado, baseou-se na FCSRT. Estes resultados parecem indicar que este teste pode ser útil na diferenciação de grupos de pessoas saudáveis e com demência, independentemente do grau de escolaridade. Jerônimo (2016) também não identificou diferenças significativas na comparação entre idosos saudáveis de alta e baixa escolaridade na tarefa de ME da BALE.

Estes estudos parecem corroborar os pressupostos apresentados por Baddeley (2011b), ao trazerem evidência para o fato de que o desenvolvimento dos sintomas da DA acarreta falhas na ME. Ainda assim, a ME, enquanto responsável pela recordação de fatos específicos, não atua de maneira isolada neste processo. Eysenck (2011) apresenta um exemplo que ilustra esta afirmação. Quando uma pessoa lembra de um encontro que teve com um amigo em uma cafeteria, vai necessitar da ME para realizar esta recordação, mas impressões sobre o sabor do café e o conhecimento das características de uma cafeteria não são associadas à ação da ME. Neste sentido, a *memória semântica (MS)* atua em conjunto com a ME, resgatando informações gerais armazenadas. Izquierdo (2011, p. 30) conceitua a MS de maneira simples, como contendo as memórias de “conhecimentos gerais”. Eysenck (2011) corrobora e é ainda mais específico, associando a MS ao nosso vocabulário, trabalho e atividades de interesse, sobre as quais registramos mais informações na memória do que outras pessoas.

No envelhecimento saudável, a MS parece ser uma das mais preservadas, ainda que seja comum uma queixa de idosos sobre o esquecimento de palavras. Além disso, este tipo de memória se mantém sempre ativo, visto que lida com um tipo de conhecimento tido como “cristalizado” (ABRIZQUETA-GOMEZ, 2013, p. 175). De acordo com Baddeley (2011b), um deterioro na MS na DA parece progredir juntamente com a evolução da demência. Pessoas com DA parecem apresentar dificuldades em tarefas que objetivam mensurar a capacidade da MS, como na nomeação de figuras.

Há ainda na divisão dos construtos de memória a *memória operacional*, ou *memória de trabalho (MT)*. A MT tem uma característica que a difere dos outros tipos de memória: ela não está ligada ao processo de armazenamento de informações. Pelo contrário, ela é uma memória temporária, essencial para a manipulação de fatos necessários em um determinado momento, que pode ser poucos minutos (IZQUIERDO, 2011). Um dos modelos mais conhecidos de MT foi o proposto por Baddeley (1986), que centraliza a ação no executivo central, apoiado pela alça fonológica e o esboço visuoespacial. A alça fonológica é responsável pelo apoio sonoro das palavras, enquanto que o esboço visuoespacial contribui com o processamento do cenário espacial e visual. Tudo isso é gerenciado pelo executivo central, que se utiliza de outros subprocessos (como a atenção) para manipular as informações necessárias. Ainda assim, por ser uma memória de caráter transitório, necessita buscar informações armazenadas na memória de longo prazo, ação necessária na complementação da atividade dos três primeiros componentes do modelo. Por esta razão, Baddeley (2000) incluiu

em seu modelo a noção do *buffer* episódico, isto é, um componente multidimensional responsável pelas diferentes ligações entre a MT e a ME de longo prazo.

Por conta de sua atuação junto ao executivo central, a MT é comumente referenciada na literatura como um componente das chamadas *funções executivas (FEs)*. Ainda que sejam muitas as definições de FEs, de Paula et. al. (2013, p.227) definem-nas como “funções cognitivas complexas, relacionadas ao comportamento dirigido a metas e que possibilitam ao indivíduo a solução de problemas novos, o planejamento em escala temporal, a análise de custo e benefício no processo de tomada de decisões e a autorregulação comportamental”. Para os autores, *funções executivas* é um termo guarda-chuva designado a abarcar processos cognitivos de alta ordem, fundamentais ao comportamento humano em sociedade.

Dentre as opções disponíveis de testes neuropsicológicos para avaliação da MT encontra-se a tarefa de ordenamento ascendente de dígitos (Digit Span) (FONSECA et al., 2009). Nesta tarefa, o participante deve manipular uma sequência numérica na memória, visando repeti-la em uma das duas modalidades do exame: direta ou indiretamente. Na forma direta, o participante deve repetir a sequência na mesma ordem em que escutar do aplicador. Por exemplo, se ouvir a sequência 1 -3 - 4, deve repetir 1 - 3 - 4. Já na versão indireta, o participante deve repetir a sequência numérica ao contrário, isto é, 4 -3 - 1, para o exemplo anterior. Soares, Soares e Caixeta (2012) associam a aplicação da ordem direta do teste com a avaliação da atenção, enquanto que a ordem indireta se encarrega, mais especificamente, da MT.

Lemos et. al. (2014) não encontraram diferenças significativas na comparação entre grupo controle e grupo DA para os escores da versão direta do teste de ordenamento de dígitos. Por outro lado, houve diferença significativa entre grupo controle e DA na comparação de dados para a versão indireta do teste. Jerônimo (2016) também encontrou diferenças significativas na comparação entre grupos de CCL e controle e DA e controle para a ordem inversa do teste de ordenamento de dígitos. Em ambos os casos o grupo controle apresentou melhor desempenho, isto é, escores mais elevados para a tarefa, indicando que este teste pode auxiliar na distinção entre grupos de idosos saudáveis ou com algum tipo de comprometimento cognitivo. Jerônimo (2016) ainda comparou o desempenho de dois grupos de idosos saudáveis de escolaridades distintas (alta e baixa) no teste de ordenamento de dígitos (ordem inversa). A autora encontrou diferença significativa entre os grupos, com vantagem para o mais escolarizado, na avaliação de MT, sugerindo que o grau de escolaridade pode apresentar benefícios no funcionamento deste construto.

Soares, Brandão e Lacerda (2012) apontam para a importância da avaliação neuropsicológica de maneira a observarem-se as alterações no âmbito linguístico-cognitivo do idoso acometido pela DA. De acordo com elas, a avaliação não somente auxilia no processo de diagnóstico da demência de uma maneira diferenciada, mais voltada às particularidades de cada indivíduo, como também contribui para o tratamento da doença em âmbito multidisciplinar.

Levando em consideração a característica multidisciplinar da avaliação da linguagem e da cognição, apresentamos, na subseção seguinte, uma revisão de trabalhos que se utilizam da ferramenta da análise automática de textos, a qual tem potencial para aplicação com fins de diagnóstico de demências. Elaborada a partir da contribuição de diferentes áreas do conhecimento, a análise automática surge com a proposta de identificar, no presente estudo, marcadores de demências no discurso a partir do estudo de métricas linguísticas extraídas de softwares treinados com a técnica do Aprendizado de Máquina.

2.2.2 Análise automática como ferramenta para avaliação da produção de narrativas

Que a narrativa faz parte da vida em sociedade não há dúvidas. Discutimos também, até aqui, que processos demenciais como a DA podem interferir na qualidade deste tipo de produção textual. Desta forma, pesquisadores ao redor do mundo vêm elaborando ferramentas que visam a avaliar as produções narrativas de idosos com demências. Dentre estas ferramentas se destaca uma das mais recentes: a da análise automática da linguagem.

A análise automática consiste na utilização de softwares treinados para identificar, através da análise de textos, métricas linguísticas de nível lexical, sintático, semântico (CUNHA, 2015). Estes textos costumam ser produto da transcrição de produções orais. Além da parte escrita, também existem softwares que analisam o áudio originário da transcrição, para a extração de métricas acústicas, conforme veremos mais adiante na apresentação de alguns estudos empregando a técnica. Utilizamos como exemplo o software Coh-Matrix Dementia (ALUÍSIO; CUNHA; SCARTON, 2016), empregado na análise de dados desta dissertação. Esta ferramenta foi desenvolvida com o objetivo de auxiliar no diagnóstico de demências como a do tipo Alzheimer por meio da comparação do desempenho entre saudáveis, indivíduos com CCL e com DA em construtos linguísticos, chamados métricas. A

análise proposta pelo software é feita a partir da transcrição³ de textos orais e se baseia em métricas linguísticas (incluindo complexidade semântica, sintática, número de palavras por frase, entre outras) que podem indicar alterações na linguagem na comparação com saudáveis, como o início de um comprometimento da linguagem. Para tanto, o software é alimentado com dicionários, corpora e árvores sintáticas originárias da técnica de *parsing* (ALUÍSIO; CUNHA; SCARTON, 2016), as quais devem ser produzidas para cada língua, de acordo com suas particularidades estruturais. Apresentamos, a seguir, um quadro reproduzido a partir da aba *Help* disponível no site do Coh Metrix Dementia, em que são apresentadas todas as 73 métricas extraídas pelo software, seguidas do cálculo realizado pelo software para análise numérica.

Quadro 1: Descrição das 73 métricas extraídas pelo software Coh-Metrix Dementia

Contagem Básica	
Incidência de adjetivos	Incidência de adjetivos de um texto. (número de adjetivos/(número de palavras/1000))
Incidência de advérbios	Incidência de advérbios de um texto. (número de advérbios/(número de palavras/1000))
Incidência de palavras de conteúdo	Incidência de palavras de conteúdo. (número de palavras de conteúdo/(número de palavras/1000))
Índice Flesch	O Índice de Legibilidade de Flesch busca uma correlação entre tamanhos médios de palavras e sentenças e a facilidade de leitura. $ILF = 248.835 - [1.015 \times (\text{Número de palavras por sentença})] - [84.6 \times (\text{Número de sílabas do texto} / \text{Número de palavras do texto})]$
Incidência de palavras funcionais	Incidência de artigos, proposições, pronomes, conjunções e interjeições de um texto. (número de palavras funcionais/(número de palavras/1000)).
Sentenças por parágrafos	Número de sentenças dividido pelo número de parágrafos.
Sílabas por palavras de Conteúdo	Número médio de sílabas dividido pelos substantivos, verbos, adjetivos e advérbios do texto.
Palavras por sentenças	Número de palavras dividido pelo número de sentenças.
Incidência de substantivos	Incidência de substantivos em um texto. (número de substantivos/(número de palavras/1000)).
Número de parágrafos	São considerados parágrafos as quebras de linha.
Número de sentenças	Número de sentenças em um texto.

³ Mais características do processo de transcrição de textos para leitura pelo software Coh-Metrix Dementia serão apresentadas no capítulo que trata do método.

Número de palavras	Número de palavras em um texto.
Incidência de pronomes	Incidência de pronomes em um texto. (número de pronomes/(número de palavras/1000)).
Incidência de verbos	incidência de verbos em um texto. (número de verbos/(número de palavras/1000)).
Operadores Lógicos	
Incidência do operador lógico E	Incidência do operador lógico E em um texto. (frequência do operador lógico E / (número de palavras/1000))
Incidência do operador lógico SE	Incidência do operador lógico SE em um texto (desconsidera quando o SE é um pronome). (frequência do operador lógico SE / (número de palavras/1000)).
Incidência do operador lógico OU	Incidência do operador lógico OU em um texto. (frequência do operador lógico OU / (número de palavras/1000)).
Incidência de negação	Incidência de negações (não, nem, nenhum, nenhuma, nada, nunca e jamais). (número de negações/(número de palavras/1000)).
Incidência de operadores lógicos	Incidência de operadores lógicos (e, ou, se, negações e um número de condições) em um texto. (número de operadores lógicos/(número de palavras/1000)).
Frequências	
Frequência das palavras de conteúdo	Média de todas as frequências das palavras de conteúdo (substantivos, verbos, advérbios e adjetivos) encontradas no texto. O valor da frequência das palavras é retirado da lista de frequências do corpus Banco de Português (BP), compilado por Tony Sardinha da PUC-SP. Soma-se o valor da frequência de cada uma das palavras de conteúdo para obter-se o valor da métrica.
Frequência da palavra de conteúdo mais rara	Primeiramente identifica-se a menor frequência dentre todas as palavras de conteúdo (substantivos, verbos, advérbios e adjetivos) em cada sentença. Depois, calcula-se uma média de todas as frequências mínimas. A palavra com a menor frequência é a mais rara da sentença.
Hiperônimos	
Hiperônimos de verbos	Para cada verbo soma-se o número de hiperônimos e divide o total pelo número de verbos.
Pronomes, Tipos e Tokens	
Índice de Brunet	$W = N^{**} (V^{**} - 0.165)$ N é o número de palavras lexicais, e V é o número total de tokens usados. Os valores de W típicos variam entre 10 e 20, sendo que uma fala mais rica produz valores menores

Estatística de Honoré	$R = 100 * \log N / (1 - (V_1 / V))$ em que N é o número total de tokens, V ₁ é o número de palavras do vocabulário que aparecem uma única vez, e V é o número de palavras lexicais.
Cláusulas por sentença	Calcula o número médio de cláusulas por sentença. Definiu-se que uma cláusula em uma sentença é caracterizada pela presença de um sintagma verbal.
Pronomes por sintagmas	Média do número de pronomes que aparecem em um texto pelo número de sintagmas nominais.
Incidência de pronomes pessoais	Incidência de pronomes pessoais (eu, tu, ele/ela, nós, nós, eles/elas, você e vocês) em um texto. (número de pronomes pessoais/(número de palavras/1000))
Relação tipo por token	Número de palavras únicas dividido pelo número de tokens dessas palavras. Cada palavra única é um tipo. Cada instância desta palavra é um token.
Conectivos	
Incidência de Conectivos	Incidência de todos os conectivos que aparecem em um texto. Para esta métrica (e as demais que contam conectivos) compilam-se listas de conectivos classificados em duas dimensões. A primeira dimensão divide os conectivos em positivos e negativos (conectivos positivos estendem eventos, enquanto que conectivos negativos param eventos). A segunda dimensão divide os conectivos de acordo com o tipo de coesão: aditivos, temporais, lógicos e causais. (número de conectivos/(número de palavras/1000)).
Incidência de conectivos classificados como aditivos negativos	Incidência de todos os conectivos negativos que aparecem em um texto. (número de conectivos positivos/(número de palavras/1000))
Incidência de conectivos classificados como aditivos positivos	Incidência de todos os conectivos positivos que aparecem em um texto. (número de conectivos positivos/ (número de palavras/1000)).
Incidência de conectivos classificados como causais negativos	Incidência de todos os conectivos causais negativos que aparecem em um texto. (número de conectivos causais negativos/(número de palavras/1000))
Incidência de conectivos classificados como causais positivos	Incidência de todos os conectivos causais positivos que aparecem em um texto. (número de conectivos causais positivos/(número de palavras/1000)).
Conectivos lógicos negativos	Incidência de todos os conectivos lógicos negativos que aparecem em um texto. (número de conectivos lógicos negativos/(número de palavras/1000))

Conectivos lógicos positivos	Incidência de todos os conectivos lógicos positivos que aparecem em um texto. (número de conectivos lógicos positivos/(número de palavras/1000))
Incidência de conectivos classificados como temporais negativos	Incidência de todos os conectivos temporais negativos que aparecem em um texto. (número de conectivos temporais negativos/(número de palavras/1000))
Incidência de conectivos classificados como temporais positivos	Incidência de todos os conectivos temporais positivos que aparecem em um texto. (número de conectivos temporais positivos/(número de palavras/1000))
Ambiguidades	
Ambiguidade de adjetivos	Para cada adjetivo do texto soma-se o número de sentidos apresentados no TEP (base de dados lexical) e divide o total pelo número de adjetivos.
Ambiguidade de advérbios	Para cada advérbio do texto soma-se o número de sentidos apresentados no TEP (base de dados lexical) e divide o total pelo número de advérbios.
Ambiguidade de substantivos	Para cada substantivo do texto soma-se o número de sentidos apresentados no TEP (base de dados lexical) e divide o total pelo número de substantivos.
Ambiguidade de verbos	Para cada verbo do texto soma-se o número de sentidos apresentados no TEP (base de dados lexical) e divide o total pelo número de verbos.
Complexidade sintática	
Entropia cruzada	Média da entropia cruzadas das sentenças do texto. A entropia cruzada mede o nível de "surpresa" do modelo de língua diante da sentença. Valores maiores de entropia cruzada significam que a sentença possui combinações não usuais de palavras. Tipicamente, portanto, quanto maior a entropia cruzada de uma sentença, maior sua complexidade esperada. É calculada com base em um modelo de trigramas, treinado sobre um corpus que consiste na união dos corpora: Wikipedia, PLN-BR, LácioWeb, e Revista Pesquisa FAPESP.
Distância de dependência	Média da distância de dependência das sentenças de um texto. Para cada sentença do texto, a distância de dependência é calculada como a soma da distância entre as palavras associadas a cada relação de dependência. Utiliza-se da técnica de parser para calcular esta métrica.
Complexidade de Frazier	Para um dado texto, calcula-se a complexidade sintática de cada sentença, e então calcula-se a média desses valores. No cálculo da complexidade, utiliza-se o máximo da soma de trigramas dos escores das palavras. Também se utiliza de árvores sintáticas no cálculo desta métrica.

Complexidade de Yngve	Para um dado texto, calcula-se a complexidade sintática de cada sentença, e então calcula-se a média desses valores. No cálculo da complexidade, utiliza-se a média dos escores das palavras. Novamente se utiliza de árvores sintáticas.
Densidade Semântica	
Densidade de conteúdo	A densidade de conteúdo de um texto é calculada como o número de palavras de classe aberta (também denominadas palavras de conteúdo) dividido pelo número de palavras de classe fechada (ou palavras funcionais).
Constituintes	
Modificadores por sintagmas	Média do número de modificadores por sintagmas nominais. Consideramos como modificadores adjetivos, advérbios e artigos que participam de um sintagma. utiliza-se técnica de parser.
Incidência de sintagmas	Incidência de sintagmas nominais por 1000 palavras. (número de sintagmas/(número de palavras/1000)).
Palavras antes de verbos principais	Média de palavras antes de verbos principais na cláusula principal da sentença. Depende da avaliação de árvores sintáticas.
Anáforas	
Referência anafórica adjacente	Proporção de referências anafóricas entre sentenças adjacentes.
Referência anafórica	Proporção de referências anafóricas que se referem a um constituinte presente em até cinco sentenças anteriores.
Correferência	
Sobreposição de argumentos adjacentes	Proporção de sentenças adjacentes que compartilham um ou mais argumentos (substantivos, pronomes ou sintagmas nominais).
Sobreposição de argumentos	Proporção de todos os pares de sentenças que compartilham um ou mais argumentos (substantivos, pronomes ou sintagmas nominais).
Sobreposição de radicais de sentenças adjacentes	Proporção de sentenças adjacentes que compartilham radicais.
Sobreposição de radicais de palavras	Proporção de todos os pares de sentenças que compartilham radicais.
Sobreposição de palavras de conteúdo em sentenças adjacentes	Proporção de sentenças adjacentes que compartilham palavras de conteúdo (substantivos, verbos, adjetivos e advérbios).
Análise Semântica Latente (LSA)	
Média entre sentenças adjacentes	Média de similaridade entre pares de sentenças adjacentes no texto. Utiliza-se do mesmo modelo de corpus da Entropia cruzada.
Desvio padrão entre sentenças adjacentes	Desvio padrão de similaridade entre pares de sentenças adjacentes no texto.
Média de similaridade entre todos os pares de sentença no texto	Média de similaridade entre todos os pares de sentenças no texto.

Desvio padrão entre todos os pares de sentenças	Desvio padrão de similaridade entre todos os pares de sentenças no texto.
Média entre parágrafos adjacentes	Média de similaridade entre pares de parágrafos adjacentes no texto. Esta métrica é calculada do mesmo modo que a média entre sentenças adjacentes, mas utilizando-se parágrafos, ao invés de sentenças, como unidades.
Desvio padrão entre parágrafos adjacentes	Desvio padrão de similaridade entre pares de parágrafos adjacentes no texto. Esta métrica é calculada do mesmo modo que o desvio padrão entre sentenças adjacentes, mas utilizando-se parágrafos, ao invés de sentenças, como unidades.
Média de <i>givenness</i> das sentenças	Média do <i>givenness</i> da cada sentença do texto, a partir da segunda. Se o texto possui apenas uma sentença, define-se a métrica como 0,0. Define-se o <i>givenness</i> de uma sentença como a similaridade LSA entre a sentença e todo o texto que a precede.
Desvio padrão de <i>givenness</i> das sentenças	Desvio padrão do <i>givenness</i> da cada sentença do texto, a partir da segunda. Se o texto possui apenas uma sentença, define-se a métrica como 0,0.
Média do <i>span</i> das sentenças	Média do <i>span</i> da cada sentença do texto, a partir da segunda. Se o texto possui apenas uma sentença, define-se a métrica como 0,0. O <i>span</i> de uma sentença, assim como o <i>givenness</i> , é uma forma de medir a proximidade entre uma sentença e o contexto que a precede. A diferença, em termos simples, consiste no fato de que o <i>span</i> procura capturar a similaridade não apenas com o conteúdo explícito apresentado anteriormente no texto, mas também com tudo o que se pode inferir com base nesse conteúdo.
Desvio padrão do <i>span</i> das sentenças	Desvio padrão do <i>span</i> da cada sentença do texto, a partir da segunda. Se o texto possui apenas uma sentença, define-se a métrica como 0,0.
Disfluências	
Número médio de disfluências	Número de palavras em emissões disfluentes dividido pelo total de palavras (inclusas as disfluências). Disfluências são marcadas com <disf> e </disf>.
Número médio de palavras vazias	Número de palavras em emissões vazias dividido pelo total de palavras (inclusas as emissões vazias). Emissões vazias são marcadas com <empty> e </empty>.
Duração média das pausas	Duração total das pausas dividida pelo número total de palavras. Pausas são marcadas como ((pausa XX segundos)).
Duração média das pausas curtas	Número de pausas curtas dividido pelo número total de palavras. Pausas curtas são indicadas por três pontos (...).

Número médio de prolongamento de vogais	Número de prolongamentos de vogais dividido pelo total de palavras. Prolongamentos de vogais são marcados com dois pontos (:).
Número médio de palavras repetidas	Número de palavras repetidas dividido pelo total de palavras (incluindo-se as repetidas). Palavras são consideradas repetidas apenas quando aparecem em sequência. Por exemplo, em Maria foi foi ao mercado, é contada uma repetição da palavra foi, mas em Maria foi ao mercado, e então foi à padaria, não é contada repetição de foi, pois ambas não ocorrem em sequência. Não são contadas repetições no interior de emissões disfluente.
Densidade de ideias total	Número de proposições presentes no texto, para cada dez palavras. Para o cálculo das proposições, não são levadas em conta proposições vazias ou disfluente, e o cálculo é feito sobre o texto revisado, para melhor desempenho da ferramenta de extração.

Fonte: O autor (2019), adaptado de <<http://143.107.183.175:22380/help>>.

O Coh-Metrix Dementia surge no contexto brasileiro como aliado nesta prática da análise automática de texto, porém não é o único software com a mesma proposta. Conforme pesquisa na literatura, foi possível identificar que esta ferramenta tem se mostrado eficaz na delimitação de marcadores de diferentes tipos demências no texto, e vem sendo utilizada por pesquisadores em diversos países. Sendo assim, elencamos a seguir alguns estudos realizados ao longo dos últimos cinco anos que têm como tema central a análise automática de narrativas produzidas por idosos em processos demenciais, em comparação com idosos saudáveis.

