



SAÚDE

ESCOLA SUPERIOR
POLITÉCNICO SETÚBAL

SUSANA ISABEL
CENTEIO
MOREIRA DE
SOUSA

**ANÁLISE MICROLINGUÍSTICA DA
FIGURA “O ENCANTADOR DE
HISTÓRIAS”
UM CONTRIBUTO PARA
OBTENÇÃO DE DADOS
NORMATIVOS**

MICROLINGUISTIC ANALYSIS OF
THE PICTURE 'THE STORY
ENCHANTER': A CONTRIBUTION TO
THE ESTABLISHMENT OF
NORMATIVE DATA

Relatório de Projeto de Investigação do Mestrado
em Terapia da Fala

ORIENTADOR

Professor Doutor José Fonseca

COORIENTADOR

Professora Doutora Ana Castro

Julho de 2025

SUSANA ISABEL
CENTEIO
MOREIRA DE
SOUSA

**ANÁLISE MICROLINGUÍSTICA DA
FIGURA “O ENCANTADOR DE
HISTÓRIAS”
UM CONTRIBUTO PARA
OBTENÇÃO DE DADOS
NORMATIVOS**

MICROLINGUISTIC ANALYSIS OF
THE PICTURE 'THE STORY
ENCHANTER': A CONTRIBUTION TO
THE ESTABLISHMENT OF
NORMATIVE DATA

JÚRI

Presidente: Professora Doutora Dina Alves, Professora
Coordenadora ESS - IPS

Arguente: Professora Doutora Inês Sousa Lopes,
Professora Adjunta da ESS- IPL

Orientador: Professor Doutor José Fonseca, Adjunto
convidado ESS - IPS

Data da realização da prova: Julho de 2025

Relatório de