Os estudos de Orimaye et. al. (2017) e Fraser et. al. (2016) apresentam dados de participantes acometidos pela DA e extraídos da mesma base de dados, o *DementiaBank*. Os dados analisados em ambas as pesquisas foram as transcrições de uma narrativa oral, proposta a partir de uma imagem (*Cookie Theft*), conforme sugerida na bateria *Boston Diagnostic Aphasia Examination* (BDAE) (GOODGLASS; KAPLAN, 1983). O que difere os estudos são algumas métricas extraídas, que suscitam resultados parcialmente distintos entre os estudos,

Fraser et. al. (2016) desenvolveram um trabalho com o objetivo de descrever o estado da arte da acurácia para métricas incluídas na análise automática de textos. Utilizando-se dos textos narrativos de 163 participantes com DA, em comparação com 97 participantes controle, os autores extraíram métricas distintas que pudessem atuar como marcadores da DA no

discurso. Além disso, também utilizaram o áudio que deu origem às transcrições, de maneira a treinar um classificador eletrônico a diferenciar participantes com DA de participantes saudáveis não somente no nível do texto escrito, como também da oralidade. 370 métricas foram consideradas para capturar a magnitude do fenômeno linguístico, porém apenas uma média de 50 métricas são apresentadas no artigo. Dentre elas, apontamos as principais, que são complexidade sintática, constituintes gramaticais (como sintagmas), riqueza lexical, nível de informação, repetição e acústica. A necessidade de métricas acústicas se dá pelo fato de que o discurso oral de pessoas com demências geralmente tem características sonoras recorrentes, como pausas, hesitações e repetições.

Os resultados apontam para os seguintes marcadores da DA na comparação com controles saudáveis: dificuldades semânticas no nível do léxico (alta produção de pronomes referenciais em detrimento de substantivos; repetições de palavras e vocabulário restrito); anomalias acústicas (como longos períodos de pausa); dificuldades sintáticas (baixa produção de verbos, produção de fragmentos de uma sentença em detrimento de sentenças completas, produção de “não-palavras”); dificuldades em compartilhar informações (dificuldade em produzir palavras principais para a descrição das imagens do estímulo, pouca produção de sintagmas preposicionais, indicando pobreza de detalhes).

O estudo de Orimaye et. al. (2017) objetivou treinar um sistema computacional para reconhecer e diagnosticar marcadores da DA através do discurso. Como o grupo de Fraser, esses autores utilizaram a base de dados *DementiaBank*, da qual extraíram as transcrições de narrativas produzidas por dois grupos, sendo um de 99 participantes acometidos pela DA e outro com 99 controles saudáveis. Foram utilizados algoritmos de aprendizado de máquina para treinar métricas sintáticas, lexicais e de n-grama através da técnica de *parsing*, isto é, análise computacional de árvores sintáticas com base em Processamento de Linguagem Natural (PNL). Através da análise estatística das métricas, os pesquisadores puderam identificar que, em geral, são marcadores do discurso de pessoas acometidas com DA a dificuldade de produção de sentenças complexas, a repetição de palavras, assim como a revisão de agramaticalidades, gerando mais fala do que os participantes saudáveis.

Toledo et. al. (2017) analisaram as transcrições de narrativas obtidas através de tarefa de produção oral de narrativa a partir de 22 imagens que formavam a história da Cinderela. Foram incluídos no estudo 60 participantes divididos em três grupos: um de pessoas com DA; um de pessoas com CCL amnésico; e um de controles saudáveis. O objetivo do estudo, diferentemente dos outros dois apresentados anteriormente, era de analisar comprometimentos

especificamente relacionados à formação da macroestrutura do texto. Para tanto, dividiram-se as imagens em 28 proposições, das quais 11 métricas semânticas relacionadas foram extraídas pelo software Coh-Metrix Dementia. Os resultados apontam para produção mais comprometida de narrativas orais na DA, em comparação com os outros grupos. Os participantes com DA demonstraram uma tendência à descrição de cenas, em detrimento à sequência narrativa proposta pelas figuras, assim como mudança de tópico e inclusão de informações extratextuais, o que configurou em dificuldades na formação da coerência global do texto e problemas na construção da estrutura narrativa. Não foram observadas diferenças entre participantes acometidos com CCL amnésico e participantes saudáveis.

Fraser et. al. (2012), em um estudo com pacientes com afasia progressiva primária (APP), extraíram 58 métricas sob o viés da análise automática de transcrições oriundas da produção de narrativa oral a partir de uma sequência de 22 figuras que compõem a história da Cinderela. As narrativas foram produzidas por 24 participantes acometidos pela variante semântica (APP-vs), também conhecida como Demência Semântica (DS) ou pela variante não-fluente (APP-vnf) e 16 participantes saudáveis. Os resultados mostram que os classificadores treinados pela técnica do aprendizado de máquina são altamente acurados em distinguir os grupos com demência do grupo controle, e um pouco menos, mas ainda assim acurados, em distinguir entre os tipos de APP. Dentre as métricas mais significativas na distinção entre os grupos apontam-se principalmente aquelas relacionadas à sintaxe e ao léxico. Os participantes acometidos pela APP-vs apresentaram, em comparação com participantes controle, maior uso de palavras de maior frequência (especialmente substantivos) e familiaridade, muitos advérbios e pronomes demonstrativos, o que demonstra baixa riqueza lexical. Já os participantes acometidos pela APP-vnf apresentaram diminuição no ritmo de fala, uso de palavras mais curtas e também tiveram uso de palavras de conteúdo (especialmente verbos) mais frequentes e familiares do que os participantes saudáveis. Entre os grupos de participantes acometidos por APP, uma das principais diferenças apontadas pela análise das transcrições é um maior uso de frases subordinadas por participantes acometidos pela DS em relação a participantes com APP-vnf. Entretanto, este resultado pode ter sido afetado pela habilidade do sistema em identificar (ou não) sentenças subordinadas apropriadamente (FRASER, et. al., 2012).

Outro estudo que aborda a análise de discursos na APP é o de Marcotte et. al. (2017). Os autores apresentam duas análises principais: a primeira dá conta das métricas extraídas das transcrições de textos narrativos; a segunda correlaciona os dados linguísticos com as

métricas do exame DTI (*Diffusion Tensor Imaging*), que analisa o grau de preservação e acometimento de feixes neuronais (ou tratos) com coleta por meio de MRI (Imagem por Ressonância Magnética) estrutural. Para tanto, os autores analisaram a produção narrativa de 25 participantes diagnosticados com Afasia Progressiva Não-Fluente (APP-vnf) ou Afasia Progressiva Primária de variante semântica (APP-vs), e 18 participantes controle saudáveis pareados por idade e escolaridade. A tarefa utilizada para coleta de dados foi uma entrevista, na qual os participantes respondiam a perguntas gerais sobre o dia a dia. Uma dessas perguntas (sobre o que os participantes faziam todos os dias) rendia respostas mais longas, com mais de 100 palavras, que então foram transcritas para serem analisadas. Foram incluídas na análise 1 métrica de fluência, 26 métricas de complexidade sintática, 17 métricas lexicais, 16 métricas semânticas e 2 métricas de riqueza lexical. Já as imagens estruturais do cérebro foram analisadas a partir da quantificação da difusividade radial (DR), sendo que quanto maior o número da DR, maior o grau de degeneração da mielina, fator que indica comprometimento da região afetada (MARCOTTE, et. al., 2017). Os principais resultados da análise automática apontaram para diminuição de complexidade sintática em participantes com APP-nf em comparação com controles; baixa riqueza lexical em participantes com APP-vs em comparação com controles; e apenas uma diferença no número de palavras produzidas por minuto entre as duas variantes de APP. As correlações entre as métricas do discurso e danos à substância branca cerebral foram mínimas na comparação entre os grupos com demência. Por outro lado, houve diferenças significativas na correlação de dados entre os grupos demenciais e controles saudáveis. Em geral, a APP-vs caracteriza-se por danos bilaterais nas conexões ventrais, fato que pode diferenciar a APP-vs da APP-vnf. Identificou-se correlação significativa entre o uso de verbos e maior DR no fascículo longitudinal superior (FLS) esquerdo, o que aponta para importância do FLS no processamento verbal. Também se identificou que a baixa riqueza lexical pode estar relacionada a comprometimentos no fascículo longitudinal inferior (FLI) esquerdo, o que corrobora estudos prévios pontuados pelos autores.

A partir da análise dos resultados dos estudos reportados é possível identificar similaridades e diferenças no que diz respeito a marcadores discursivos de demências extraídos de programas treinados pelo aprendizado de máquina para distinguir produções de pessoas saudáveis e com demência. Os estudos com DA parecem apontar para declínios na estrutura narrativa, assim como nas questões de coerência global e macroestrutura do texto, conforme indica o trabalho de Toledo et. al. (2017). Por outro lado, os estudos de Orimaye et.

al. (2017) e Fraser et. al. (2016) apontam principalmente para comprometimentos relacionados à produção de sentenças mais complexas, repetição de palavras e inserção de comentários extratextuais. O trabalho de Fraser et. al. (2016) ainda atenta para a questão das métricas acústicas, que sinalizam a ocorrência frequente de pausas e revisões de erros gramaticais comuns às pessoas acometidas pela DA.

O mesmo ocorre na comparação entre os estudos com APP. Tanto a pesquisa de Fraser et. al. (2012), quanto a de Marcotte et. al. (2017) chamam a atenção para a baixa riqueza lexical e uso de vocabulário mais frequente na SD, e diminuição da complexidade sintática na APP-vnf. Fraser et. al. (2012) ainda apontam para baixa riqueza lexical também na APP-vnf e fatores de distinção entre as variações da APP (associados à produção de sentenças subordinadas) não corroborados pelo estudo de Marcotte, et. al. (2017). Este último contribui para a relação dos dados das métricas aos estudos dos correlatos neuronais a elas associados. Apresentamos, no Quadro 2, um resumo dos achados principais de cada artigo aqui discutido.

Quadro 2: Resumo de estudos que abordam a análise automática de textos

Estudo	Marcotte, et. al. 2017	Orimaye, et. al. 2017	Toledo, et. al. 2017	Fraser, et. al. 2016	Fraser, et. al. 2014
Demência	APP (vs e vnf).	DA.	DA; CCL.	DA.	APP (vs e vnf).
Objetivos	Investigar a correlação entre redes neurais de substância branca associadas a falhas no discurso de pessoas acometidas com APP-vnf e APP-vs, utilizando técnicas de tractografia e análise automática computacional.	Apontar marcadores da DA no discurso produzido a partir da tarefa de descrição de uma figura através da análise automática computacional.	Investigar alterações no nível de macroestrutura do discurso através de ferramenta computacional de análise automática.	Apontar marcadores da DA no discurso produzido a partir da tarefa de descrição de uma figura através da análise automática computacional.	Apresentar um método de análise e classificação de discurso na APP com base em técnica computacional.
Participantes	25 participantes com APP-vnf ou APP-vs e 18 participantes saudáveis.	99 participantes com DA e 99 participantes saudáveis.	60 participantes em três grupos: DA; CCL amnésico e saudáveis.	167 participantes com DA e 97 participantes saudáveis.	24 participantes com APP-vnf ou APP-vs e 16 participantes saudáveis.
Tarefa de produção discursiva	Entrevistas gravadas, das quais textos com no mínimo 100 palavras foram transcritos.	Transcrições extraídas do DementiaBank, em tarefa de descrição da imagem “Cookie Theft”.	Narrativas produzidas a partir de 22 imagens que formam a história da Cinderella foram transcritas.	Transcrições extraídas do DementiaBank, em tarefa de descrição da imagem “Cookie Theft”.	Narrativas produzidas a partir de 22 imagens que formam a história da Cinderella foram transcritas.
Número de	70 métricas	23 métricas	11 métricas	50 métricas	58 métricas

métricas analisadas	divididas entre complexidade sintática, léxico e semântica.	divididas entre sintaxe e léxico.	semânticas no nível de macroestrutura do texto.	divididas entre sintáticas, lexicais, semânticas, acústicas, entre outras, são apresentadas na correlação entre transcrições e áudios.	divididas principalmente entre classes de palavras, familiaridade, frequência, entre outras.
Principais resultados	Comprometimentos nas métricas sintáticas (APP-vnf em comparação com saudáveis); baixa riqueza lexical (APP-vnf em comparação com saudáveis); não houve diferenças nas métricas entre APP-vnf e APP-vs. Substância branca: correlação entre DR e uso de verbos indicam DR no FLS; baixa riqueza lexical indica DR no FLI.	O modelo identificou marcadores para distinguir DA de saudáveis nas áreas de sintaxe, léxico e n-grama (sintagmas).	Participantes com DA tiveram pior desempenho em coerência global, modalização, e estrutura narrativa, em comparação com saudáveis. Não se identificou diferenças entre CCL e saudáveis.	O modelo pôde identificar fatores marcadores de DA no discurso, em comparação com saudáveis, principalmente nas questões semânticas, acústicas, sintáticas e informativas. A acurácia das distinções ficou em 81%.	Maior uso de palavras de maior frequência e familiaridade nos discursos de pessoas com APP-vnf e APP-vs em relação a saudáveis. Pessoas com APP-vnf foram mais lentas e usaram palavras mais curtas que saudáveis. Pessoas com APP-vs produziram mais demonstrativos e advérbios e menos substantivos que saudáveis.

Legenda: APP-vnf: Afasia Progressiva Primária – variante não-fluente; APP-vs: Afasia Progressiva Primária – variante semântica; CCL: Comprometimento Cognitivo Leve; DA: Doença de Alzheimer; DR: difusão radial; FLI: Fascículo Longitudinal Inferior; FLS: Fascículo Longitudinal Superior.

Fonte: O autor (2019).

Em geral, todos os estudos apresentados parecem ter obtido sucesso no treinamento de classificadores automáticos pelo aprendizado de máquina em detectar marcadores discursivos das demências analisadas. Este resultado atenta para a importância de estudos desta natureza no diagnóstico precoce de demências, fator que pode contribuir para um tratamento mais focalizado nas necessidades da pessoa acometida.

Sendo assim, nesta dissertação, um dos objetivos é avaliar a produção oral de narrativa de pessoas acometidas pela Doença de Alzheimer com o auxílio da análise automática. Para tanto, optou-se pelo uso do software Coh-Metrix Dementia, por entender que este possui grande acurácia na detecção de marcadores discursivos da DA nas produções, o mesmo software utilizado no estudo de Toledo et al. (2017). Optamos por este software por ser uma

versão programada para analisar textos escritos em português brasileiro, especialmente voltada ao texto de pessoas acometidas por demências como DA ou declínios cognitivos como o CCL, sendo relevante ao contexto do nosso estudo. Além disso, acreditamos que a análise automática, no nível da sentença, pode ser grande aliada ao estudo da microestrutura, conforme definida por Kintsch e Van Dijk (1978) e Van Dijk e Kintsch (1983).

Assim, nesta dissertação buscamos responder às seguintes questões norteadoras:

Q1. Há diferenças no resultado da análise da micro e macroestrutura de narrativas produzidas por idosos saudáveis, tendo como variável a escolaridade?

Q2. Terão os idosos saudáveis de alta escolaridade melhor desempenho na correlação entre a tarefa linguística de produção de narrativa (nos níveis de micro e macroestrutura) e construtos cognitivos em comparação com idosos saudáveis de baixa escolaridade?

Q3. Há diferenças no resultado da análise da micro e macroestrutura de narrativas produzidas por idosos saudáveis na comparação com pessoas acometidas pela DA?

Q4. Terão os idosos saudáveis melhor desempenho na correlação entre a tarefa linguística de produção de narrativa (nos níveis de micro e macroestrutura) e construtos cognitivos em comparação com idosos acometidos pela DA?

No capítulo seguinte será abordado o desenho metodológico dos estudos experimentais conduzidos nesta pesquisa, assim como a delimitação dos objetivos (gerais e específicos), hipóteses e participantes de cada um deles. Além disso será descrita a tarefa utilizada para coleta de dados linguísticos, assim como os testes neuropsicológicos adotados no protocolo para mensurar os construtos cognitivos avaliados nos estudos. Por fim, apresenta-se a forma de análise dos dados adotada para cada teste/tarefa.

3 ESTUDO EXPERIMENTAL

De maneira a melhor organizar e apresentar os dados obtidos, esta dissertação se divide em dois estudos: Estudo 1 e Estudo 2. Sendo assim, neste capítulo trataremos do delineamento dos estudos, apresentando, a seguir, os objetivos, hipóteses e método adotados para a condução de ambos.

3.1 OBJETIVOS

3.1.1 Objetivo geral da dissertação

Investigar se há diferenças na produção oral de narrativas na comparação entre idosos saudáveis segundo a escolaridade, bem como entre saudáveis e acometidos por DA nos níveis micro- e macroestruturais, correlacionando-se esses dados a construtos de memória (episódica e de trabalho).

3.1.2 Objetivos específicos – Estudo 1

- Investigar se há diferenças na produção oral de narrativas no nível macroestrutural na comparação entre idosos saudáveis de alta e de baixa escolaridade.
- Investigar se há diferenças na produção oral de narrativas na comparação entre idosos saudáveis de alta e de baixa escolaridade a partir da análise da microestrutura, utilizando como objeto de análise os construtos linguísticos da avaliação automática de texto.
- Investigar se há diferenças entre grupos no desempenho em tarefas de memória episódica e de trabalho, assim como na relação entre os níveis micro e macroestruturais das narrativas produzidas por idosos saudáveis de alta versus baixa escolaridade e estes construtos cognitivos.

3.1.3 Objetivos específicos - Estudo 2

- Investigar se há diferenças na produção oral no nível da macroestrutura da narrativa entre idosos saudáveis e idosos diagnosticados com DA, ambos de baixa escolaridade.

- Investigar se há diferenças na produção oral do nível de microestrutura da narrativa, utilizando-se dos construtos linguísticos (métricas) da avaliação automática de texto, entre idosos saudáveis e idosos diagnosticados com DA, ambos de baixa escolaridade.
- Investigar se há diferenças entre grupos no desempenho em tarefas de memória episódica e de trabalho, assim como na relação entre os níveis micro e macroestruturais das narrativas produzidas por idosos saudáveis e idosos diagnosticados com DA e estes construtos cognitivos.

3.2 HIPÓTESES

3.2.1 Hipóteses - Estudo 1

- Idosos saudáveis de alta escolaridade obterão desempenho superior (ou seja, escores mais elevados nas tarefas de compreensão da macroestrutura do texto) na produção do nível de macroestrutura na comparação com idosos saudáveis de baixa escolaridade.
- Idosos saudáveis de baixa escolaridade demonstrarão desempenho inferior na produção do nível de microestrutura do texto, observado na análise das métricas, em especial na complexidade sintática, entre outros.
- Espera-se que não haja diferenças entre os grupos saudáveis originadas pelo grau de escolaridade no desempenho da memória episódica, mas que haja diferença entre grupos no desempenho da memória de trabalho; na correlação entre o desempenho linguístico dos grupos (métricas de micro- e total de macroestrutura) e construtos cognitivos (memória de trabalho e episódica), espera-se que haja correlação positiva.

3.2.2 Hipóteses – Estudo 2

- Idosos saudáveis obterão melhor desempenho (escores mais altos nas tarefas de compreensão da macroestrutura do texto) na produção do nível da macroestrutura na comparação com idosos diagnosticados com DA.
- Pessoas acometidas pela DA demonstrarão desempenho inferior no nível da microestrutura do texto, observada na análise das métricas, em especial no uso de pronomes, na complexidade sintática, entre outras.

- Haverá diferenças significativas entre os grupos para tarefas de memória episódica e de memória de trabalho, além de uma correlação positiva entre o desempenho linguístico dos grupos de idosos saudáveis e de idosos acometidos pela DA (métricas de micro- e total de macroestrutura) e construtos cognitivos (memória de trabalho e episódica); especificamente, escores mais baixos nos construtos da avaliação neuropsicológica do grupo DA na comparação com o saudável correlacionarão com desempenho linguístico mais baixo, tanto na produção da micro- quanto da macroestrutura.

3.3 MÉTODO

Esta pesquisa é de cunho experimental, de abordagem qualitativa e quantitativa e natureza exploratório-explicativa. As próximas seções darão conta de caracterizar os grupos de participantes, assim como delimitar os testes neuropsicológicos e linguísticos realizados.

3.3.1 Aspectos éticos

O protocolo adotado nesta pesquisa foi aprovado pelo Comitê de Ética em Pesquisa sob o número de parecer 560.073, no dia 06 de março de 2014, registro CAAE 21006913.0.0000.5336.

3.3.2 Participantes

Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) (ANEXO A). Visando a uma melhor visualização dos grupos, a relação de participantes envolvidos será apresentada a partir da divisão entre os estudos.

Estudo 1: dois grupos de idosos saudáveis, sendo um de alta (+ 9 anos de escolaridade) (SAE), composto de 28 participantes, entre 62 e 82 anos; e um de baixa escolaridade (de 2 a 8 anos de escolaridade) (SBE), composto de 35 participantes, pareados por idade entre 63 e 82 anos.

Estudo 2: dois grupos de idosos de baixa escolaridade (de 2 a 8 anos de escolaridade), sendo um grupo controle de idosos saudáveis (GC), composto de 35 participantes (o mesmo do Estudo 1); e um de idosos acometidos pela DA (DA), composto de 12 participantes, pareados por idade entre 66 e 84 anos.

Os critérios de inclusão e exclusão de participantes são apresentados na próxima seção, e, em seguida, são apresentadas as tarefas realizadas para caracterização das amostras.

3.3.2.1 Critérios de inclusão e exclusão de participantes

Para a seleção dos participantes, alguns critérios foram adotados. Dentre os critérios de inclusão encontram-se a idade (entre 62 e 84 anos); escolaridade (entre 2 e 8 anos de ensino regular no grupo de baixa escolaridade; e +9 anos para o grupo de alta escolaridade); grau de demência 1 ou 2, medido pela escala *Clinical Dementia Rating* (CDR) (MORRIS, 1993; MONTAÑO, RAMOS, 2005) (para o grupo com DA). Todos os participantes que compõem o grupo DA foram diagnosticados previamente por profissionais da área da neurologia, e encaminhados para a pesquisa a partir do ambulatório do SUS de um hospital da cidade de Porto Alegre. Dos critérios de exclusão para os dois estudos destacamos história de alcoolismo, de doença neurológica (que não DA), depressão não tratada, problemas de visão e/ou audição não corrigidos e grau de escolaridade abaixo de 2 anos de ensino regular. Também foram excluídos participantes bilíngues. A inclusão e/ou exclusão de participantes se deu por meio do resultado da administração de instrumentos de caracterização de amostra, que serão apresentados detalhadamente na seção a seguir.

3.3.3 Instrumentos de seleção e caracterização dos participantes dos Estudos 1 e 2

Os participantes saudáveis desta pesquisa foram recrutados através de divulgação do estudo em grupos de convivência para idosos, igrejas, eventos, entre outros. Já os participantes acometidos pela DA foram indicados para esta pesquisa a partir do setor de neurologia de um hospital que atende a rede pública. Por esta razão, o diagnóstico de possível DA, assim como o grau de acometimento da demência, foram estabelecidos por neurologistas integrantes do grupo de pesquisa por meio de avaliação clínica, a partir de exames

neurológicos, neuropsicológicos e de neuroimagem. As sessões de coletas de dados foram realizadas em dois encontros presenciais. O primeiro, de duração média de 1h30min, tinha caráter de entrevista, em que foram aplicados os testes de caracterização de amostra, neuropsicológicos e linguísticos. A partir deste primeiro encontro, os participantes eram convidados, se aptos, a um segundo encontro para coleta de dados de neuroimagem, que aconteceram no Instituto do Cérebro do Rio Grande do Sul, com duração média de uma hora (entre preparação e exame). Nesta pesquisa não serão incluídos os dados de neuroimagem (VBM, Voxel-Based Morphometry, e DTI, Diffusion Tensor Imaging) coletados. Além disso, outros dados linguísticos, neuropsicológicos e de funcionalidade também foram coletados por meio da aplicação de tarefas, como parte do estudo guarda-chuva em que este se insere, mas como não são foco deste estudo não serão aqui reportados. A seguir são detalhados os testes e questionários de caracterização da amostra dos estudos aqui descritos.

a) Questionário de condições de saúde, aspectos sociodemográficos e socioculturais (Adaptado de Fonseca et al., 2012)

O questionário tem por objetivo mapear dados relativos às condições clínicas dos participantes, como medicamentos utilizados e tratamentos médicos realizados, assim como morbidades e/ou patologias das quais o participante fosse acometido.

b) Mini-Exame do Estado Mental (MEEM) (CHAVES; IZQUIERDO, 1992)

O Mini-Exame do Estado Mental – MEEM é a adaptação para o português brasileiro do teste que investiga possíveis indícios de quadros demenciais. É um exame breve que inclui itens que avaliam a linguagem, a atenção e a orientação espaço-temporal, entre outros construtos. Ao todo, o participante pode acumular 30 pontos ao responder às perguntas, sendo o ponto de corte acima 22 pontos para participantes saudáveis com escolaridade entre 1 e 11 anos, e acima de 28 pontos para participantes saudáveis com mais de 11 anos de escolaridade.

c) Escala de Depressão Geriátrica (GDS) (ALMEIDA; ALMEIDA, 1999)

Composta de 15 perguntas que investigam a motivação, a autoestima, o pessimismo, entre outros, a Escala de Depressão Geriátrica – GDS tem como objetivo apontar uma suspeita de quadro depressivo no participante. A soma das respostas negativas ou afirmativas é necessária para a avaliação do participante. Quanto maior o resultado da soma das respostas, maior a chance de o participante apresentar algum quadro de depressão.

d) Questionário de hábitos de leitura e escrita (adaptado de Pawlowski et. al., 2012)

O questionário investiga os hábitos atuais e anteriores de leitura e escrita dos participantes. O objetivo desta distinção é avaliar, no caso dos participantes com DA, a relação deles com a leitura e a escrita antes e durante o quadro demencial. Os primeiros protocolos aplicados nesta pesquisa incluíam somente a versão dos hábitos atuais de leitura e escrita. Por representarem a maioria das respostas coletadas, optou-se pela análise dos dados relativos aos hábitos atuais somente. Sendo assim, pontua-se a frequência da seguinte forma: todos os dias (4 pontos), alguns dias por semana (3 pontos), uma vez por semana (2 pontos), raramente (1 ponto) e nunca (0 pontos).

e) Questionário para caracterização do status socioeconômico e condição social (ABEP, 2014)

Utiliza-se o Questionário do Critério Brasil (Associação Brasileira de Empresas e Pesquisa, 2014), que trata de quantificar itens como eletrodomésticos, eletroeletrônicos, automóveis, assim como banheiros, grau de instrução do (a) chefe da família, e a presença de empregada mensalista. A pontuação se dá pela quantidade de cada item declarada pelo participante, e tem como objetivo traçar uma estimativa do seu padrão socioeconômico. A subdivisão das classes se dá da seguinte forma: A1 = 42-46; A2 = 35-41; B1 = 29-34; B2 = 23-28; C1 = 18-22; C2 = 14-17; D = 8=13; E = 0-7.

3.3.4 Instrumentos para a coleta de dados neuropsicológicos

Para avaliação de desempenho neuropsicológico foram utilizados os seguintes testes, compilados na Bateria de Avaliação da Linguagem no Envelhecimento (BALE) (HÜBNER, et. al., no prelo).

a) Avaliação de memória episódica verbal

Utiliza-se a tarefa de Aprendizagem Verbal adaptada do teste *Free and Cued Selective Reminding Test with Immediate Recall* (FCSRT+IR) (BUSCHKE, GROBER, 2012), que é um teste de nomeação e memorização de figuras de diferentes categorias semânticas, uma subtarefa da BALE. Ao todo, 16 figuras são apresentadas em quatro cartelas, sendo quatro figuras em cada cartela. Primeiramente se realiza uma rodada de treino, em que o participante deve identificar a figura pela categoria semântica (por exemplo, “aponte a fruta”, “aponte o

jogo”, etc.). Após a identificação, o participante deve nomear a figura, por exemplo “que fruta é esta?”, “morango”, e assim por diante. Após o treino, é solicitado que o participante conte até o número vinte, de trás para a frente. Aí então inicia-se o teste, em que o participante deve recordar (sem pista, primeiramente) as figuras que acabou de ver nas cartelas. Após isso, o aplicador pode dar uma pista, que é a categoria semântica da figura a qual o participante não conseguiu recordar da primeira vez. Este processo é repetido por mais duas vezes, sendo que entre a penúltima e a última vez há um intervalo de 20 min para avaliação da recordação tardia. A pontuação total do teste é de 48 pontos.

b) Avaliação de memória de trabalho

Utiliza-se o teste *Digit Span*, ou Ordenamento Ascendente de Dígitos (subteste do NEUPSILIN, de FONSECA et al., 2009). Neste teste, o participante deve repetir uma sequência numérica em duas maneiras, direta e indireta. Na primeira etapa (direta), o aplicador lê uma sequência que pode ir de dois a oito números, e o participante deve repeti-la na mesma ordem lida pelo aplicador. Na segunda etapa (indireta), o aplicador lê uma sequência de números que vai de dois a sete números, porém o participante deve ler na ordem inversa, isto é, começando de trás para a frente. A pontuação total do teste é de 30 pontos, sendo 16 na primeira etapa e 14 na segunda.

3.3.5 Instrumento para a coleta de dados linguísticos

Para a avaliação de produção oral de narrativa, nesta pesquisa emprega-se a história *The Dog Story* (LE BOEUF, 1976). A tarefa, que integra a BALE, consiste na elaboração de uma narrativa oral a partir de uma sequência de imagens (Figura 1) que foram exibidas aos participantes uma ao lado da outra, em cenas de aproximadamente 5cm x 5cm cada, de maneira a reforçar a ideia de sequência de informações, comum ao gênero narrativo.

Figura 1: The Dog Story

Fonte: Le Boeuf (1976)

O avaliador indica que o participante deve analisar as figuras e elaborar uma história a partir delas. Quando o participante demonstra que compreendeu a instrução, inicia a produção da narrativa. O áudio é gravado para ser submetido à análise. Não há tempo limite para a execução da tarefa, nem interferência do aplicador, a não ser por frases de encorajamento como “e o que mais?”, “hum...”, etc.

A partir da gravação em áudio da produção narrativa do participante, é possível pontuar os aspectos da macroestrutura do texto, conforme o manual de aplicação da BALE (Quadro 3). No item ESTRUTURA estão descritas as seis macroproposições que dão conta dos aspectos principais da história (macroestrutura), e foram divididas levando em consideração os estágios de elaboração da estrutura narrativa conforme Adam (2008). Estas proposições foram avaliadas por juízes especialistas como sendo as principais unidades de sentido obtidas a partir da sequência de imagens apresentada. A pontuação máxima dos aspectos macroestruturais é de 12 pontos. Nos itens de desempenho linguístico são avaliadas características mais gerais do texto, como o número de pausas, repetições, estrutura narrativa, entre outros. As pontuações máximas são 09 (Desempenho Linguístico I) e 06 (Desempenho Linguístico II).

Quadro 3: Critérios de avaliação da tarefa de produção oral de narrativa

NARRATIVA			
<i>Vou lhe mostrar uma história composta de sete figuras. Cada figura representa um momento da história, que tem início, meio e fim. Vou lhe pedir para olhar bem cada uma das figuras e tentar entender a história que elas contam. Depois disso, vou pedir que o(a) senhor(a) me conte essa história. (Mostrar as figuras e esperar). Está pronto(a)? Podemos começar?</i>			
ESTRUTURA	S	N	
1. Um garoto vê um cachorro (de rua/perdido) na rua/calçada.	(2)	(1)	(0)
2. O garoto leva (decide levar) o cachorro para casa.	(2)	(1)	(0)
3. O garoto esconde o cachorro no guarda-roupa/roupa/armário.	(2)	(1)	(0)
4. A mãe encontra o cachorro.	(2)	(1)	(0)
5. A mãe pede ao garoto algumas explicações / O garoto implora à mãe para ficar com o cachorro.	(2)	(1)	(0)
6. A mãe permite ao garoto ficar com o cachorro / A mãe ajuda o filho/garoto a construir a casinha para o cachorro.	(2)	(1)	(0)
			TOTAL: ___/12
DESEMPENHO LINGUÍSTICO I	N	S	
Ausência de sequência de narrativa	(2)	(1)	(0)
Predominância de descrição de cenas	(2)	(1)	(0)
Inclusão de informações intrusivas / não existentes na história	(2)	(1)	(0)
Não reconhecimento de personagens	(3)	(2)	(1)
			TOTAL: ___/9
DESEMPENHO LINGUÍSTICO II	N	S	
Dificuldade de busca de palavras	(2)	(1)	(0)
Repetições imediatas de palavras	(2)	(1)	(0)
Revisões	(2)	(1)	(0)
			TOTAL: ___/6

Fonte: HÜBNER et. al, no prelo.

A análise da microestrutura foi realizada a partir das métricas obtidas no software Coh-Metrix Dementia⁴ (ALUÍSIO; CUNHA; SCARTON, 2016). Este software deriva do Coh-Metrix_Port (SCARTON; ALUÍSIO, 2010), que também analisa métricas linguísticas em textos. A diferença é que o Coh-Metrix Dementia surge com a proposta de apontar indícios do aparecimento de declínios cognitivos (primeiramente CCL e DA) a partir da análise da produção de narrativas. Para tanto, o software se vale de ferramentas de Processamento de Línguas Naturais (PLN) e de Aprendizado de Máquina, assim como as técnicas de *parsing*, que auxiliam na análise de componentes linguísticos do texto. Ao todo, são mais de 70 métricas para calcular densidade semântica, complexidade sintática, coerência textual, entre outras (CUNHA, 2015). As métricas selecionadas para análise são: contagem básica de palavras (Incidência de Adjetivos, Incidência de Advérbios, Incidência de Palavras de Conteúdo, Índice Flesch, Incidência de Palavras Funcionais, Incidência de Substantivos, Número de Palavras, Incidência de Pronomes, Incidência de Verbos); operadores lógicos (incidência de Operadores Lógicos); frequência das palavras de conteúdo (Frequência das Palavras de Conteúdo); pronomes, tipos e tokens (Índice Brunet, Estatística de Honoré, Incidência de Pronomes Pessoais, Relação Tipo por Token); Conectivos (Incidência de conectivos); complexidade sintática (Complexidade de Frazier, Complexidade de Yngve); densidade semântica (Densidade de Conteúdo) e disfluências (Número Médio de Disfluências, Número Médio de Emissões Vazias, Duração Média das Pausas, Densidade de Ideias).

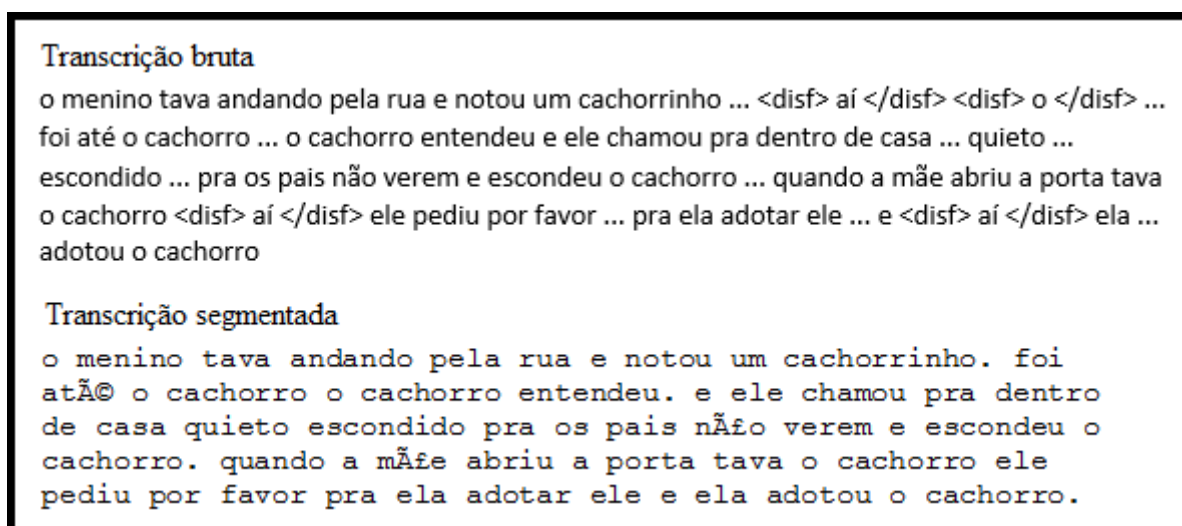
Todos os áudios foram transcritos seguindo as orientações do software e dos colegas parceiros do grupo de pesquisa, autores do software. Indicativos de pausas curtas foram registrados com reticências, pausas mais longas foram informadas como ((pausa XX segundos)), prolongamentos de vogais foram registrados com ::, comentários extra-textuais foram registrados entre <empty> </empty>, e disfluências foram registradas como <disf> </disf>. Em seguida, todas as transcrições foram submetidas à ferramenta DeepBonDD⁵ (*a Deep neural approach to Boundary and Disfluency Detection*), um segmentador automático de sentenças (TREVISIO, 2018). O objetivo de segmentar as sentenças no texto produzido por populações clínicas é o de extrair, automaticamente, as marcas de pausas e disfluências presentes no texto, delimitando os limites das sentenças produzidas. A ferramenta é desenvolvida pelo mesmo grupo responsável pelo Coh-Metrix Dementia, e promete refinar

⁴ Disponível em: <<http://143.107.183.175:22380/>>. Acesso em 21 jun. 2018.

⁵ Disponível em: <<https://mtreviso.github.io/deepbond/index.html>>. Acesso em 12. Dez, 2018.

ainda mais os resultados do extrator de métricas a partir desta segmentação. Além disso, o processo de segmentação automático tem o objetivo de ser uma ferramenta que economiza o tempo de uma segmentação manual, que exigiria atuação de profissional adequado para realizar o trabalho. A Figura 02 apresenta um exemplo do texto transcrito e segmentado.

Figura 2: Modelo de transcrição e segmentação de texto



Fonte: O autor (2019)

Para a extração de métricas do software Coh-Metrix Dementia, são incluídas as duas versões da transcrição: bruta e segmentada. O objetivo dessa alimentação dupla é o de contabilizar as métricas de disfluências, disponíveis apenas na transcrição bruta. Os resultados das métricas são disponibilizados em unidades numéricas, a partir da fórmula aplicada no cálculo de cada uma delas, conforme demonstrado anteriormente. Sendo assim, na subseção a seguir serão apresentados os procedimentos para análise de todos os dados coletados nos estudos propostos nesta dissertação.

3.3.6 Procedimentos para análise de dados

Os dados estatísticos da pesquisa foram computados pelo pesquisador com supervisão estatística. O software utilizado para análise dos dados foi o *Statistical Package for Social Sciences* (SPSS) versão 17.0. As comparações entre grupos foram realizadas a partir de um teste-t paramétrico, com valor de $p < 0,05$ para resultados significativos.

Também foram realizados testes de normalidade Shapiro-Wilk para os dados das métricas de microestrutura, o que resultou em uma análise com teste não-paramétrico Mann-

Whitney para as variáveis de disfluências e de frequência de palavras de conteúdo. Para a análise das correlações entre grupos e variáveis (linguísticas x neuropsicológicas) foram realizados testes de correlação de Pearson, também com valor de $p < 0,05$ para resultados significativos.

Além da análise estatística quantitativa, o desempenho dos grupos nos textos foram explorados de maneira qualitativa, com o objetivo de identificar características e complementar os dados obtidos através da utilização da pontuação macroestrutural e das métricas responsáveis pela análise da microestrutura.

4 APRESENTAÇÃO DOS RESULTADOS

Neste capítulo, os dados serão apresentados observando-se a condução dos dois estudos que compõem a dissertação. Na primeira seção serão disponibilizados os dados referentes aos questionários sociodemográficos e tarefas de funcionalidade. Após, serão apresentados os dados obtidos a partir da administração de tarefas neuropsicológicas. Em seguida, os dados linguísticos, os quais serão subdivididos em análise da macro e da microestrutura textual, seguidos de uma análise também de cunho qualitativo. E, por último, os dados das correlações entre dados neuropsicológicos e linguísticos.

4.1 DADOS SOCIODEMOGRÁFICOS E COGNITIVOS

Nesta seção serão apresentados os dados de caracterização sociodemográfica dos participantes, assim como dados dos testes de avaliação cognitiva. De maneira a melhor organizar os resultados, estes serão divididos por estudos, conforme divisão descrita na seção do método.

Um número maior de participantes havia sido pré-selecionado para compor o estudo. Porém, ao final da revisão dos critérios de inclusão e exclusão, o número total ficou em 76 participantes, considerando que o grupo de idosos saudáveis de baixa escolaridade é o mesmo para os estudos 1 e 2. Destes 76 participantes, foi necessário, ainda, excluir um dos participantes (do grupo SAE) devido à falta de dados linguísticos para análise da microestrutura. Esta exclusão acarretou uma diferença significativa na idade entre os grupos, até então controlada.

4.1.1 Dados sociodemográficos e cognitivos dos participantes do Estudo 1

Os dados de caracterização dos grupos incluídos no Estudo 1 estão divididos entre idade, escolaridade, status socioeconômico, Mini Exame do Estado Mental, Escala de Depressão Geriátrica, hábitos de leitura e hábitos de escrita. A Tabela 1 apresenta os valores de média, desvio padrão, e p (quando significativo na distinção entre grupos).

Tabela 1: Dados sociodemográficos e cognitivos dos participantes do Estudo 1

	SAE (N=28)		SBE (N=35)		<i>p</i>
	M	DP	M	DP	
Idade	69,46	5,15	72,20	5,12	,040
Escolaridade	13,96	3,20	5,29	2,25	,000
Status socioeconômico	25,46	5,65	21,27	4,48	,002
MEEM	27,93	2,15	26,71	2,94	-
Hábitos de leitura atuais	8,93	3,23	4,83	2,94	,000
Hábitos de escrita atuais	5,71	2,38	3,17	3,17	,001

Nota: SAE = Grupo de idosos saudáveis de alta escolaridade, SBE = Grupo de idosos saudáveis de baixa escolaridade, N = número de participantes, M = média, DP = desvio padrão, MEEM = Mini Exame do Estado Mental, GDS = Escala de Depressão Geriátrica.

Fonte: O autor (2019)

A diferença de idade entre os grupos se mostrou significativa ($p = ,040$), sendo a média de idade do grupo SAE ($M = 69,46$) menor que a do grupo SBE ($M = 72,20$). Salientamos que a necessidade de exclusão de um dos participantes, devido ao fato de que este não havia completado a tarefa linguística, acarretou esta diferença, não observada anteriormente. No entanto, salienta-se que a significância nos demais dados não sofreu alteração a partir da exclusão do participante. Como esperado, há diferença significativa ($p = ,000$) no grau de escolaridade entre os grupos. Há ainda diferenças significativas entre as variáveis de hábitos de leitura ($p = ,000$) e escrita ($p = ,001$), indicando que os indivíduos mais escolarizados leem e escrevem mais do que os menos escolarizados. Os dados do MEEM apontaram não haver declínios cognitivos. Houve diferença significativa no padrão socioeconômico ($p = ,002$), indicando associação entre grau de escolaridade mais elevado e padrão socioeconômico também mais elevado, isto é, a média do grupo SAE caracteriza-o sob a classe B2, enquanto que a média do grupo SBE caracteriza-o sob o nível C1, de acordo com a ABEP (2014). Essa diferença significativa pode estar atrelada ao grau de escolaridade, fazendo com que pessoas mais escolarizadas tenham mais acesso a bens culturais, o que pode impactar no desempenho desse grupo na tarefa linguística de produção de narrativa.

Após a apresentação dos dados sociodemográficos de caracterização da amostra de participantes incluída no Estudo 1, apresentaremos, na subseção a seguir, os mesmos dados referentes aos participantes do Estudo 2.

4.1.2 Dados sociodemográficos e cognitivos dos participantes do Estudo 2

Os dados de caracterização dos grupos incluídos no Estudo 2 estão divididos entre idade, escolaridade, status socioeconômico, Mini Exame do Estado Mental, Escala de Depressão Geriátrica, hábitos de leitura e hábitos de escrita. A Tabela 2 apresenta os valores de média, desvio padrão, e *p* (quando significativo na distinção entre grupos).

Tabela 2: Dados sociodemográficos e cognitivos dos participantes do Estudo 2

	DA (N=12)		GC (N=35)		<i>p</i>
	M	DP	M	DP	
Idade	74,58	5,83	72,20	5,12	-
Escolaridade	5,00	1,53	5,29	2,25	-
Status socioeconômico	20,29	5,21	21,27	4,48	-
MEEM	20,75	3,22	26,71	2,94	,000
Hábitos de leitura atuais	3,91	4,43	4,83	2,94	-
Hábitos de escrita atuais	1,50	1,90	3,17	3,17	-

Nota: DA = Grupo de idosos acometidos pela Doença de Alzheimer, GC = Grupo controle, N = número de participantes, M = média, DP = desvio padrão, MEEM = Mini Exame do Estado Mental, GDS = Escala de Depressão Geriátrica.

Fonte: O autor (2019)

As diferenças de idade e escolaridade entre os grupos não se mostraram significativas, indicando homogeneidade nos grupos. Também não houve diferenças significativas entre as variáveis de hábitos de leitura e escrita, cujos escores apontam para baixa frequência para ambas as atividades; ainda que, em média, os participantes do grupo controle leiam e escrevam mais do que os participantes do grupo DA. Conforme esperado, no MEEM a diferença entre grupos foi significativa ($p = ,000$), indicando comprometimentos cognitivos

para o grupo DA. Não houve diferença significativa no padrão socioeconômico. As médias classificam os grupos como C1, de acordo com subdivisão da ABEP (2014). Este dado pode indicar uma relação maior do padrão socioeconômico com o grau de escolaridade, como também observada na comparação entre os grupos descritos no Estudo 1, ou ainda com os hábitos de leitura e escrita, isto é, pessoas mais escolarizadas parecem ter maior padrão socioeconômico e hábitos de leitura e escrita mais frequentes.

Concluída a apresentação dos dados de caracterização de ambos os estudos abordados nesta dissertação, introduzimos na seção seguinte a apresentação dos dados obtidos através da administração das tarefas neuropsicológicas adotadas nesta pesquisa.

4.2 DADOS NEUROPSICOLÓGICOS DOS PARTICIPANTES

Os dados neuropsicológicos foram coletados a partir de tarefas elaboradas para investigar dois construtos de memória: episódica e de trabalho, conforme apresentado na seção do método. Novamente apresentamos os dados de acordo com a divisão entre os estudos.

4.2.1 Dados neuropsicológicos dos participantes do Estudo 1

A Tabela 3 apresenta os valores de média, desvio padrão, e *p* (quando significativo na distinção entre grupos) para as variáveis neuropsicológicas de memória episódica (recordação sem pista, recordação com pista e recordação total) e memória de trabalho (digit span – ordem direta, digit span – ordem indireta, digit span total) para os participantes descritos no Estudo 1.

Tabela 3: Dados neuropsicológicos dos participantes do Estudo 1

	SAE (N=28)		SBE (N=33*/35)		<i>p</i>
	M	DP	M	DP	
ME – Recordação sem pista	31,18	5,85	30,45	8,37	-

	SAE (N=28)		SBE (N=33*/35)		p
	M	DP	M	DP	
ME – Recordação com pista	16,25	5,44	16,00	6,18	-
ME – Recordação total	47,50	,839	47,21	3,79	-
MT – Digit Span – Ordem direta	8,11	2,14	6,69	1,47	,005
MT – Digit Span – Ordem indireta	4,04	1,64	3,40	1,53	-
MT – Digit Span Total	12,14	3,33	10,09	2,47	,007

Nota: SAE = Grupo de idosos saudáveis de alta escolaridade, SBE = Grupo de idosos saudáveis de baixa escolaridade, N = número de participantes, M = média, DP = desvio padrão, ME = memória episódica, MT = memória de trabalho.

* Para os dados da tarefa de ME, o grupo SBE tem N = 33, pois a tarefa não foi aplicada em dois participantes.

Fonte: O autor (2019)

Os resultados para os testes de ME não apresentaram diferenças significativas entre os grupos para nenhuma das etapas da tarefa. Isto pode indicar que a preservação da ME independe do grau de escolaridade. Por outro lado, os dados relacionados com o construto de MT apontaram diferença significativa no teste de span de dígitos em ordem direta ($p = ,005$) e no total da soma dos scores das ordens direta e indireta ($p = ,007$), evidenciando melhor desempenho para o grupo mais escolarizado. Esses resultados podem indicar que há uma relação positiva entre o grau de escolaridade e a atenção, isto é, quanto maior a escolaridade, melhor o desempenho deste construto, visto que, conforme discutido nos pressupostos teóricos, a ordem direta do teste de dígitos está mais relacionada a esta função executiva. O foco principal desta análise está na avaliação de MT, que está associada à versão indireta do teste, cuja diferença não se mostrou significativa entre os grupos, ainda que o grupo mais escolarizado tenha obtido escores superiores em relação ao grupo menos escolarizado para este construto.

Seguimos, na subseção a seguir, com a apresentação dos dados relativos às tarefas neuropsicológicas aplicadas aos participantes do Estudo 2.

4.2.2 Dados neuropsicológicos dos participantes do Estudo 2

A Tabela 4 apresenta os valores de média, desvio padrão, e *p* (quando significativo na distinção entre grupos) para as variáveis neuropsicológicas de memória episódica (recordação sem pista, recordação com pista e recordação total) e memória de trabalho (digit span – ordem direta, digit span – ordem indireta, digit span total) para os participantes descritos no Estudo 2.

Tabela 4: Dados neuropsicológicos dos participantes do Estudo 2

	DA (N=12)		GC (N=33*/35)		<i>p</i>
	M	DP	M	DP	
ME – Recordação sem pista	11,25	12,54	30,45	8,37	,000
ME – Recordação com pista	14,83	10,58	16,00	6,18	-
ME – Recordação total	26,08	17,50	47,21	3,79	,002
MT – Digit Span (Ordem direta)	6,55	2,33	6,69	1,47	-
MT – Digit Span (Ordem indireta)	2,64	1,28	3,40	1,53	-
MT – Digit Span Total	9,18	3,25	10,09	2,47	-

Nota: DA = Grupo de idosos acometidos pela Doença de Alzheimer, GC = Grupo controle, N = número de participantes, M = média, DP = desvio padrão, ME = memória episódica, MT = memória de trabalho.

* Para os dados da tarefa de ME, o grupo SBE tem N = 33, pois a tarefa não foi aplicada em dois participantes.

Fonte: O autor (2019)

Diferentemente dos resultados obtidos para a comparação entre grupos proposta no Estudo 1, observa-se uma diferença significativa entre grupos na tarefa de ME, mais especificamente na recordação de palavras sem pista semântica ($p = ,000$) e no total de recordação das palavras ($p = ,002$). Esta diferença evidencia que o grupo DA tem piores

escores que o GC na avaliação da ME, apontando para declínio deste construto cognitivo em quadros demenciais do tipo Alzheimer. Não foram observadas diferenças significativas entre os grupos para nenhuma das etapas da tarefa de ordenamento de dígitos, ainda que o GC tenha apresentado médias superiores ao grupo DA nas três análises realizadas. Estes dados parecem sugerir que as funções de MT permanecem inalteradas em estágios iniciais da DA, ou pelo menos não sofreram influência do quadro demencial no momento da testagem. Ainda é possível associar a falta de diferença significativa ao baixo número de participantes que compõem o grupo DA.

Após a apresentação de dados relativos à administração de testes neuropsicológicos, apresentaremos, na seção a seguir, os dados oriundos da tarefa linguística. Assim como nesta seção, os dados serão separados por estudo, sendo que, primeiramente, são apresentados os dados de macro e microestrutura para a comparação de grupos proposta no Estudo 1, e, em seguida, os dados de macro e microestrutura para a comparação de grupos proposta no Estudo 2.

4.3 DADOS LINGUÍSTICOS DA TAREFA DE NARRATIVA

Os dados linguísticos apresentados nesta seção foram coletados a partir da administração de uma tarefa de produção oral de narrativa a partir de sequência de figuras. Cada subseção tratará de um dos estudos descritos na dissertação, iniciando pela apresentação dos dados da macroestrutura e seguidas dos dados da microestrutura. Além dos resultados de análise estatística (quantitativos), exemplos de transcrições de textos orais produzidos serão utilizados como ilustração para uma análise mais profunda do processo de produção textual (qualitativa).

4.3.1 Dados linguísticos dos participantes do Estudo 1

Os dados de macroproposições responsáveis pela elaboração da macroestrutura da narrativa (Total Macro), assim como de desempenho linguístico, são apresentados na Tabela 5, que inclui os valores de média, desvio padrão, e *p* (quando significativo na distinção entre grupos) para os grupos que compõem o Estudo 1.

Tabela 5: Dados de macroestrutura e desempenho linguístico dos participantes do Estudo 1

	SAE (N=28)		SBE (N=35)		p
	M	DP	M	DP	
Total Macro	8,86	2,94	5,86	3,90	,001
Total DLI	7,82	2,01	6,20	2,99	,013
Total DLII	5,39	1,47	4,83	1,90	-

Nota: SAE = Grupo de idosos saudáveis de alta escolaridade, SBE = Grupo de idosos saudáveis de baixa escolaridade, N = número de participantes, M = média, DP = desvio padrão, DLI = desempenho linguístico I, DLII = desempenho linguístico II.

Fonte: O autor (2019)

Os resultados apontam diferença significativa ($p = ,001$) entre os grupos na variável Total Macro, sendo que o grupo SAE obteve melhores escores que o grupo SBE. Esta diferença permite concluir que o grau de escolaridade pode estar associado com a produção do nível macroestrutural do texto, isto é, pessoas mais escolarizadas parecem perceber e construir de maneira mais eficaz as ideias principais do texto, pelo menos em tarefas de produção textual com base em sequência de figuras. Dados das transcrições também permitem observar esta diferença. As figuras (3 e 4) a seguir ilustram diferenças significativas entre os textos de pessoas saudáveis de alta e baixa escolaridade.

Figura 3: Exemplo de transcrição do grupo SAE

C08_22414_1600

Essa é a história do Joãozinho ... ou melhor do Pedrinho ... <empty> é o nome do meu outro filho </empty> ... Pedrinho tava na rua voltando do colégio ... quando bateu o olho num cachorrinho ... e o cachorrinho tava muito assustado porque:: ele tava abandonado e ninguém dava bola pra ele ... <disf> aí </disf> o Pedrinho olhou e chamou ele e o cachorrinho já começou a balançar o rabo ... e disse "olha eu vou te levar pra casa mas com uma condição ... tu entra e fica quietinho porque lá em casa não é um... pode ser que nós não seja muito bem-vindo" ... e o cachorro como todo bom cachorro entendeu ... aí ele colocou o cachorro no seu quarto dentro do seu guarda roupa ... e o cachorro ficou bem acomodadinho ali no quentinho ... a essas alturas ele já tinha trazido uma comidinha e o cachorro tava bem feliz ... até que chega a hora da mamãe arrumar o quarto e dá de cara com o cachorro ... "que é isso menino?" ... "mãe por favor é um bichinho bom tava faminto tava necessitado vamos atendê-lo vamos socorrê-lo" ... a mãe achou meio ruim meio ruim mas... como todo jeito concordou ... e passaram no pátio a fazer uma casinha pro cachorro que é um lugar mais adequado do que dentro do roupeiro pra um bichinho desses morar

Fonte: O autor (2019)

Como é possível observar na Figura 3, todas as macroproposições esperadas na pontuação da tarefa são produzidas. Primeiramente, é possível perceber que o participante reconhece os três personagens (o cachorro, o menino e a mãe) e identifica a história contada a partir da sequência de figuras, relacionando os personagens nas diferentes ações performadas ao longo da história. Nota-se, ainda, que todas as etapas de construção da narrativa, conforme propostas por Adam (2008), são respeitadas, sendo elas o Estágio Inicial (Pn1): o menino encontra o cachorro na rua; o Nó Desencadeador (Pn2): o menino leva o cachorro para casa e esconde-o no roupeiro; Re-ação ou Avaliação (Pn3): a mãe encontra o animal e pede explicações; Desenlace (Pn4): o menino implora para ficar com o cachorro e a mãe concorda; Situação Final (Pn5): a mãe auxilia o menino na construção de uma casinha para o cachorro.

Além dos aspectos mais estruturais da narrativa, é possível perceber também características de quem conta uma história, como a presença de um narrador em terceira pessoa, a nomeação do personagem (Pedrinho), e as falas em primeira pessoa do menino, como, por exemplo, quando ele implora para a mãe para que permita que ele fique com o cachorrinho. Ambos os aspectos destacados nesta construção narrativa retomam as noções de Van Dijk (1983), que compreende o processo de construção de texto narrativo como dividido em dois aspectos linguísticos principais: o semântico e o pragmático. O aspecto semântico, no exemplo da Figura 3, está mais relacionado à estrutura da narrativa e construção linguística em si, enquanto que o aspecto pragmático se relaciona mais com essa modulação do discurso para se encaixar à proposta de contar uma história. A Figura 4, por outro lado, apresenta um exemplo de texto produzido por idoso saudável de baixa escolaridade.

Figura 4: Exemplo de transcrição do grupo SBE

A54_101115_1500

num cantinho ... dessa paisagem existia um ... criança e um cachorrinho <disf> né </disf>
brincando conversando e ele tentando dar comidinha pra ele ... lá estava o outro chamando
para que voltasse pra dentro de casa <empty> é casa isso aqui? </empty> voltasse pra dentro
de casa ... o bichinho estava entretido com o outro não estava entendendo ... lá atrás o outro
latiu pra chamar atenção <disf> del </disf> do outro ... chamou o de trás ... se deu conta e
olhou ... e o outro continuou como estava ... <disf> ai </disf> ficou tudo em silêncio <disf> seu
</disf> sua família estava esperando os bichinho para dar os seus <empty> ah não sei
</empty> os seus alimentos ... até que () ... com o tempo nenhum se aproximou-se deles

Fonte: O autor (2019)

Neste exemplo, diferentemente do anterior, não há identificação de todos os personagens da história. São identificados o cachorro e uma criança (não há especificação sobre o sexo do personagem), mas a mãe, não. Além disso, a participante utiliza a expressão “outro” para se referir a personagens da história, o que dificulta a identificação deles. Também não há respeito à construção da estrutura narrativa, estando deficitárias as delimitações de macroproposições, conforme sugeridas por Adam (2008). Nota-se, também, um maior número de comentários extra-textuais e revisões, como *seu sua família, del do outro*. Neste sentido, retomamos a análise dos aspectos linguísticos conforme apresentadas na Tabela 5.

Também se observou diferença significativa no Total de Desempenho Linguístico I ($p = ,013$), em que o grupo SAE obteve maiores escores que o grupo SBE. Dentre as características de desempenho linguístico encontram-se a ausência de sequência narrativa, a predominância da descrição de cenas, a inclusão de informações intrusivas ou não existentes na história e o não reconhecimento de personagens. A comparação entre os exemplos apresentados nas figuras 3 e 4 permite constatar essa diferença qualitativa entre os textos.

Por fim, não foi observada diferença significativa entre os grupos para o Total de Desempenho Linguístico II. Ainda que o grupo SAE tenha média superior ao grupo SBE, não parecem ser características do texto de pessoas menos escolarizadas a dificuldade na busca de palavras, as repetições imediatas de palavras e as revisões, ainda que possam ocorrer, como observado no exemplo da Figura 4.

Apresentamos, a seguir, os dados relativos à microestrutura dos textos. A Tabela 6 organiza os dados obtidos através da extração automática de métricas linguísticas com apoio do software Coh-Metrix Dementia (ALUÍSIO; CUNHA; SCARTON, 2016). São apresentados os valores de média, desvio padrão e p (quando significativo na distinção entre grupos) para a comparação entre grupos SAE e SBE.

Tabela 6: Dados de microestrutura dos participantes do Estudo 1

	SAE		SBE		p
	(N=28)		(N=35)		
	M	DP	M	DP	
Incidência de adjetivos	36,85	18,00	55,32	27,62	,003
Incidência de advérbios	64,76	30,45	68,48	31,81	-

	SAE		SBE		<i>p</i>
	(N=28)		(N=35)		
	M	DP	M	DP	
Incidência de palavras de conteúdo	613,47	43,35	621,71	55,29	-
Índice Flesch	82,24	10,59	84,29	8,90	-
Incidência de palavras funcionais	380,86	48,98	368,30	58,39	-
Incidência de substantivos	296,15	49,71	296,06	43,00	-
Número de palavras	144,89	65,66	152,91	74,08	-
Incidência de pronomes	84,55	25,64	90,61	30,56	-
Incidência de verbos	215,70	29,01	201,83	32,89	-
Incidência de operadores lógicos	54,63	18,89	56,96	20,58	-
Índice Brunet	11,71	1,73	12,00	1,48	-
Estatística de Honoré	607,11	95,21	599,05	111,95	-
Relação tipo por token	,498	,109	,466	,098	-
Incidência de pronomes pessoais	50,24	20,46	48,65	24,02	-
Incidência de conectivos	97,61	27,53	87,81	28,20	-
Complexidade de Frazier	7,22	,589	6,99	,651	-
Complexidade de Yngve	2,30	,278	2,37	,352	-
Densidade de conteúdo	1,65	,349	1,75	,441	-
Densidade total	2,71	,753	2,70	,784	-

	SAE (N=28)		SBE (N=35)		<i>p</i>
	M	DP	M	DP	
de ideias					
Frequência das palavras de conteúdo	219854,42	161179,97	181319,80	124787,92	-
Número médio de disfluências	,064	,041	,083	,051	-
Número médio de emissões vazias	,070	,097	,126	,121	,047
Média de duração das pausas	,025	,076	,059	,121	,039

Nota: SAE = Grupo de idosos saudáveis de alta escolaridade, SBE = Grupo de idosos saudáveis de baixa escolaridade, N = número de participantes, M = média, DP = desvio padrão.

Fonte: O autor (2019)

A comparação dos dados entre os grupos não demonstrou diferenças significativas para a maioria das métricas, ainda que, na grande maioria dos casos, as médias do grupo SAE tenham sido superiores às do grupo SBE. Houve diferença significativa no uso de adjetivos ($p = ,003$), indicando incidência mais frequente de adjetivos no texto de pessoas menos escolarizadas na comparação com mais escolarizadas. Além disso, houve também diferença significativa entre o número de emissões vazias, isto é, comentários que não compõem a história ($p = ,047$), e duração de pausas ($p = ,039$), com maior incidência no grupo menos escolarizado. Os resultados para as métricas de disfluências parecem evidenciar que os textos produzidos por pessoas menos escolarizadas são, em geral, mais pausados e mais interrompidos pela adição de informações extratextuais do que o texto de pessoas mais escolarizadas. Os exemplos a seguir permitem observar esta tendência.

Figura 5: Exemplo de transcrição do grupo SBE (pausas e emissões vazias)

V04_100118_1500

o Alexandre tava indo na rua pra ir na casa de alguém <disf> né </disf> ... na época ainda tava frio ... ele tá com cachecol ... <disf> e </disf> <disf> ai </disf> ... chamou o cachorrinho pra fazer um agrado ... <disf> aqui </disf> ... <empty> que que ele ta fazendo? </empty> ... ele ta entrando ... ele tem uma pasta ... <disf> aí </disf> <disf> já </disf> tem uma pasta ... <empty> aqui não tem </empty> ... ou tem ali naquela mão ... <empty> e eu não notei </empty> ... ((pausa 06 segundos)) ele tá indo pra escola ... tá com a pasta ali e o cachorrinho seguiu ele ... <disf> aqui </disf> ... <empty> ou </empty> ele mandou o cachorrinho sair ... <empty> ou </empty> ele deu alguma coisa pro cachorro comer ... ((pausa 04 segundos)) em casa depois ele trouxe o cachorro e guardou ... e a mãe dele abriu o armário e achou o cachorro ali <empty> isso me aconteceu com um gato quando eu era pequena ... consegui ter o gato três dias dentro de casa ... meu pai não gostava de gato ... cachorro podia pedir trinta ... ele dava e não se importava ... mas gato ele não gostava ... eu consegui ter o gato três dias dentro de casa ... depois um dia o diabo do gato miou ... ((risos)) ... ai ele me pegou pela orelha assim ... e disse ... "a senhora vai devolver esse bicho ... lá de onde que a senhora ganhou" </empty> <disf> ai </disf> ele tá implorando pra mãe dele pra ficar com o cachorro... porque ele gostou do cachorrinho <disf> né </disf> ... lá já tá arrumando uma casa ali pro cachorro com o amigo dele ((pausa 04 segundos)) é:: ... é o amigo dele aquele ali <disf> né? </disf> ... tem um serrote ... um alicate ... ((pausa 03 segundos)) e fizeram uma casa pro cachorro ... ((pausa 03 segundos)) <empty> é isso ai </empty>

Fonte: O autor (2019)

O exemplo apresentado na Figura 5 foi produzido por uma participante do grupo SBE. É possível identificar um texto extenso, aparentemente sem grandes problemas de construção microestrutural, ao mesmo tempo que as relações intrafrasais são preservadas, não sendo observadas falhas de coerência local. As sentenças parecem ser sintaticamente bem construídas, ainda que sejam consideradas sentenças simples. Entretanto, notam-se comentários não pertinentes à construção da narrativa baseada em figuras, como, por exemplo, uma lembrança de cunho pessoal (relacionada ao gato que a participante teve na infância) reavivada a partir da resolução da tarefa. Além da intrusão de informações, é possível notar um grande número de pausas longas, totalizando aproximadamente 20 segundos de pausas ao longo da produção textual. Esta tendência não parece ser observada nos textos produzidos por participantes do grupo mais escolarizado, como podemos observar no exemplo apresentado na Figura 06.

Figura 6: Exemplo de transcrição do grupo SAE (microestrutura)

E14_031115_1424

então a história conta na verdade que <disf> uma </disf> um menino estava caminhando pela rua e ele viu um cachorro ... e o cachorro pareceu a ele um cachorro carente ... ele chama o cachorro fazendo um gesto ... e vai indo pra casa ... só que ele pede silêncio para o cachorro e sai pé ante pé ... pra esconder dos pais ... ele coloca dentro do roupeiro ... depois ele tira ... a mãe vai abrir o roupeiro e encontra o cachorro ... e ele pede pelo amor de Deus fazendo um gesto ... que a mãe aceitasse o cachorro ... e o bicho do lado dele ... até que ele encontra a mãe ... <disf> que </disf> <disf> tava </disf> ... que fez uma casinha pro cachorro ... <empty> e esse é o final feliz </empty>

Fonte: O autor (2019)

Observa-se, neste exemplo, uma construção mais fluida, sem a presença de pausas longas e sem a inclusão de informações não pertinentes à construção da narrativa. Mais especificamente, no nível de microestrutura, é possível observar sentenças mais elaboradas, com utilização da conjunção “que”, por exemplo. A análise da comparação entre ambos os textos permite levantar a hipótese de que não há grandes diferenças no nível lexical na construção da microestrutura, ainda que, sintaticamente, o texto de pessoas mais escolarizadas parece ser melhor elaborado que o texto de pessoas menos escolarizadas, a julgar pela média superior na métrica de Complexidade de Frazier, que está relacionada à complexidade sintática do texto. É necessário, porém, uma nova avaliação, mais conclusiva, para atestar tal afirmativa.

Observadas as diferenças linguísticas no nível microestrutural entre os grupos, serão apresentados, na subseção seguinte, os dados linguísticos de micro e macroestrutura para o Estudo 2, que compreende a comparação de produção oral de narrativa entre os grupos de idosos saudáveis de baixa escolaridade e acometidos pela DA, também de baixa escolaridade.

4.3.2 Dados linguísticos dos participantes do Estudo 2

Os dados de proposições (Total Macro), assim como de desempenho linguístico, são apresentados na Tabela 7, que inclui os valores de média, desvio padrão, e p (quando significativo na distinção entre grupos) para os grupos que compõem o Estudo 2.

Tabela 7: Dados de macroestrutura e desempenho linguístico dos participantes do Estudo 2

	DA (N=12)		GC (N=35)		P
	M	DP	M	DP	
Total Macro	4,08	4,48	5,86	3,90	-
Total DLI	4,50	2,87	6,20	2,99	-
Total DLII	3,58	2,31	4,83	1,90	-

Nota: DA = Grupo de idosos acometidos pela Doença de Alzheimer, GC = Grupo controle, N = número de participantes, M = média, DP = desvio padrão, DLI = desempenho linguístico I, DLII = desempenho linguístico II.

Fonte: O autor (2019)

Não foram observadas diferenças significativas entre os grupos para o total da construção do nível de macroestrutura nem para desempenho linguístico. O GC obteve melhores médias do que o grupo DA em todas as variáveis, porém sem alcançar diferença significativa. Acredita-se que a ausência de significância estatística possa estar relacionada ao número de participantes reduzidos no grupo DA. Observa-se, porém, na análise qualitativa dos textos produzidos por pessoas acometidas pela DA, características descritas na avaliação do desempenho linguístico (ausência de sequência narrativa, não reconhecimento de personagens, etc), conforme exemplos.

Figura 7: Exemplo de transcrição do grupo DA (macroestrutura)

B38_271016_1530

chamei o cachorro ... o cachorro veio ... ((pausa 04 segundos)) parou ali ... ((pausa 06 segundos)) <disf> ah </disf> <empty> me pegou hein </empty> ... ((pausa 03 segundos)) a mulher abriu a porta ... e pegou <empty> não sei o que não sei o que </empty> ... ((pausa 05 segundos)) <empty> não sei minha filha </empty>

Fonte: O autor (2019)

A Figura 7 apresenta um exemplo de narrativa produzida por participante do grupo DA. Nota-se a identificação de personagens (cachorro e mulher), ainda que não de maneira específica. O menino é tratado em primeira pessoa, como se o participante fizesse parte da história, como visto na sentença *chamei o cachorro*. O esquema de Adam (2008) não é respeitado, faltando informações necessárias para a transformação de Pn1 em Pn5,

característica mais importante da construção narrativa, de acordo com o autor. Nota-se, além disso, presença frequente de pausas longas, totalizando 18 segundos de pausa. Além disso, percebemos, no exemplo da Figura 8, outras características apontadas na avaliação do desempenho linguístico.

Figura 8: Exemplo de transcrição do grupo DA (descrição de cena)

A01_250414_1550

um gurizinho ... outro gurizinho chamando o cachorrinho ... <disf> é </disf> ... a guriazinha ... a senhora abrindo a porta ... a outra conversando com o outro gurizinho ... e o cachorrinho ... o cachorrinho tá aqui ... pegando o gurizinho querendo pegar ou brincar <empty> não sei </empty> ... aqui é outro gurizinho ... o cachorrinho <disf> né? </disf> ... isso é uma casinha do cachorro ... e o gurizinho brincando ali <disf> ó </disf> ... em cima da casinha do cachorro ... <empty> foi o que eu entendi </empty>

Fonte: O autor (2019)

No texto da Figura 8 não se identificam personagens, visto que há predominância na descrição de cada figura em detrimento da construção de uma sequência narrativa, ilustrada na utilização de palavras como *aqui*, *ali*, *outro*, para se referir aos quadrinhos e aos personagens. Não há relação de sentido entre uma figura ou outra, ou até mesmo relação entre o menino que aparece em uma figura e na outra. A macroestrutura da narrativa não foi construída, pelo fato de as macroproposições, isto é, as ações de cada estágio da produção narrativa, não terem sido sequer mencionadas. Possíveis justificativas para a não-significância na comparação entre a construção de macroestrutura no grupo DA e SBE podem ser atribuídas a dificuldades na identificação das figuras por parte dos próprios participantes do grupo SBE.

Além da construção da macroestrutura das narrativas, apresentamos, a seguir, os dados relativos à microestrutura dos textos produzidos pelos participantes dos grupos analisados no Estudo 2. A Tabela 8 organiza os dados obtidos através da extração automática de métricas linguísticas com apoio do software Coh-Metrix Dementia (ALUÍSIO; CUNHA; SCARTON, 2016), assim como apresentados para o Estudo 1. São incluídos os valores de média, desvio padrão e p (quando significativo na distinção entre grupos).

Tabela 8: Dados de microestrutura dos participantes do Estudo 2

	DA (N=12)		GC (N=35)		P
	M	DP	M	DP	
Incidência de adjetivos	62,57	46,70	55,32	27,62	-
Incidência de advérbios	98,12	45,84	68,48	31,81	,017
Incidência de palavras de conteúdo	692,31	50,32	621,71	55,29	,000
Índice Flesch	82,78	19,54	84,29	8,90	-
Incidência de palavras funcionais	279,28	80,25	368,30	58,39	,000
Incidência de substantivos	317,68	48,94	296,06	43,00	-
Número de palavras	108	57,29	152,91	74,08	-
Incidência de pronomes	86	37,76	90,61	30,56	-
Incidência de verbos	213,92	36,59	201,83	32,89	-
Incidência de operadores lógicos	47,49	24,12	56,96	20,58	-
Índice Brunet	11,25	2,47	12,00	1,48	-
Estatística de Honoré	563,78	215,87	599,05	111,95	-
Relação tipo por token	,489	,156	,466	,098	-
Incidência de pronomes pessoais	31,05	29,82	48,65	24,02	,045
Incidência de conectivos	61,57	24,48	87,81	28,20	,006
Complexidade de Frazier	6,76	1,06	6,99	,651	-

	DA (N=12)		GC (N=35)		P
	M	DP	M	DP	
Complexidade de Yngve	2,49	,735	2,37	,352	-
Densidade de conteúdo	3,41	3,82	1,75	,441	-
Densidade total de ideias	3,08	,793	2,70	,784	-
Frequência das palavras de conteúdo	228172,33	134430,39	181319,80	124787,92	-
Número médio de disfluências	,086	,068	,083	,051	-
Número médio de emissões vazias	,169	,133	,126	,121	-
Média de duração das pausas	,257	,347	,059	,121	,005

Nota: DA = Grupo de idosos acometidos pela Doença de Alzheimer, GC = Grupo controle, N = número de participantes, M = média, DP = desvio padrão.

Fonte: O autor (2019)

Como é possível identificar a partir da análise da tabela, as métricas que mostraram diferença significativa entre grupos são as de incidência de advérbios ($p=,017$), incidência de palavras de conteúdo ($p=,000$), incidência de palavras funcionais ($p=,000$), incidência de pronomes pessoais ($p=,045$), incidência de conectivos ($p=,006$), e média de duração das pausas ($p=,005$). A análise desses dados requer atenção. Na métrica de incidência de advérbio, por exemplo, o grupo DA teve média superior ao grupo GC, o que indica que advérbios foram mais incidentes no texto de pessoas com DA do que de pessoas saudáveis. O mesmo se repetiu para a métrica de incidência de palavras de conteúdo (verbos, adjetivos, substantivos e advérbios). Este dado, porém, não necessariamente indica que o texto produzido por pessoas com DA é mais elaborado que o texto produzido por pessoas saudáveis, visto que há vasta produção de palavras de conteúdo, mas nem sempre elas estão inseridas em sentenças gramaticalmente bem construídas. A Figura 9 traz um exemplo de texto produzido por participante diagnosticado com DA que reforça esta análise.

Figura 9: Exemplo de transcrição do grupo DA (palavras de conteúdo)

B35_300916_1400

casa ... ((pausa 05 segundos)) homem mulher ... criança menina menino cachorrinho ... ((pausa 07 minutos)) menina cachorrinho cachorrinho menino ... ((pausa 04 segundos)) homi ... moça ... moça ... menino ... cachorrinho ... menino ... cachorrinho ... ((pausa 05 segundos)) menino ... ((pausa 07 segundos)) e aparelho de serrá

Fonte: o autor (2019)

Na Figura 9 é possível notar um texto constituído basicamente de palavras de conteúdo, não sendo incluídas palavras funcionais como preposições, conjunções, artigos, etc. O texto é composto de mais de 20 palavras, mas nenhuma delas se relaciona na composição de uma sentença gramaticalmente possível, evidenciando falha na relação de coerência local. Esta característica não se nota nos dados obtidos através da análise dos textos produzidos por pessoas saudáveis, visto que na comparação entre grupos para as métricas de incidência de palavras funcionais (como preposições, artigos, pronomes, entre outros) e incidência de conectivos, o grupo saudável teve maiores médias, diferenciando-se significativamente do grupo DA. Ainda, mesmo que a diferença não tenha se mostrado significativa, o grupo saudável também obteve melhores médias nas métricas incidência de operadores lógicos e complexidade de Frazier (complexidade sintática), o que reforça a hipótese de que o nível microestrutural pode sofrer influência do quadro demencial. Além disso, houve diferença significativa na quantidade de pausas longas, havendo vantagens para o grupo controle nesta métrica.

Figura 10: Exemplo de transcrição do grupo DA (pausas)

A11_160514_1450

um menino com o cachorrinho ... ((pausa 12 segundos)) ele:: ... tá dando <disf> aqui </disf> ele tá dando ... alguma coisa pro cachorro comer ... ((pausa 04 segundos)) <disf> né </disf> ... tá dando algo <disf> aqui </disf> pro cachorro ... ((pausa 17 segundos)) <disf> aqui </disf> ele tá dando algo pro cachorro <disf> né </disf> ... <disf> e </disf> <disf> aqui </disf> tem uma moça ... abrindo uma porta <empty> tipo </empty> dum roupeiro <empty> sei lá </empty> ... um armário ... tá o:: guri ... com as mãos postas pra cima como se fosse rezando ... pra uma senhora que <empty> tá ali </empty> tem o cachorro do lado ... <empty> ali tem .</empty> .. o guri o cachorro o guri também ... <empty> que eu possa enxergar né </empty>

Fonte: O autor (2019)

Em outros exemplos apresentados até então, assim como o ilustrado na Figura 10, a presença de pausas longas é recorrente no discurso de DA. Neste último exemplo, vemos um total de 33 segundos de pausa, além de marcas de disfluências, prolongamento de vogais e emissões vazias. Estes dados permitem concluir que a construção do nível de microestrutura na DA parece afetada, principalmente na elaboração de sentenças mais complexas, mas também na organização das ideias. De forma geral, ainda que estatisticamente os dados não tenham delimitado diferenças significativas, tanto a macro quanto a microestrutura parecem afetadas no quadro demencial do Alzheimer. Buscando compreender quais fatores neuropsicológicos podem influenciar nessas deficiências, apresentamos, na subseção a seguir, os dados de correlações para os grupos analisados.

4.3.3 Dados das correlações de tarefas neuropsicológicas e linguísticas

Nesta subseção serão apresentados os dados das correlações de tarefas neuropsicológicas e linguísticas. Diferentemente da divisão por estudos elaborada nas subseções anteriores, os dados aqui analisados serão disponibilizados por grupos. Sendo assim, iniciaremos pelo grupo SAE, seguido do grupo SBE (que é GC para o Estudo 2), e finalizando com as informações referentes ao grupo DA.

A Tabela 9 apresenta os resultados da correlação de dados neuropsicológicos com o total de macroestrutura e desempenho linguístico I e II para o grupo SAE. São incluídos os valores de média, desvio padrão, e *p* (quando significativo para correlação positiva ou negativa).

Tabela 9: Resultados de correlação de dados neuropsicológicos x macroestrutura

	SAE			SBE			DA		
	(N = 28)			(N = 35)			(N = 12)		
	N	CP	<i>p</i>	N	CP	<i>p</i>	N	CP	<i>P</i>
ME (TG) x TM	28	,210	-	33	-,146	-	12	,604	,038
ME (RSP) x TM	28	,206	-	33	-,154	-	12	,802	,002
MT (DS-OI) x TM:	28	,446	,018	35	,103	-	11	,118	-

Nota: SAE = Grupo saudável de alta escolaridade, N = número de participantes, M = média, DP = desvio padrão, CP = correlação de Pearson, TM = Total de Macro, ME (TG) = Memória Episódica (Total Geral), ME (RSP) = Memória Episódica (Recordação sem pista), MT (DS-OI) = Memória de Trabalho (Digit Span – ordem indireta).
Fonte: O autor (2019)

Para o teste de correlação de Pearson, quanto mais próximo de 1 (ou -1) for o valor, mais forte é o grau de relação entre as variáveis. Nota-se, para o grupo saudável de alta escolaridade, apenas uma correlação positiva entre a ordem indireta do teste de ordenamento de dígitos (que avalia o construto de memória de trabalho) com o total de macroestrutura (CP = ,446; $p = ,018$), o que quer dizer que, quanto maior os escores no exame de MT, maiores os resultados de macroestrutura. Esta correlação sugere que o alto grau de escolaridade pode estar associado a um melhor funcionamento da MT, característica não observada nos outros grupos. Não foram observadas correlações significativas para o grupo de idosos saudáveis de baixa escolaridade.

Em relação à ME, o grupo DA apresentou fortes correlações positivas entre Total de ME (CP = ,604; $p = ,038$) e recordação sem pista (CP = ,802; $p = ,002$) e a construção de macroestrutura. Este dado permite compreender que o desempenho da ME está fortemente relacionado ao processo de construção da macroestrutura textual para este grupo, logo, conforme aumentam (ou diminuem) os escores na tarefa de ME, o mesmo é observado na análise da macroestrutura. Esta correlação pode complementar a análise comparativa entre DA e GC, ao passo que, por mais que a diferença de escores não tenha se demonstrado significativa, as dificuldades de construção de macroestrutura em pessoas acometidas pela DA podem estar associadas às falhas de ME observadas, inclusive, na distinção entre grupos.

A Tabela 10 apresenta as correlações de dados neuropsicológicos e de microestrutura. Por conta do grande número de métricas analisadas, serão apresentados somente os dados cuja correlação se mostrou significativa para algum dos três grupos analisados. Para tanto, são incluídos na tabela valores de média, desvio padrão, e p (quando significativo para correlação positiva ou negativa).

Tabela 10: Resultados de correlação de dados neuropsicológicos x microestrutura

	SAE (N = 28)			SBE (N = 35)			DA (N = 12)		
	N	CP	p	N	CP	p	N	CP	p
ME (TG) x Incidência de palavras funcionais	28	-	-	33	-,364	,037	12	-	-
ME (TG) x Incidência de palavras de conteúdo	28	-	-	33	-	-	12	-,720	,008

	SAE (N = 28)			SBE (N = 35)			DA (N = 12)		
	N	CP	<i>p</i>	N	CP	<i>p</i>	N	CP	<i>p</i>
ME (TG) x Complexidade de Frazier	28	-	-	33	-	-	12	,596	,041
ME (TG) x Incidência de substantivos	28	-	-	33	,358	,041	12	-	-
ME (RSP) x Incidência de conectivos	28	-	-	33	-,475	,005	12	,595	,041
MT (DS-OI) x Incidência de palavras de conteúdo	28	-,455	,015	35	-	-	12	-	-
MT (DS-OI) x Incidência de palavras funcionais	28	,448	,017	35	-	-	12	-	-
MT (DS-OI) x Incidência de Conectivos	28	-	-	35	,371	,028	12	-	-
MT (DS-OI) x Estatística de Honoré	28	,409	,031	35	-	-	12	,681	,021
MT (DS-OI) x relação Tipo por Token	28	-	-	35	-	-	12	,681	,021
MT (DS-OI) x Densidade de Conteúdo	28	-,450	,016	35	-	-	12	-	-

Nota: SAE = Grupo saudável de alta escolaridade, N = número de participantes, M = média, DP = desvio padrão, ME (TG) = Memória Episódica (Total Geral), ME (RSP) = Memória Episódica (Recordação sem pista), MT (DS-OI) = Memória de Trabalho (Digit Span – ordem indireta).

Fonte: O autor (2019)

Para o grupo SAE, não foram observadas correlações entre a memória episódica e o nível de microestrutura. Para a correlação com a MT, observou-se correlação positiva entre incidência de palavras funcionais (CP = ,448; *p* = ,017) e estatística de Honoré (CP = ,409; *p* = ,031) e o teste de ordenamento de dígitos, forma indireta. Esta relação indica que os escores para as métricas são mais altos conforme o aumento dos escores para a tarefa de MT. Essa

correlação possibilita concluir que, pelo menos para o estilo de tarefa de construção de narrativa a partir de sequência de figura, a MT parece contribuir para a escolha de palavras, evitando as repetições (a estatística de Honoré avalia a relação entre a utilização de palavras uma única vez e a repetição delas ao longo do texto) nos textos produzidos por pessoas mais escolarizadas. Houve, além disso, duas correlações negativas entre escores de MT e métricas linguísticas. As correlações negativas acontecem quando uma variável se beneficia da diminuição da outra, isto é, o aumento de escore de uma variável está condicionado à diminuição do escore de outra. Assim como no verificado em relação à análise da macroestrutura, a ME não correlacionou com métricas da microestrutura no grupo mais escolarizado.

Ainda no caso do grupo SAE, a incidência de palavras de conteúdo ($CP = -,455$; $p = ,015$) e a densidade de conteúdo ($CP = -,450$; $p = ,016$) parecem não depender da funcionalidade da MT, podendo estar mais associadas a outros construtos cognitivos não abordados nesta análise.

O grupo de participantes saudáveis de baixa escolaridade correlacionou positivamente os dados de ME (total) e incidência de substantivos ($CP = ,358$; $p = ,041$), evidenciando a relação da memória episódica com a identificação de objetos, pessoas, lugares, etc, palavras que são descritas por substantivos. Correlacionou negativamente os dados de ME (total) e a incidência de palavras funcionais ($CP = -,364$; $p = ,037$). Esta correlação pode estar associada ao fato de que a escolha por palavras funcionais (pronomes, artigos, etc) não esteja condicionada ao bom funcionamento da ME, sendo associada talvez a outros construtos. A correlação negativa entre ME (recordação sem pista) e incidência de conectivos ($CP = -,475$; $p = ,005$) demonstra que também não parece haver forte relação entre a ME e a o uso de conectivos, estando mais associada à MT, conforme aponta a correlação positiva entre a forma indireta do teste de ordenamento de dígitos e o construto de MT para este grupo ($CP = ,371$; $p = ,028$).

Para o grupo DA, observou-se correlações positivas entre os aspectos microestruturais tanto para ME quanto para MT. A complexidade de Frazier, métrica que avalia complexidade sintática do texto, correlacionou positivamente com os escores totais do teste de ME ($CP = ,596$; $p = ,041$), indicando que, quanto melhor o desempenho na ME, melhor também é a construção sintática do texto. Este dado também indica que, se observadas falhas na ME, também serão observadas falhas na métrica sintática, resultado observado na análise da comparação entre grupos (ainda que a diferença entre GC e DA para esta métrica não tenha se mostrado significativa). A correlação entre a etapa de recordação sem pista do teste de ME e a

métrica de incidência de conectivos também se mostrou positiva ($CP = ,595$; $p = ,041$), reforçando a relação entre a complexidade sintática da produção textual e a funcionalidade da ME em quadros demenciais. Houve, ainda, correlação negativa entre o total de ME e a métrica de incidência de palavras de conteúdo ($CP = -,720$; $p = ,008$), indicando que não há uma relação dependente entre o desempenho da ME e a utilização de palavras de conteúdo (como substantivos, verbos, adjetivos, etc). É possível que a capacidade de produção de palavras de conteúdo esteja relacionada a outros construtos, como a memória semântica, conforme discutido no capítulo que trata do referencial teórico. Seria necessário, porém, uma avaliação mais focada nesta relação para confirmar tal hipótese.

O construto de MT também apresentou correlações positivas com métricas linguísticas. Mais especificamente, as métricas de estatística de Honoré ($CP = ,681$; $p = ,021$) e relação tipo por token ($CP = ,681$; $p = ,021$) correlacionaram positivamente com os escores da versão indireta do teste de ordenamento de dígitos, indicando que as repetições de palavras (características da DA) têm relação com o desempenho da MT. Porém, não é possível afirmar se esta relação está condicionada à variável demência ou à baixa escolaridade, visto que não houve diferenças significativas entre os grupos comparados no Estudo 2, com estas características (ainda que a média dos participantes saudáveis de baixa escolaridade tenha sido superior à dos participantes com DA).

Após a apresentação dos dados obtidos através das comparações entre grupos e correlações entre construtos propostas pelos dois estudos conduzidos nesta dissertação, passamos, no próximo capítulo, para a discussão dos resultados, a ser apresentada tendo como objetivo confirmar ou não as hipóteses levantadas para cada um deles.

5 DISCUSSÃO DOS RESULTADOS

De maneira a manter a forma de organização desta dissertação, os resultados serão discutidos de acordo com a divisão de estudos propostos. Sendo assim, na primeira seção apresentaremos as hipóteses levantadas sobre as comparações entre grupos presentes no Estudo 1. Na seção seguinte será realizada a discussão das hipóteses levantadas sobre o Estudo 2.

5.1 DISCUSSÃO DAS HIPÓTESES SOBRE O ESTUDO 1

Nesta etapa da dissertação, serão discutidas as hipóteses elencadas para as análises realizadas no Estudo 1. Em linhas gerais, buscou-se investigar se há diferenças na produção oral de narrativas, tanto no nível macro quanto microestrutural, entre idosos divididos pelo grau de escolaridade (alta e baixa). Além disso, também foi proposta uma investigação sobre a relação entre o desempenho na construção destes níveis do texto e construtos cognitivos de memória de trabalho (MT) e memória episódica (ME). A partir da redação desses objetivos, foram elencadas três hipóteses para o estudo, as quais serão discutidas individualmente, a seguir.

A primeira hipótese levantada era a de que idosos saudáveis de alta escolaridade obteriam desempenho superior (ou seja, escores mais elevados na tarefa...) na produção do nível de macroestrutura na comparação com idosos saudáveis de baixa escolaridade. Com base nos resultados obtidos, nossa hipótese foi corroborada. Foram identificadas diferenças significativas na análise estatística, indicando melhor construção de macroestrutura por participantes mais escolarizados. Além disso, os componentes de desempenho linguístico (como a ausência de sequência narrativa, a predominância de descrição de cenas, a inclusão de informações não pertinentes à história e o não reconhecimento de personagens) também diferenciaram os grupos significativamente, sendo que os menos escolarizados demonstraram piores escores totais.

A análise qualitativa permitiu constatar estas características nas narrativas produzidas por pessoas menos escolarizadas, ao passo que foi possível identificar dificuldades na elaboração dos estágios da narrativa conforme propostos por Adam (2008). Jerônimo (2016) também encontrou dados semelhantes em seu estudo. De acordo com a autora, idosos menos

escolarizados têm mais dificuldade na construção da estrutura de narrativa, assim como na construção de coerência global, que está diretamente relacionada com este construto, conforme indicam Van Dijk e Kintsch (1983).

É importante ressaltar que Salles e Brandão (2013) apontam a inclusão de informações pessoais no discurso como uma característica natural do discurso do idoso, sem mencionar claramente o fator escolaridade. Para as autoras, este fenômeno pode estar mais relacionado ao contexto social ou à identidade do falante, sendo a intenção comunicativa o aspecto mais relevante, o que chamam de hipótese pragmática. Retomamos, então, dados relevantes que dizem respeito ao padrão socioeconômico e cultural (hábitos de leitura e escrita) dos participantes. Na comparação entre grupos, idosos mais escolarizados foram classificados como B2 na escala ABEP (2014), enquanto que os idosos menos escolarizados foram classificados como C1. Além disso, idosos menos escolarizados se diferenciaram significativamente de idosos mais escolarizados no que tange aos hábitos de leitura e escrita, novamente apontando vantagem para o grupo mais escolarizado. A diferença do padrão socioeconômico e cultural pode ser uma variável de distinção de grupo, indicando que pessoas com mais alta escolaridade tendem a ler e escrever mais. O padrão socioeconômico pode estar atrelado à ocupação profissional – assim, profissões mais acadêmicas ou que requerem maior quantidade de leitura e escrita, ou ainda, maior desenvolvimento acadêmico, tendem a favorecer a habilidade de leitura, como verificado no presente estudo. No entanto, não se podem fazer asserções muito contundentes, uma vez que as amostras dos dois grupos são pequenas, além de que o delineamento do estudo não teve como objetivo maior focar nessa discussão.

Essas comparações parecem sugerir que a escolaridade tem impacto na construção do nível macroestrutural de textos, corroborando os achados de Stern (2012) em sua proposta de reserva cognitiva no envelhecimento. Mesmo para as diferenças não apontadas como significativas na análise estatística, as médias do grupo SAE foram superiores às médias do grupo SBE para todas as análises de macroestrutura realizadas. Além disso, a análise qualitativa também apontou diferenças referenciadas na literatura na distinção entre grupos.

Mais um exemplo disso é a questão das pausas. O discurso de idosos menos escolarizados pareceu mais pausado que o de idosos mais escolarizados, característica que também apareceu na comparação das métricas extraídas pelo software Coh-Metrix Dementia. A ferramenta de análise automática tem se mostrado útil para a avaliação de produção textual, conforme analisado no capítulo da fundamentação teórica, o que gerou nossa segunda hipótese para este estudo, de que idosos saudáveis de baixa escolaridade demonstrariam

desempenho inferior na produção do nível de microestrutura do texto, principalmente em métricas de complexidade sintática, entre outras. A avaliação dos dados permitiu constatar que esta hipótese não foi corroborada.

Na distinção entre grupos saudáveis cuja variável é o grau de escolaridade, a única métrica linguística que evidenciou diferenças estatisticamente significativas foi a de incidência de adjetivos. Por outro lado, ainda que não tenham sido apontadas diferenças significativas entre os grupos para as métricas de análise sintática, foi possível identificar valores de médias mais altos para o grupo de alta escolaridade nas métricas de complexidade de Frazier, incidência de palavras funcionais e incidência de conectivos. Na análise qualitativa, essas diferenças também surgiram, caracterizando a narrativa de idosos menos escolarizados como mais simples sintaticamente.

Acredita-se que a falta de significância no resultado das métricas na comparação entre grupos possa estar associada ao processo de segmentação automática das sentenças do texto. Por ainda estar sendo alimentado com textos escritos por pessoas acometidas por demências, a ferramenta continua sendo aprimorada. E há o caso de algumas palavras isoladas, como regionalismos, como *gurizinho*, *guria*, que não são reconhecidas por não comporem os dicionários que alimentam a ferramenta. A Figura 11 apresenta um exemplo de comparação entre uma transcrição manual e a segmentação automática.

Figura 11: Diferença transcrição X segmentação

Texto segmentado:

um homem uma senhora passando na calçada em sentido oposto e tinha um cachorrinho sentado na beira da calçada. o menino chamou o cachorrinho e levou até em casa foi chamando até em casa chegou a entrar em casa. quando a mãe e não na porta da casa daí a mãe abre a porta e vê o repreende. o menino ela coloca as mãos cintura falando com parado. mas depois o cachorrinho já dentro de casa tem uma casinha ele o menino com a amiguinha e até tem um serrote ali fizeram uma casinha tem uma ali fizeram uma cachorrinho até com as patinhas em cima da casinha. e então o cachorrinho vai ficar. na casa

Transcrição

<dist> um </dist> <dist> homem </dist> <dist> e::... </dist> um homem uma senhora passando na calçada em sentido oposto ... e tinha um cachorrinho sentado na beira da calçada ... <disf> aí </disf> o menino chamou o cachorrinho ... e levou até em casa ... foi chamando até em casa ... em casa... chegou a entrar em casa ... <empty> me parece assim </empty> ... quando a mãe ... <disf> ah </disf> e não tá na porta da casa daí a mãe abre a porta e vê o cachorrinho ... <disf> é:: </disf> <empty> pelo jeito </empty> ela repreende o menino ... ela coloca as mãos na cintura ... tá falando com o menino ... o cachorrinho tá parado ... mas depois o cachorrinho já tá dentro de casa ... <disf> numa </disf> <disf> cas </disf> tem uma casinha ... ele tá tá: o menino com a amiguinha e:: até tem um serrote ali fizeram uma casinha tem <disf> um </disf> <disf> u::m </disf> uma torquesa ali <empty> qualquer coisa </empty> fizeram <disf> uma </disf> <disf> ãh </disf> uma casinha pro cachorrinho ... o cachorrinho tá até com as patinhas em cima da casinha ... e:: então o cachorrinho vai ficar na casa ...

Fonte: O autor (2019)

É possível identificar, no exemplo acima citado, que o software pode apresentar dificuldades na delimitação de algumas sentenças, como, por exemplo, logo no final, em que o segmentador delimita o final da sentença após *...vai ficar.*, fazendo com que o complemento *na casa* se separe do restante da sentença. Essa quebra pode influenciar no resultado das métricas que derivam da análise de *parsers* (árvores sintáticas), como, por exemplo, a Complexidade de Frazier, a Complexidade de Yngve, e a Densidade de Conteúdo, que não mostraram diferença significativo no estudo aqui apresentado. Esta limitação do software é prevista pelo desenvolvedor, que defende que, quanto mais análises de transcrições de textos orais forem realizadas, melhor será o resultado final da segmentação, visto que o sistema atualmente é alimentado principalmente de modelos de texto escrito, que se diferem da expressão oral (TREVISÓ, 2017). Apesar das pequenas limitações atinentes a uma técnica em refinamento constante, a ferramenta apresenta excelentes resultados. Treviso et. al. (2018) compararam a segmentação manual com o desempenho do DeepBonD em resultado de métricas extraídas de transcrições de narrativas produzidas por pessoas acometidas por CCL e controles e encontraram diferença significativa entre os métodos de segmentação apenas para a análise da métrica de modificadores por sintagma nominal. Este dado confirma a eficácia e acurácia da ferramenta da delimitação de sentenças automaticamente, o que pode contribuir para que, muito em breve, seja possível realizar uma análise totalmente automática de textos orais transcritos, o que torna o processo mais ágil e acessível a outros profissionais, que não somente da área de Letras.

De maneira geral, as duas primeiras hipóteses estão relacionadas ao desempenho linguístico dos participantes. Como pudemos observar, diferentemente do que se esperava, foram encontradas diferenças estatisticamente significativas apenas para a construção de macroestrutura, diferenciando os grupos de escolaridades distintas com vantagem para os mais escolarizados. A análise da microestrutura não permitiu encontrar diferenças entre as métricas extraídas, a não ser na incidência de adjetivos, que beneficia, novamente, os participantes mais escolarizados. A análise qualitativa permitiu, porém, discutir algumas diferenças relativas à complexidade sintática dos textos. Passamos, a seguir, para a discussão da terceira hipótese levantada neste estudo, que dá conta da correlação entre testes neuropsicológicos e desempenho linguístico nos diferentes graus de escolaridade.

A hipótese 3 postulava que não haveria diferenças entre os grupos saudáveis originadas pelo grau de escolaridade no desempenho da memória episódica, mas esperava encontrar diferença entre grupos no desempenho da memória de trabalho. Já na correlação entre o desempenho linguístico dos grupos (métricas da micro- e total de macroestrutura) e

construtos cognitivos, esperava-se que houvesse correlação positiva. Esta hipótese foi parcialmente corroborada. Primeiramente, não houve diferença significativa entre grupos para o teste de ME, o que sugere que este construto não é afetado pelo grau de escolaridade. Estes achados corroboram os estudos de Hübner et. al. (2018) e Jerônimo (2016), que não encontraram diferenças significativas para a variável escolaridade na distinção entre grupos de idosos. Esperava-se encontrar uma diferença significativa para MT, mais especificamente no teste de ordem indireta, conforme identificado por Jerônimo (2016), que não foi constatada na comparação de resultados do nosso estudo, ainda que a média do grupo mais escolarizado tenha sido superior a do menos escolarizado. Foi identificada uma diferença significativa na ordem direta do teste (que avalia a atenção) e no escore total, da soma das duas etapas, apontando vantagem para o grupo mais escolarizado. Estes dados não corroboram os encontrados por Jerônimo (2016), que identificou diferença significativa entre idosos saudáveis de alta e baixa escolaridade na ordem indireta do teste, apontando vantagem para o grupo saudável de alta escolaridade. De acordo com os resultados da autora, o desempenho nesta etapa do teste poderia sofrer um impacto do grau de escolaridade, que não foi identificado neste estudo. Entretanto, o fato de a soma total ter apresentado diferença significativa entre os grupos faz com que fique evidente uma diferença no desempenho da MT de maneira geral, possivelmente oriunda da diferença entre o grau de escolaridade.

Os resultados das comparações entre grupos para testes neuropsicológicos parecem corroborar os resultados encontrados nos testes de correlação. O grupo saudável de alta escolaridade apresentou correlação positiva entre o desempenho da MT (ordem indireta) e a construção da macroestrutura da narrativa. Voltando aos dados linguísticos, o grupo menos escolarizado apresentou piores escores, obtendo diferença significativa na comparação com o mais escolarizado para a construção deste nível do texto. Logo, este dado sugere que a MT pode atuar na melhor construção do nível macroestrutural do texto, trazendo vantagem, neste caso, ao grupo mais escolarizado, visto que a mesma correlação não foi identificada no grupo menos escolarizado. De fato, nenhum dos construtos de memória avaliados (MT e ME) correlacionaram com a construção de macroestrutura para o grupo menos escolarizado, o que sugere desempenho inferior em idosos com menor grau de instrução.

Na correlação entre construtos de memória e métricas de microestrutura, observou-se correlação positiva entre o desempenho da MT e as métricas de incidência de palavras funcionais e estatística de Honoré para o grupo mais escolarizado. Este dado sugere que, quanto maiores os escores na avaliação de MT, melhor o desempenho nas métricas. Esta correlação não foi identificada para o grupo menos escolarizado, cuja única correlação de

microestrutura com MT ficou associada à incidência de conectivos. Acredita-se que esta falha de MT para o grupo menos escolarizada possa estar associada ao déficit de inibição. Conforme visto anteriormente, a hipótese do déficit de inibição relaciona as falhas na MT às falhas na função executiva da inibição, que acaba permitindo que haja interferências no desempenho na MT durante a produção do discurso (SALLES; BRANDÃO, 2013). Entretanto, não é possível afirmar, com os dados coletados nesta pesquisa, se esta hipótese é confirmada, ou ainda se está relacionada ao grau de escolaridade ou alguma outra variável, como a idade, por exemplo.

Encontrou-se, ainda uma relação negativa entre a MT e as métricas de palavras de conteúdo e densidade de conteúdo, para o grupo mais escolarizado. Este dado sugere que a seleção de palavras como substantivos, verbos, adjetivos e advérbios não esteja relacionada ao bom funcionamento da MT. Neste sentido, sugere-se avaliar o desempenho da memória semântica (MS) na correlação com estas métricas, visto que é neste construto que se observa o armazenamento do vocabulário (EINSECK, 2011). Não é possível afirmar, entretanto, com os testes aqui administrados e analisados, se há diferenças no acesso à MS oriundas do grau de escolaridade.

O grupo menos escolarizado apresentou correlação positiva entre o desempenho na memória episódica e a incidência de substantivos no texto. Acredita-se que a utilização de substantivos na descrição dos personagens da história, assim como a repetição deles, possa explicar esta relação. Além disso, conforme exemplo apresentado na análise dos dados, é comum no texto do idoso o não reconhecimento específico de personagens, sendo frequentes um indicativo de todas as pessoas que compõem a cena, ou ainda da indicação *outro(a)* para se referir aos personagens da história. Novamente associamos os resultados a uma possível relação com a memória semântica. As memórias episódica e semântica se relacionam em muitos casos, visto que é preciso acessar informações armazenadas na MS para descrever um evento específico (Eysenck, 2011). No caso da história do cachorrinho, o participante precisou construir um modelo cognitivo do texto a partir da sequência de figuras na ME, que pode ter atuado em conjunto com a MS na construção da narrativa. Novamente, não é possível afirmar tal relação, sendo necessários estudos complementares, com foco no desempenho da MS e possível relação com estes construtos.

Ainda sobre as correlações entre microestrutura e memória episódica, o grupo saudável de baixa apresentou correlação negativa entre o desempenho do construto e o resultado das métricas de incidência de palavras funcionais e conectivos. Aparentemente, a correlação positiva entre MT e incidência de conectivos para este grupo parece explicar que a

demanda de componentes sintáticos mais complexos, como o uso de preposições, conjunções e conectivos, possa estar mais relacionada ao bom funcionamento da MT, e não da ME, que pareceu, para este grupo, mais relacionada ao desempenho de palavras de conteúdo (substantivos).

Não foi observada relação entre ME e construção da narrativa para o grupo mais escolarizado, mas aspectos da correlação positiva entre ME e microestrutura para o grupo menos escolarizado levantaram a hipótese de uma possível influência da memória semântica na produção textual, hipótese esta que necessita estudo complementar para ser confirmada.

Concluídas as discussões dos dados obtidos para o estudo 1, longe de delimitá-las, apresentaremos, na seção a seguir, a discussão dos dados obtidos através das análises propostas no estudo 2, cuja variável dependente é a Doença de Alzheimer. Novamente, a discussão será conduzida tendo como base a confirmação das hipóteses levantadas sobre o estudo.

5.2 DISCUSSÃO DAS HIPÓTESES SOBRE O ESTUDO 2

Para o estudo dois, objetivos semelhantes aos do Estudo 1 foram redigidos, com apenas uma diferença: ao invés de analisar a escolaridade como variável, o foco do Estudo 2 foi nas diferenças oriundas do quadro demencial estabelecido pelo diagnóstico da Doença de Alzheimer. Sendo assim, um grupo de idosos de baixa escolaridade com DA foi comparado a um grupo de idosos saudáveis também de baixa escolaridade visando a investigar as diferenças na construção da micro e macroestruturas de uma narrativa, assim como as relações de MT e ME com estes processos. Destes objetivos também foram levantadas três hipóteses, que serão discutidas a seguir.

A primeira hipótese assumida para este estudo foi a de que idosos saudáveis obteriam melhor desempenho (escores mais altos) na produção do nível da macroestrutura na comparação com idosos diagnosticados com DA. Esta hipótese não foi corroborada. A análise estatística quantitativa não diferenciou os grupos significativamente no total de macroestrutura, nem nas duas esferas de análise de desempenho linguístico. Entretanto, a análise qualitativa permitiu observar diferenças referenciadas na literatura, como falhas na construção de sentenças (palavras soltas), desrespeito à estrutura narrativa, predominância na descrição de cenas, pausas longas, apontando vantagem para o grupo controle. Foi possível caracterizar a narrativa produzida por pessoas com DA como menos coerente, de sentenças

mais simples, com falhas de estrutura e desrespeito aos estágios propostos por Adam (2008). Também, como característica da tarefa, foi possível identificar uma maior incidência de descrição de cenas em detrimento de sequencialidade comum ao texto narrativo nos exemplos produzidos por idosos com DA. Estes dados corroboram os achados de Kavé e Dassa (2018), Soares, Brandão e Lacerda (2012) e Brandão e Parente (2010), que descreveram características semelhantes sobre o discurso de pessoas com DA.

Acredita-se, portanto, que a semelhança de resultados entre os grupos possa estar associada ao baixo grau de escolaridade de ambos, que, conforme descrito por Parente et. al. (2009), pode dificultar na diferenciação entre os grupos. Outro fator que pode ter influenciado nos resultados é o baixo número de participantes que compõem o grupo DA. Não é possível afirmar que um grupo maior de narrativas produzidas por pessoas acometidas pela DA diferenciaria os grupos, mas entende-se que, quanto maior a amostra, maior é a força de explicação dos resultados. A análise da microestrutura, todavia, deixou bastante evidente a diferença das produções entre grupos. Nossa hipótese inicial era a de que pessoas acometidas pela DA demonstrariam déficits na produção no nível da microestrutura do texto, observada na análise das métricas, em especial no uso de pronomes, na complexidade sintática, entre outras. A hipótese foi corroborada. A comparação entre grupos permitiu observar uma distinção significativa no desempenho de produção de microestrutura em cinco métricas linguísticas (incidência de advérbios, incidência de palavras de conteúdo, incidência de palavras funcionais, incidência de pronomes pessoais e incidência de conectivos) e uma métrica de disfluência (duração média de pausas). Os resultados de métricas corroboram estudos de Orimaye et. al. (2017) e Fraser et. al. (2016), que encontraram resultados semelhantes aos nossos na comparação de textos produzidos por pessoas com DA e controles saudáveis. As narrativas produzidas por DA evidenciaram maior uso de advérbios e palavras de conteúdo no texto, o que não necessariamente configura o texto de DA como mais elaborado que o texto de controles saudáveis, como nossa análise qualitativa permitiu constatar, e como mostram os estudos de Fraser et al (2016) e de Orimaye et al. (2017). Mais especificamente, Fraser et. al. (2016) também identificaram no texto de DAs vocabulário restrito, uso das mesmas palavras, e construção de fragmentos de sentenças em detrimento à construção de sentenças completas. Da mesma forma, Orimaye et. al. (2017) identificaram falhas de sintaxe e léxico, como a repetição frequente de palavras e diminuição na construção de sentenças mais complexas. Neste sentido, os dados obtidos neste estudo também se relacionam aos encontrados naquelas pesquisas. Ainda que não significativa, foi observada diferença na frequência de palavras de conteúdo, sendo que controles saudáveis utilizaram

palavras menos frequentes do que pessoas com DA, o que pode ser explicado pela maior dificuldade de acesso lexical de palavras menos frequentes por pessoas com DA, oriundas do empobrecimento do vocabulário, conforme indicam Sales, Brandão e Lacerda (2012). Da mesma forma, salientamos a diferença, ainda que não significativa, na métrica de relação tipo por token, que apontou vantagem para o grupo controle na comparação com DA. A utilização de palavras de conteúdo mais frequentes na DA pode ter relação com esta métrica pela evidência da repetição de palavras oriundas do quadro demencial, responsáveis, também, pelo empobrecimento do vocabulário já apontado na literatura (SALES, BRANDÃO, LACERDA, 2012). Também foi identificada diferença significativa com maior média para o grupo controle na incidência de pronomes pessoais, dado que esperávamos fosse diferente, a julgar pelos resultados de Fraser et. al. (2016), que apontam que, em geral, pessoas com DA costumam utilizar mais pronomes referenciais em detrimento do uso de substantivos, característica que não ficou evidente na comparação realizada neste estudo. Conforme pesquisa bibliográfica, fica evidente uma maior incidência do uso de pronomes (de maneira geral) na DA, porém nem sempre de maneira adequada, isto é, idosos acometidos pela DA tendem a utilizar pronomes vagos, sem referentes ou com referentes indefinidos, ambíguos (KAVÉ E DASSA, 2018; VARGAS, 2015; SOARES, BRANDÃO, LACERDA, 2012).

Em relação às pausas, novamente corroboramos os achados do estudo de Fraser et. al. (2016). Identificamos diferença significativa na duração média das pausas na comparação entre grupo controle e DA, sendo que o grupo DA apresentou construções com pausas mais longas do que controles saudáveis. Fraser et. al. (2016), a partir da análise de médias acústicas, também identificaram aumento na ocorrência de pausas, característica que diferenciou grupos controle e DA. Essa característica aponta novamente para o benefício oriundo do uso de ferramentas de análise automática de textos. Treviso et. al. (2018) postulam, como plano de aprimoramento para ferramenta DeepBonD de segmentação de sentenças, a implementação da função de segmentação de sentença a partir do arquivo de áudio, em conjunto com a forma escrita. Para os autores, esta nova função da ferramenta pode refinar ainda mais as análises de disfluências como pausa e hesitações, características do discurso na DA.

Em linhas gerais, a análise da micro e da macroestrutura do texto permitiu observar algumas características específicas da produção narrativas de idosos acometidos pela DA. A nível de macroestrutura, nossos dados não identificaram diferenças significativas entre o grupo controle e o grupo DA. A análise qualitativa, por outro lado, permitiu observar características pontuadas em estudos realizados anteriormente, como o desrespeito à

sequência narrativa e a descrição de cenas, característica mais associada ao tipo de tarefa aplicada. A microestrutura, a partir da avaliação das métricas, permitiu observar que as narrativas produzidas por pessoas acometidas pela DA eram constituídas, em maioria, de sentenças fragmentadas, com menor incidência de elementos de conexão ou palavras funcionais, na comparação com controles saudáveis.

Assim, partimos, a seguir, para a discussão dos dados neuropsicológicos obtidos na comparação entre grupos, assim como as correlações realizadas com os dados das tarefas linguísticas.

A terceira hipótese levantada para este estudo era a de que haveria diferenças significativas entre os grupos para tarefas de memória episódica e de memória de trabalho, além de uma correlação positiva entre o desempenho linguístico dos grupos de idosos saudáveis e de idosos acometidos pela DA (métricas de micro- e total de macroestrutura) e construtos cognitivos (memória de trabalho e episódica). Nossa hipótese foi parcialmente corroborada. Conforme esperado, os resultados da comparação de desempenho dos grupos no teste de memória episódica evidenciaram diferença significativa entre eles, sendo que pessoas acometidas pela DA obtiveram piores escores do que controles saudáveis. Os dados corroboram estudos anteriores de Grober, Veroffe e Lipton (2018) e Baddeley (2011b). De acordo com Grober, Veroffe e Lipton (2018), pessoas acometidas pela DA demonstram declínio primeiramente na etapa de recordação sem pista no teste *Free and Cued Selective Reminding Test* (FCSRT). Nossos dados permitem observar esta tendência, uma vez que controles saudáveis obtiveram melhor desempenho que o grupo DA nesta etapa do teste. Esta falha na recordação sem pista às vezes é equalizada na etapa da recordação com pista, que favorecem o grupo controle, já que o grupo DA tende a não se beneficiar com as pistas, o que faz com que falhas no total do teste sejam fortes indicativos de DA (GROBER, VEROFFE E LIPTON, 2018). Também identificamos esta diferença, em que o grupo DA diferiu significativamente do grupo controle no total para a tarefa de ME. Baddeley (2011b) também relaciona os déficits de ME com o surgimento de sintomas da DA.

Ademais, uma falha na ME poderia evidenciar falhas na construção da macroestrutura do texto, visto que, segundo o modelo situacional de produção textual de Van Dijk e Kintsch (1983), uma construção mental do sentido do texto é armazenada na ME como um estágio de pré-produção textual, que envolve ainda estratégias de planejamento, entre outras. A correlação de dados de total de macroestrutura e total de ME mostrou-se significativa para o grupo DA, reforçando a alta demanda do construto cognitivo no processo de elaboração da esfera macroestrutural do texto. Logo, uma falha na ME pode acarretar falhas na produção da

macroestrutura, correlação não identificada para o grupo controle, conforme discutido na seção anterior. O desempenho na tarefa de ME não está relacionado com a formação da macroestrutura em idosos saudáveis, pelo menos não na presente amostra, o que pode indicar que esta característica esteja condicionada à variável demência.

Novamente destacamos a acurácia da tarefa de ME baseada em recordação livre e com pista para o diagnóstico de DA. Conforme avaliado, nossos dados corroboram estudos consultados (TEICHMANN, ET. AL., 2017; LEMOS ET. AL., 2017; LEMOS ET. AL., 2014), trazendo ainda mais evidências de que a tarefa é sensível a esta diferença entre idosos saudáveis e com DA, ou seja, é eficiente no diagnóstico da DA na comparação com controles saudáveis. Além disso, a tarefa parece não ser influenciada pelo nível de escolaridade, o que ameniza a preocupação de que pessoas menos escolarizadas podem ter o diagnóstico afetado por dificuldades cognitivas oriundas do baixo grau de escolaridade (e/ou sociocultural) (PARENTE, ET. AL., 2009). Ademais, apontamos para a importância de avaliações neuropsicológicas deste tipo, que podem contribuir para o diagnóstico precoce, auxiliando também na identificação de marcadores da doença.

Para a análise de microestrutura, o grupo DA correlacionou negativamente o total de ME com a métrica de incidência de palavras de conteúdo, o que indica que as variáveis não evoluem paralelamente. Este dado parece evidenciar que o deterioro na ME não está relacionado com o desempenho na utilização de palavras de conteúdo no texto, isto é, os substantivos, verbos, adjetivos e advérbios. Este dado parece estar relacionado com a análise qualitativa das transcrições, ao passo que, conforme observado anteriormente, os participantes com DA não diferenciaram significativamente dos participantes saudáveis para as métricas de palavras de conteúdo, o que indica que a falha de ME parece não impactar no acesso às palavras, que, como também vimos anteriormente, poderia estar mais relacionado à memória semântica, hipótese que não pode ser respondida neste estudo. Por outro lado, o desempenho da ME correlacionou positivamente com as métricas de incidência de conectivos e complexidade de Frazier, o que coloca as falhas na ME como associadas às falhas na complexidade sintática das sentenças. Novamente retomamos os pressupostos de Van Dijk e Kintsch (1983), que relacionam a ME com o processo de produção textual no nível do encadeamento das sentenças.

A análise da MT, por outro lado, não evidenciou diferenças entre os grupos para os resultados da tarefa de ordenamento de dígitos ordem indireta, diferenciando-se dos resultados encontrados por Jerônimo (2016) e Lemos et. al. (2014), que identificaram diferenças significativas entre DA e controles saudáveis nesta etapa da tarefa. Entretanto,

foram identificadas diferenças significativas na ordem direta do teste, assim como no total da soma entre as ordens direta e indireta, o que indica que pode, sim, haver diferenças no desempenho da MT oriundas do quadro demencial.

Retomamos, neste sentido, a questão da escolaridade novamente como uma influência nos resultados de tarefas cognitivas. A falta de diferença significativa entre os grupos GC e DA (ambos de baixa escolaridade) parece evidenciar e corroborar os achados de Parente et. al. (2009), além de ir ao encontro da advertência da *American Psychiatric Association* (2014) de que a baixa escolaridade pode interferir nos resultados e atrapalhar diagnósticos de demências, visto que pessoas menos escolarizadas parecem ter desempenhos inferiores em tarefas cognitivas, como também observamos a partir da análise dos dados deste estudo. Esta influência pode ocorrer também no desempenho em tarefas linguísticas, visto que o padrão socioeconômico mais baixo, relacionado à escolaridade mais baixa, pareceu atrelado aos hábitos de leitura e escrita também inferiores, que, por sua vez, podem impactar no desempenho linguístico. Logo, parece pertinente explorar ainda mais os impactos do grau de escolarização na comparação entre grupos controle e populações clínicas em tarefas neuropsicológicas e linguísticas, visando a encontrar parâmetros para análise, assim como pontos de corte específicos para a variável escolaridade.





A correlação de dados de MT com dados linguísticos não demonstrou relação do construto com o total de macroestrutura para o grupo DA, nem para o grupo controle. Em relação à microestrutura, foi identificada correlação positiva dos escores do teste de ordenamento de dígitos ordem indireta com as métricas de estatística de Honoré e relação tipo por token. Ambas as métricas têm relação com o uso e repetição de vocabulário na construção das narrativas. Este resultado permite fazer uma relação com outros dados obtidos na comparação entre grupos, que dá conta de explicitar que o texto de DA costuma ter vocabulário simples, com palavras repetidas e mais frequentes lexicalmente do que as narrativas produzidas por participantes saudáveis. Logo, essa correlação positiva parece indicar que esta característica está condicionada ao funcionamento da MT, ou seja, quanto melhor o desempenho da MT, melhor o resultado nas métricas analisadas.




Em geral, as hipóteses elaboradas para o Estudo 2 foram corroboradas parcialmente pelos testes estatísticos, e complementadas a partir das análises qualitativas realizadas. Foi possível constatar que o construto de memória episódica, amplamente documentado na literatura como um dos principais déficits na DA, desde os estágios iniciais, pareceu relacionado com o desempenho dessa população na elaboração tanto da macro quanto da microestrutura. As narrativas na DA, por esta razão, parecem ter problemas de estrutura, não

respeitando os estágios conforme propostos por Adam (2008). Também ficaram evidentes o uso de vocabulário mais simples, de palavras mais frequentes, e de complexidade sintática baixa (ainda que não tenha se diferenciado significativamente do grupo controle na análise estatística, mas evidente na análise qualitativa).

O Quadro 4, abaixo, apresenta um resumo da discussão das hipóteses do Estudo 1 e do Estudo 2, indicando se foram ou não corroboradas.

Quadro 4: Resumo da discussão das hipóteses

	Hipótese	Corroborada	Discussão
Estudo 1	- Idosos saudáveis de alta escolaridade obterão desempenho superior (ou seja, escores mais elevados nas tarefas de compreensão da macroestrutura do texto) na produção do nível de macroestrutura na comparação com idosos saudáveis de baixa escolaridade.		Foram identificadas diferenças significativas na comparação entre grupos SAE e SBE na construção da macroestrutura.
	- Idosos saudáveis de baixa escolaridade demonstrarão desempenho inferior na produção do nível de microestrutura do texto, observado na análise das métricas, em especial na complexidade sintática, entre outros.		Somente a métrica de incidência de adjetivos distinguiu os grupos SAE e SBE significativamente.
	- Espera-se que não haja diferenças entre os grupos saudáveis originadas pelo grau de escolaridade no desempenho da memória episódica, mas que haja	 	Foram identificadas correlações entre construtos de ME e MT e níveis de micro e macroestrutura, com divergências entre grupos.

	diferença entre grupos no desempenho da memória de trabalho; na correlação entre o desempenho linguístico dos grupos (métricas de micro- e total de macroestrutura) e construtos cognitivos (memória de trabalho e episódica), espera-se que haja correlação positiva.		
Estudo 2	- Idosos saudáveis obterão melhor desempenho (escores mais altos nas tarefas de compreensão da macroestrutura do texto) na produção do nível da macroestrutura na comparação com idosos diagnosticados com DA.		Não foram identificadas diferenças significativas entre grupos DA e GC na construção de macroestrutura.
	- Pessoas acometidas pela DA demonstrarão desempenho inferior no nível da microestrutura do texto, observada na análise das métricas, em especial no uso de pronomes, na complexidade sintática, entre outras.		Houve diferença significativa na incidência de métricas mais relacionadas ao léxico na construção de microestrutura.
	- Haverá diferenças significativas entre os grupos para tarefas de memória episódica e de		Houve diferença significativa nos dados de ME, assim como correlações entre construtos de MT e ME

	<p>memória de trabalho, além de uma correlação positiva entre o desempenho linguístico dos grupos de idosos saudáveis e de idosos acometidos pela DA (métricas de micro- e total de macroestrutura) e construtos cognitivos (memória de trabalho e episódica); especificamente, escores mais baixos nos construtos da avaliação neuropsicológica do grupo DA na comparação com o saudável correlacionarão com desempenho linguístico mais baixo, tanto na produção da micro- quanto da macroestrutura.</p>		<p>e níveis de micro e macroestrutura das narrativas.</p>
--	--	--	---

Nota: SAE = saudáveis de alta escolaridade; SBE = saudáveis de baixa escolaridade; DA = Doença de Alzheimer; GC = grupo controle; ME = memória episódica; MT = memória de trabalho.

Fonte: O autor (2019)

Partimos, portanto, para as considerações finais desta pesquisa, em que serão apresentados, resumidamente, os objetivos, principais resultados, limitações, diretrizes para estudos futuros e contribuições dos estudos aqui debatidos.

6 CONSIDERAÇÕES FINAIS

Esta dissertação foi organizada em dois estudos. O primeiro, com foco na escolaridade, visava a comparar dois grupos (um de idosos saudáveis de alta escolaridade e outro de idosos saudáveis de baixa escolaridade) a fim de investigar as diferenças oriundas da produção oral de narrativa no nível de macroestrutura (a partir do uso de macroproposições), na microestrutura (com base na análise de métricas extraídas pela análise automática de textos) e construtos de memória de trabalho (MT) e memória episódica (ME) (de acordo com escores oriundos da administração de testes neuropsicológicos). O segundo estudo também tinha por objetivo realizar uma comparação entre dois grupos, sendo estes divididos entre idosos saudáveis de baixa escolaridade e idosos acometidos pela DA, também de baixa escolaridade. Neste estudo, a variável de observação foi a demência. Para tanto, foram investigados os mesmos construtos descritos para o primeiro estudo.

O Estudo 1 foi composto de dois grupos de participantes idosos, um de alta escolaridade (SAE, N=28) e outro de baixa escolaridade (SBE, N=35). O Estudo 2 foi composto de dois grupos de participantes idosos, sendo um de idosos saudáveis (GC, N=35) e outro de idosos acometidos pela doença de Alzheimer (DA, N=12). Ambos os grupos foram avaliados neuropsicologicamente a partir da administração de testes de memória episódica e memória de trabalho. Também realizaram uma tarefa linguística de produção oral de narrativa a partir da sequência de figuras denominada *The Dog Story* (História do Cachorro), de Le Boeuf (1976). A partir da construção narrativa, foram analisados os níveis de macroestrutura, a partir da produção de macroproposições, e de microestrutura, em métricas extraídas pelo software Coh-Metrix Dementia (ALUÍSIO, CUNHA E SCARTON, 2016).

Para o Estudo 1, três hipóteses foram levantadas. A primeira era a de que idosos mais escolarizados teriam melhor desempenho na produção da macroestrutura das narrativas na comparação com idosos menos escolarizados. A hipótese foi corroborada. A análise estatística apontou diferenças significativas para o total de macroestrutura, evidenciando vantagem para o grupo de alta escolaridade. As principais diferenças ficaram por conta do desrespeito à estrutura narrativa, conforme proposta por Adam (2008), inclusão de informações extra-textuais e predominância de descrição de cenas no texto de pessoas de baixa escolaridade, afetando a construção do nível macroestrutural das narrativas.

A segunda hipótese levantada era a de que a produção do nível microestrutural seria mais deficiente nas narrativas produzidas por pessoas menos escolarizadas, em comparação

com mais escolarizadas, a partir da análise das métricas extraídas da análise automática. A hipótese não foi corroborada, visto que apenas a métrica de incidência de adjetivos diferenciou os grupos. Entretanto, a análise qualitativa permitiu observar que, em geral, o texto de idosos menos escolarizadas tem menor complexidade sintática do que o texto dos mais escolarizados.

A terceira hipótese redigida para o Estudo 1 dizia respeito aos resultados dos testes neuropsicológicos e sua relação com desempenho linguístico. Esperava-se que idosos saudáveis de alta escolaridade não apresentassem resultados significativamente diferentes na comparação com menos escolarizados para o teste de memória episódica, mas esperava-se que houvesse diferença significativa no processamento de memória de trabalho. Além disso, também se esperava que houvesse uma correlação positiva entre os dados de macro e microestrutura e os construtos cognitivos de memória episódica e de trabalho. A hipótese foi parcialmente corroborada. Idosos saudáveis de alta escolaridade não diferiram dos menos escolarizados nos escores para a avaliação de memória episódica, indicando que o teste utilizado não é afetado pelo grau de escolaridade, o que era nossa expectativa. Em relação à MT, a hipótese foi parcialmente corroborada, pois saudáveis de alta escolaridade obtiveram melhor desempenho na MT Total e na ordem direta de dígitos, mas não na ordem inversa de dígitos. No entanto, os escores de MT e total de macroestrutura correlacionaram positivamente, indicando que quanto melhor o desempenho na MT, melhor a construção da macroestrutura. Saudáveis de baixa escolaridade não correlacionaram os dados de ME com os de macroestrutura. No nível de microestrutura, foi identificada correlação positiva para as métricas de incidência de palavras funcionais e estatística de Honoré, e correlação negativa entre as métricas de palavras de conteúdo e densidade de conteúdo com os escores de MT para o grupo saudável de alta escolaridade. No grupo saudável de baixa escolaridade correlacionaram-se os escores de ME com as métricas de incidência de substantivos positivamente, e incidência de palavras funcionais e incidência de conectivos negativamente. Além disso, no grupo de baixa escolaridade também correlacionaram-se positivamente os dados de MT com a métrica de incidência de conectivos.

Em geral, os dados apresentados neste estudo permitem corroborar a hipótese da reserva cognitiva proposta por Stern et. al. (2003), Stern (2012) e Reuter-Lorenz e Park (2014), que postulam que a escolaridade e os hábitos de leitura e escrita ao longo da vida têm impacto positivo no processo de envelhecimento mais saudável. Observou-se influência positiva do grau de escolaridade em ambos os níveis de produção da narrativa, evidentes,

mais especificamente, de maneira quantitativa e qualitativa (para construção de macroestrutura), e qualitativa (na microestrutura). Também ficaram evidentes características do discurso do idoso como a diminuição da complexidade sintática, o acréscimo de informações pessoais, conforme descritos por Salles e Brandão (2013). Acredita-se, ainda, que a não observância desses fatores nos textos produzidos pelos participantes mais escolarizados esteja relacionada à reserva cognitiva.

Para o Estudo 2, outras três hipóteses foram levantadas. A primeira era de que idosos saudáveis teriam melhor desempenho na construção do nível de macroestrutura. A hipótese não foi corroborada. A análise estatística não identificou diferenças significativas para o total de macroestrutura e desempenho linguístico, porém, a análise qualitativa indicou diferenças condizentes com a literatura para as características do texto de DA. Dentre as principais características surgiram as falhas na identificação dos personagens da história, dificuldades na construção dos estágios da narrativa, predominância na descrição de cenas, pausas e hesitações.

Novamente retomamos a hipótese de que o grau de escolaridade possa ter influenciado este resultado, visto que essas são características comuns no discurso do idoso, mesmo saudável. Reforça essa hipótese a orientação da American Psychiatric Association (2014), corroborada pelos achados de Parente et. al. (2009), que colocam o baixo grau de escolaridade como um fator de interferência no desempenho de idosos em tarefas de avaliação neuropsicológica. Acredita-se que esta relação também possa surgir na análise linguística, sendo necessárias, entretanto, investigações futuras para atestar tal afirmação.

Por outro lado, a análise da microestrutura permitiu corroborar a segunda hipótese levantada para o estudo, que esperava que a análise das métricas diferenciase os grupos significativamente, especialmente na complexidade sintática e uso de pronomes, entre outros. Não foram identificadas diferenças significativas na distinção entre grupos para a análise das métricas sintáticas. Entretanto, foram identificadas diferenças significativas entre as métricas de incidência de advérbios, incidência de palavras de conteúdo, incidência de palavras funcionais, incidência de pronomes pessoais e incidência de conectivos. A análise qualitativa desses resultados permitiu compreender que, em geral, o texto de DA é composto de fragmentos de sentenças, apresentando falhas de palavras funcionais e/ou conectivos. Além disso, evidenciou que pessoas com DA utilizam palavras de maior frequência lexical na comparação com idosos saudáveis. Características semelhantes foram reportadas por Soares, Brandão e Lacerda (2012) e Brandão e Parente (2010).

Os dados microlinguísticos corroboraram alguns achados, e não corroboraram outros, na comparação com estudos de Orimaye et. al. (2017) e Fraser et. al. (2016) na observação de métricas linguísticas extraídas dos textos. A análise automática dos textos demonstrou potencial na distinção entre grupos de idosos saudáveis e acometidos por demências, evidenciando, mais especificamente, o software Coh-Metrix Dementia (ALUÍSIO, CUNHA E SCARTON, 2016) como um grande aliado de linguistas e profissionais da área da saúde como complementar no diagnóstico de demências a partir de análise de discurso.

Para a terceira hipótese do estudo, era esperada uma diferença significativa entre os escores de ME e MT na comparação entre grupos, assim como correlação positiva entre aspectos micro e macrolinguísticos e construtos cognitivos. A hipótese foi parcialmente corroborada. A tarefa de ME distinguiu significativamente os grupos, evidenciando déficits para o grupo DA. Não foram identificadas diferenças significativas no teste de MT, diferentemente do que se esperava. Dados de ME correlacionaram positivamente com formação de macroestrutura e com as métricas de complexidade de Frazier e incidência de conectivos para o grupo DA, evidenciando que as falhas da macroestrutura podem estar associadas aos déficits de ME para a população clínica. Também foram identificadas correlações positivas entre MT e as métricas de estatística de Honoré e relação tipo por token para o grupo clínico, indicando que as repetições e pobreza de vocabulário podem estar associadas ao funcionamento da MT.

Outro importante dado oriundo desta análise é a acurácia da tarefa de avaliação de ME, baseada no aprendizado e recordação de figuras, para o diagnóstico de demências. Neste sentido, corroborou-se os estudos de Teichmann et. al. (2017), Lemos et. al. (2017) e Lemos et. al. (2014), que também apontaram a ferramenta de avaliação como acurada na distinção de grupos pela variável demência. Além disso, a tarefa também não se mostrou sensível à diferença do grau de escolaridade dos grupos, contemplando a preocupação apresentada pela *American Psychiatric Association* (2014).

Entende-se que o número reduzido de participantes possa ter atuado na interpretação dos dados e resultados aqui discutidos, apontando estes fatores como limitações desta pesquisa. A análise do construto de memória semântica poderia ter sido incluída, a julgar pela possível relação do construto com o uso de vocabulário, conforme observado na discussão dos dados, e será realizada para as futuras publicações, uma vez que esses dados foram coletados. Sugere-se, para estudos futuros, replicar os mesmos testes em um número maior de participantes, de outras regiões do país, controlando-se as variáveis de idade e padrão

socioeconômico, com o objetivo de identificar se as tendências estatísticas aqui apontadas podem diferenciar de acordo com a quantidade de respostas e diferenças socioculturais. Também indica-se a correlação dos dados aqui apresentados com dados de volumetria cerebral (VBM) e de tractografia (DTI), cuja coleta encontra-se em fase final pelo nosso grupo de pesquisa. Recomenda-se, igualmente, a incorporação de ferramentas automáticas de análise de textos, assim como de segmentação de sentenças, com o objetivo de refinar ainda mais a funcionalidade desta forma inovadora de tratamento de dados na comparação entre outras populações clínicas e participantes saudáveis.

Entende-se que esta pesquisa contribui para fomentar os estudos na área da psicolinguística cognitiva com participantes falantes de português brasileiro. Ressalta-se, mais uma vez, a importância de se analisar o discurso de populações clínicas (neste caso, na DA), de maneira a ampliar nosso conhecimento das características específicas da produção discursiva, a fim de amparar o diagnóstico precoce dessas patologias. Assim, espera-se que as análises e dados aqui apresentados sirvam para auxiliar no processo de avaliação de discurso em populações clínicas, especialmente de idosos acometidos pela DA, trazendo evidências linguísticas sobre o assunto. Também se compreende que a presente pesquisa intensifica os estudos na área da análise automática de textos, apresentando esta técnica como potencial aliada no diagnóstico precoce de doenças a partir da observância de aspectos linguísticos que costumam dar indícios de comprometimentos cognitivos desde os primeiros estágios de seu desenvolvimento. Espera-se que esta pesquisa seja útil para profissionais da área da linguística, sugerindo novas tarefas para coleta de dados linguísticos, assim como profissionais das áreas da psicologia cognitiva, fonoaudiologia, neurologia e demais profissionais da saúde que têm papel fundamental na busca pela melhor qualidade de vida de pessoas acometidas por demências.

REFERÊNCIAS

ABEP. **Critério de classificação econômica Brasil**, 2014. Disponível em: <<http://www.abep.org/criterio-brasil>>. Acesso em 20/09/2016.

ABRISQUETA-GOMEZ, J. Memória e envelhecimento cognitivo saudável. In: MALLOY-DINIZ, L. F.; FUENTES, D.; COSENZA, R. M. (Orgs.). **Neuropsicologia do Envelhecimento: uma abordagem multidimensional**. Porto Alegre: Artmed, 2013.

ADAM, J-M. **A linguística textual: introdução à análise textual dos discursos**. São Paulo: Cortez, 2008. 373 p.

ALMEIDA, O. P.; ALMEIDA, S. A. Short versions of the geriatric depression scale: a study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. **International Journal of Geriatric Psychiatry**, v. 14, n. 10, p. 858–865, 1999.

ALUÍSIO, S. M.; CUNHA, A. SCARTON, C. Evaluating progression of Alzheimer's disease by regression and classification methods in a narrative language test in Portuguese, In: **12th International Conference on the Computational Processing of Portuguese**. Tomar; 2016, p. 109–114

AMERICAN PSYCHIATRIC ASSOCIATION. **Manual diagnóstico e estatístico de transtornos mentais (DSM-5)**. Trad. Maria Inês Correa Nascimento et. al. Revisão técnica Aristides Volpato Cordioli et. al. Porto Alegre: Artmed, 2014. 948 p.

ARDILA, A., OSTROSKY-SOLIS, F., ROSSELLI, M., GÓMEZ, C. Age-Related Cognitive Decline During Normal Aging: The Complex Effect of Education. **Archives of Clinical Neuropsychology**, v. 15, n.6, p.495-513, 2000. doi: 10.1016/S0887-6177(99)00040-2

BADDELEY, A. **Working Memory**. Oxford: Claridon Press, 1986.

BADDELEY, A. The episodic buffer: a new component of working memory? **Trends of Cognition Sciences**. 2000. v.4, n.11. p. 417-423. Disponível em: <https://ac.els-cdn.com/S1364661300015382/1-s2.0-S1364661300015382-main.pdf?_tid=103e2393-de97-4540-be6a-37ab1ab1ee25&acdnat=1527869377_6d6d3875164ef5cd3829e7313392db85>. Acesso em 01 jun. 2018.

BADDELEY, A. Memória episódica: organizando e lembrando. (2011a) In.: BADDELEY, A; ANDERSON, M. C.; EYSENCK, M. W. **Memória**. Trad. Cornélia Stolting. Porto Alegre: Artmed, 2011. 472 p.

BADDELEY, A. A memória e o envelhecimento. (2011b) In.: BADDELEY, A; ANDERSON, M. C.; EYSENCK, M. W. **Memória**. Trad. Cornélia Stolting. Porto Alegre: Artmed, 2011. 472 p.

BRANDÃO, L; PARENTE, M. A. Compreensão e produção do discurso oral em portadores da doença de Alzheimer. In: ORTIZ, K.Z (Org.) **Distúrbios neurológicos adquiridos: linguagem e cognição**. São Paulo: Manole, 2010.

BUSCHKE, H.; GROBER, H. **Free and Cued Selective Reminding with Immediate Recall (FCSRT+IR)**: FCSRT+IR Instructions for English and Spanish 2.21.12. New York: Albert Einstein College of Medicine, 2012.

CABEZA, R. et al. Hemispheric Asymmetry Reduction in Older Adults: The HAROLD Model. **Psychology and Aging**, v. 17, n. 1, p. 85-100, 2002.

CAPUANO, A. M. N. Alterações de memória e suas correlações com a linguagem. In.: ORTIZ, K. Z. (Org.). **Distúrbios neurológicos adquiridos: linguagem e cognição**. 2a ed. São Paulo: Manole, 2010. 510 p.

CHAVES, M.; IZQUIERDO, I. Differential diagnosis between dementia and depression: A study of efficiency increment. **Acta Neurologica Scandinavica**, v. 85, p. 378-382, 1992.

COSENZA, R. M.; MALLOY-DINIZ, L. F. Declínio cognitivo, plasticidade cerebral e o papel da estimulação cognitiva na maturidade. In: MALLOY-DINIZ, L. F.; FUENTES, D.; COSENZA, R. M. (Orgs.). **Neuropsicologia do Envelhecimento: uma abordagem multidimensional**. Porto Alegre: Artmed, 2013.

CUNHA, A. L. V. **Coh-Matrix-Dementia**: análise automática de distúrbios de linguagem nas demências utilizando Processamento de Línguas Naturais. 2015. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2015. doi:10.11606/D.55.2016.tde-31082016-095132. Acesso em: 03 dez. 2017.

DUBOIS, B.; FELDMAN, H.H.; JACOVA, C.; HAMPEL, H.; MOLINUEVO, J. L.; BLENNOW, K.; DEKOSKY, S. T.; GAUTHIER, S.; SELKOE, D.; BATEMAN, R.; CAPPA, S.; CRUTCH, S.; ENGELBPRGHS, S.; FRISONI, G. B.; FOX, N. C.; GALASKO, D.; HABERT, M.; JICHA, G. A.; NORDBERG, A.; PASQUIER, F.; RABINOVICI, G.; ROBERT, P.; ROWE, C.; SALLOWAY, S.; SARAZIN, M.; EPELBAUM, S.; DE SOUZA, L. C.; VELLAS, B.; VISSER, P. J.; SCHNEIDER, L.; STERN, Y.; SCHELTENS, P.; CUMMINGS, J. L. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. **The Lancet Neurology**, v. 13, n.6, p. 614-629, 2014. DOI: [https://doi.org/10.1016/S1474-4422\(14\)70090-0](https://doi.org/10.1016/S1474-4422(14)70090-0). Acesso em: 09 Dez. 2018.

EYSENCK, M. W. Memória semântica e conhecimento armazenado. In.: BADDELEY, A; ANDERSON, M. C.; EYSENCK, M. W. **Memória**. Trad. Cornélia Stolting. Porto Alegre: Artmed, 2011. 472 p.

FONSECA, R. P.; SALLES, J. F.; PARENTE, M. A. M. P. **Instrumento de Avaliação Neuropsicológica Breve Neupsilin**. São Paulo: Vetor Editora, 2009.

FONSECA, R. P., ZIMMERMANN, N., PAWLOWSKI, J., OLIVEIRA, C.R., GINDRI, G., SCHERER, L. C., RODRIGUES, J. C., PARENTE, M. A. M. P. Métodos em avaliação neuropsicológica. In: LANDEIRA-FERNANDEZ, J; FUKUJIMA, S. (Org.). **Métodos em neurociência**. São Paulo: Manole, 2012. p. 266-296

FRASER, K. C.; MELTZER, J. A.; RUDZICZ, F. Linguistic Features Identify Alzheimer's Disease in Narrative Speech. **Journal of Alzheimer's Disease**, v.49, p.407-422, 2016. DOI: 10.3233/JAD-150520. Acesso em: 13 Dez. 2017.

FRASER, K. C.; MELTZER, J. A.; GRAHAM, N. L.; LEONARD, C.; HIRST, G.; BLACK, S. E.; ROCHON, E. Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex*, v.55, p.43-60, 2012. DOI: <http://dx.doi.org/10.1016/j.cortex.2012.12.006>. Acesso em: 13 dez. 2017.

GOODGLASS, H.; KAPLAN, E. **The Boston Diagnostic Aphasia Examination**. Filadélfia: Lea & Febinger, 1983.

GROBER, E.; BUSCHKE, H.; CRYSTAL, H.; BANG, S.; DRESNER, R. Screening for dementia by memory testing. *Neurology*, v. 38, n.6, p. 900-903, 1988.

GROBER, E.; VEROFF, A. E.; LIPTON, R. B. Temporal unfolding of declining episodic memory on the Free and Cued Selective Reminding Test in the predementia phase of Alzheimer's disease: Implications for clinical trials. *Alzheimer's and Dementia*, v.10, p. 161-171, 2018. DOI: 10.1016/j.dadm.2017.12.004.

HÜBNER, L. C.; SIQUEIRA, L. S.; LOUREIRO, F. S.; SMIDARLE, A. D.; TESSARO, B.; SIQUEIRA, E. C. G.; JERÔNIMO, G. M.; GARCIA, V. R. M. **Bateria de Avaliação da Linguagem no Envelhecimento (BALE)**. In: FONSECA, R. P.; ZIMMERMANN, N. et al. Tarefas de avaliação neuropsicológica para adultos: memórias e linguagem. São Paulo: Memnon. (no prelo)

HÜBNER, L. C.; LOUREIRO, F.; TESSARO, B.; SIQUEIRA, E. C. G.; JERÔNIMO, G. M.; GOMES, I.; SCHILLING, L. P. Naming and verbal learning in adults with Alzheimer's disease, mild cognitive impairment and in healthy aging, with low educational levels. *Arquivos de Neuropsiquiatria*, v.76, n.2, p. 93-99, 2018. DOI: <http://dx.doi.org/10.1590/0004-282x20170190>.

IBGE. Idosos já são 17,3% da população do Rio Grande do Sul. **Zero Hora**, Porto Alegre, 4 dez. 2015. Disponível em: <<http://zh.clicrbs.com.br/rs/vida-e-estilo/noticia/2015/12/idosos-ja-sao-17-3-da-populacao-do-rio-grande-do-sul-4921945.html>>. Acesso em 5 dez. 2017.

IZQUIERDO, I. **Memória**. 2ª ed., rev. e ampl. Porto Alegre: Artmed, 2011. 133 p.

JERÔNIMO, G. M **Produção de narrativas orais no envelhecimento sadio, no comprometimento cognitivo leve e na doença de Alzheimer e sua relação com construtos cognitivos e escolaridade**. 2016. Tese (Doutorado em Linguística_ - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2016. Disponível em: <<http://tede2.pucrs.br/tede2/handle/tede/6592?mode=simple>>. Acesso em: 03 dez. 2017.

JERÔNIMO, G. M.; HÜBNER, L. C. Abordagem neurolinguística do texto narrativo: um enfoque teórico. **Linguagem em (Dis)curso – LemD**, Tubarão, SC, v. 14, n. 2, p. 411-429, maio/ago. 2014.

KAVÉ, G; DASSA, A. Severity of Alzheimer's disease and language features in picture descriptions. *Aphasiology*. v. 32, n.1, p. 27-40, 2018. DOI: <https://doi.org/10.1080/02687038.2017.1303441>.

KINTSCH, W.; VAN DIJK, T.A. Toward a model of text comprehension and production. **Psychological Review**, v. 85, n. 5, p. 363-394, 1978.

LABOV, W.; WALETZKY, J. Narrative analysis: oral versions of personal experience. In: HELM, J. **Essays on the verbal and visual arts**. Washington: University of Washington Press, p.12-44, 1967.

LE BOEUF, C. **Raconte – 55 historiettes en images**. Paris: L'Ecole, 1976.

LEFÈVRE, B. H. W. F.; CAIXETA, L. Semiologia neuropsicológica na avaliação da doença de Alzheimer. In: CAIXETA, L., et. al. **Doença de Alzheimer**. Porto Alegre: Artmed, 2012. 504 p.

LE MOS, R.; DURO, D.; SIMÕES, M. R.; SANTANA, I. The Free and Cued Selective Reminding Test Distinguishes Frontotemporal Dementia From Alzheimer's Disease. **Archives of Clinical Neuropsychology**, v.29, p. 670-679, 2014. DOI: 10.1093/arclin/acu031.

LE MOS, R.; MARÔCO, J.; SIMÕES, M. R.; SANTIAGO, B.; TOMÁS, J.; SANTANA, I. The free and cued selective reminding test for predicting progression to Alzheimer's disease in patients with mild cognitive impairment: A prospective longitudinal study. **Journal of Neuropsychology**, v.11, p.40-55, 2017. DOI: 10.1111/jnp.12075.

LE MOS, R.; MARTINS, C.; SIMÕES, M. R.; SANTANA, I. Estudo de adaptação do Teste de Recordação Selectiva Livre e Guiada para a população portuguesa [Adaptation Study of the Free and Cued Selective Reminding Test for the portuguese population]. **Avaliação Psicológica**. v.11, p.49–61, 2012. Disponível em: < <http://pepsic.bvsalud.org/pdf/avp/v11n1/v11n1a06.pdf>>. Acesso em Dez. 10, 2018.

LIMA, T. M.; BRANDÃO, L.; PARENTE, M. A.; PENA-CASANOVA, J. Doença de Alzheimer: cognição e discurso narrativo com apoio em figuras. **Revista CEFAC**. v.16. n.4. p.1168-1177, 2014. Disponível em: < http://professor.ufrgs.br/lenisabrandao/files/en_1982-0216-rcefac-16-4-1168.pdf>. Acesso em 21 mai. 2018.

MANSUR, L. L. Linguagem. In: MALLOY-DINIZ, L. F.; FUENTES, D.; MATTOS, P.; ABREU, N (orgs.). **Avaliação Neuropsicológica**. Porto Alegre: Artmed, 2010. 423 p.

MARCOTTE, K.; GRAHAM, N. L.; FRASER, K. C.; MELTZER, J. A.; TANG-WAI, D. F.; CHOW, T. W.; FREEDMAN, M.; LEONARD, C.; BLACK, S. E.; ROCHON, E. White Matter Disruption and Connected Speech in non-Fluent and Semantic Variants of Primary Progressive Aphasia. **Dementia and Geriatric Cognitive Disorders**, n. 7, p;52-73, 2017. DOI: 10.1159/000456710. Acesso em: 13 dez. 2017.

MARCUSCHI, L. A. **Produção textual, análise de gêneros e compreensão**. São Paulo: Parábola Editorial, 2008. 295 p.

MONTAÑO, M. B.; RAMOS, L. Validade da versão em português da Clinical Dementia Rating. **Revista de Saúde Pública**, v.39, n.6, p. 912-917, 2005.

MORRIS J.C. The Clinical Dementia Rating (CDR): Current version and scoring rules. **Neurology**, v.43, 2412-2414, 1993.

ORIMAYE, S. O.; WONG, J. S-M.; GOLDEN, K. J.; WONG, C. P; SOYIRI, I. N. Predicting probable Alzheimer's disease using linguistic deficits and biomarkers. **BMC Bioinformatics**, v.18, n.34, p.1-13. DOI: <https://doi.org/10.1186/s12859-016-1456-0>. Acesso em: 13 dez. 2017.

PARENTE, M. A. M. P., FONSECA, R. P., SCHERER, L. C. Literacy as a determining factor for brain organization: from Lecours' contribution to the present day. **Dementia & Neuropsychologia**, v. 2, n.3, p.165-172, 2008. Disponível em: <http://www.demneuropsy.com.br/imageBank/PDF/dnv02n03a04.pdf>. Acesso em: 26 dez. 2017.

PARENTE, M. A. M. P., SCHERER, L. C., ZIMMERMANN, N., FONSECA, R. P. Evidências do papel da escolaridade na organização cerebral. **Revista de Neuropsicologia Latinoamericana**, v. 1, n.1, p.72-80, 2009. Disponível em: <http://pepsic.bvsalud.org/pdf/rnl/v1n1/v1n1a09.pdf>. Acesso em: 26 dez. 2017.

PARK, D. C.; REUTER-LORENZ, P. The Adaptive Brain: Aging and Neurocognitive Scaffolding. *Annual Reviews Psychology*, v.60, p. 173-196, 2009. DOI: doi:10.1146/annurev.psych.59.103006.093656.

PASSOS, V. M. de A., GIATTI, L., BENSENOR, I., TIEMEIER, H., IKRAM, M. A., DE FIGUEIREDO, R. C., BARRETO, S. M. Education plays a greater role than age in cognitive test performance among participants of the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). **BMC Neurology**, v. 15, n.1, 2015. doi: 10.1186/s12883-015-0454-6

de PAULA, J. J.; SILVA, K. K. M.; FUENTES, D.; MALLOY-DINIZ, L. F. Funções executivas e envelhecimento. In.: MALLOY-DINIZ, L. F.; FUENTES, D.; COSENZA, R. M. (Orgs.). **Neuropsicologia do Envelhecimento: uma abordagem multidimensional**. Porto Alegre: Artmed, 2013.

de PAULA, V. J. R.; FORLENZA, O. V. Neurobiologia da doença de Alzheimer e outras demências. In.: CAIXETA, L., et. al. **Doença de Alzheimer**. Porto Alegre: Artmed, 2012. 504 p.

PAWLOWSKI, J., REMOR, E., PARENTE, M. A. M. P., SALLES, J. F., FONSECA, R. F., BANDEIRA, D. R. The influence of reading and writing habits associated with education on the neuropsychological performance of Brazilian adults. **Reading & Writing**, 25, p. 2275-2289, 2012.

RIVERO, T. S.; CANALI-PRADO, F.; VIEIRA, V. L. D.; RIVERO, A. Aspectos psicossociais do envelhecimento. In.: MALLOY-DINIZ, L. F.; FUENTES, D.; COSENZA, R. M. (Orgs.). **Neuropsicologia do Envelhecimento: uma abordagem multidimensional**. Porto Alegre: Artmed, 2013.

REUTER-LORENZ, P. A.; PARK, D.C. How does it STAC UP? Revisiting the Scaffolding Theory of Aging and Cognition. **Neuropsychology Review**, v. 24, p. 355-370, 2014.

SALLES, J. F.; BRANDÃO, L. Linguagem e Comunicação. In: MALLOY-DINIZ, L. F.; FUENTES, D.; COSENZA, R. M. (Orgs.). **Neuropsicologia do Envelhecimento**: uma abordagem multidimensional. Porto Alegre: Artmed, 2013.

SCARTON, C.; ALUÍSIO, S. Análise da Inteligibilidade de textos via ferramentas de Processamento de Língua Natural: adaptando as métricas do Coh-Metrix para o Português. **Linguamática**, 2010; v.2, p.45–62.

SOARES, C. D.; BRANDÃO, L.; LACERDA, M. C. Linguagem e discurso na doença de Alzheimer. In.: CAIXETA, L., et. al. **Doença de Alzheimer**. Porto Alegre: Artmed, 2012. 504 p.

SOARES, V. L. D.; SOARES, C. D.; CAIXETA, L. Métodos de avaliação neuropsicológica no diagnóstico da doença de Alzheimer. In.: CAIXETA, L., et. al. **Doença de Alzheimer**. Porto Alegre: Artmed, 2012. 504 p.

STERN, Y, et al. Exploring the neural basis of cognitive reserve. **Journal of Clinical and Experimental Neuropsychology**, v. 25, p. 691-701, 2003.

STERN, Y. Cognitive reserve in ageing and Alzheimer's disease. **Lancet Neurol**. v. 11, n.11, p. 1006-1012, 2012. doi:10.1016/S1474-4422(12)70191-6.

STELLA, F. Comprometimento Cognitivo Leve. In.: CAIXETA, L., et. al. **Doença de Alzheimer**. Porto Alegre: Artmed, 2012. 504 p.

TEICHMANN, M; EPELBAUM, S.; SAMRI, D.; NOGUEIRA, M. L.; MICHON, A.; HAMPEL, H.; LAMARI, F.; DUBOIS, B. Free and Cued Selective Reminding Test – accuracy for Alzheimer's and neurodegenerative disease differential diagnosis: A large-scale biomarker-characterized monocenter cohort study (ClinAD). **Alzheimer's & Dementia**, v. 13, n.8, p.913-923, 2017. DOI: <https://doi.org/10.1016/j.jalz.2016.12.014>

TOLEDO, C. M.; ALUÍSIO, S. M.; SANTOS, L. B.; BRUCKI, S. M. D.; TRÉS, E. S.; OLIVEIRA, M. O.; MANSUR, L. L. Analysis of macrolinguistic aspects of narratives from individuals with Alzheimer's disease, mild cognitive impairment, and no cognitive impairment. **Alzheimer's and Dementia**, n.10, 31-40, 2017. DOI: <http://dx.doi.org/10.1016/j.dadm.2017.08.005>.

TREVISO, M. V. **Segmentação de sentenças e detecção de disfluências em narrativas transcritas de testes neuropsicológicos**. 2017. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional – Universidade de São Paulo – São Carlos, 2017). Disponível em: <<http://www.teses.usp.br/teses/disponiveis/55/55134/tde-05022018-090740/pt-br.php>>. Acesso em 11 Dez. 2018.

TREVISO, M. V.; SANTOS, L. B.; SHULBY, C.; HÜBNER, L. C.; MANSUR, L. L.; ALUÍSIO, S. M. Detecting mild cognitive impairment in narratives in Brazilian Portuguese: first steps towards a fully automated system. *Letras de Hoje*, v.53, n.1, p. 48-58, jan-mar, 2018. DOI: <http://dx.doi.org/10.14448/1984-7726.2018.1.30955>.

VAN DIJK, T. A. La ciência del texto. Trad. Sibila Hunzinger. 1ª ed. 3ª. reimp. Barcelona: Pidós Comunicación, 1983. 309 p.

VAN DIJK, T. A. **Estructuras y funciones del discurso**: una introducción interdisciplinaria a la lingüística del texto y a los estudios del discurso. Trad. Myra Gann y Martí Mur. 11ª ed. aumentada. México (D. F.): Siglo Veintiuno, 1997. 204 p.

VAN DIJK, T. A. **Cognição, discurso e interação**. Org. e apresentação de Ingedore V. Koch. 6ª ed. 2ª reimp. São Paulo: Contexto, 2010. 207 p.

VAN DIJK, T. A.; KINTSCH, W. *Strategies of Discourse Comprehension*. New York: Academic Press, 1983. 413 p.

VARGAS, R. J. B. **A referenciação na produção de narrativas orais no envelhecimento sadio e na Doença de Alzheimer e sua relação com a escolaridade e o tipo de estímulo**. 2015. Tese (Doutorado em Linguística - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2015). Disponível em: <<http://tede2.pucrs.br:8080/tede2/bitstream/tede/6424/2/476675%20-%20Texto%20Completo.pdf>>. Acesso em: 31 mai. 2018.

ANEXOS

ANEXO A: Termo de Consentimento Livre e Esclarecido (TCLE)

Termo de Consentimento Livre e Esclarecido
Autorização para participar de um projeto de pesquisa

Nome do estudo: **Aspectos semânticos e discursivos no envelhecimento sadio, e na Doença de Alzheimer relacionados à escolaridade: um estudo longitudinal**

Instituição: Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) – Instituto de Letras, Programa de Pós-Graduação em Letras

Pesquisadora responsável: Prof^ª Dr Lilian Cristine Hubner

Telefones para contato: (51) 3320-3676 (secretaria PPGL); (51) 3320-3500 ramal 4606 (gabinete da Prof^ª Dr. Lilian Hübner); CEP da PUCRS 3320-3345.

Nome do participante: _____

1. Objetivo e benefícios do estudo

Analisar o desempenho linguístico no nível semântico e discursivo em populações de adultos idosos saudáveis e diagnosticados com provável DA, considerando-se seu nível de escolaridade, por meio de uma bateria de avaliação de linguagem. Os resultados fornecerão subsídios para uma melhor compreensão do funcionamento dos aspectos acima mencionados, bem como o suporte teórico para futuras técnicas de terapia e de reabilitação da linguagem em sujeitos com esse tipo de lesão.

2. Explicação dos procedimentos

O (a) Senhor(a) será convidado(a) a responder a perguntas e a realizar tarefas que fazem parte deste estudo, algumas destas tarefas serão gravadas e/ou filmadas. Esta aplicação será feita em *três* encontros de no máximo uma hora cada. Sua participação é voluntária. Só responderá a estas avaliações se concordar.

3. Possíveis riscos e desconfortos

O possível desconforto do participante está relacionado ao cansaço ao longo da execução das tarefas.

4. Direito de desistência

O(a) Senhor(a) pode desistir de participar a qualquer momento sem nenhum prejuízo ou consequência.

5. Sigilo

Todas as informações obtidas neste estudo poderão ser publicadas com finalidade científica, preservando-se o completo anonimato dos participantes, os quais serão identificados apenas por um número.

6. Consentimento

Declaro ter lido – ou me foram lidas – as informações acima antes de assinar este formulário. Foi-me dada oportunidade de fazer perguntas, esclarecendo totalmente as minhas dúvidas. Por este documento, tomo parte, voluntariamente, deste estudo.

Porto Alegre, _____ de _____ de 20____.

Assinatura do participante

Assinatura da testemunha

Assinatura do pesquisador responsável

Comitê de Ética em Pesquisa (CEP)
Av. Ipiranga 6681, Prédio 50 - Sala 703
Porto Alegre/RS - Brasil - CEP 90619-900
Telefone: Fone/Fax: [\(51\) 3320.3345](tel:5133203345)
E-mail: cep@pucrs.br
Horário de funcionamento: Segunda a sexta-
feira, das 08h às 12h e das 13h30 às 17h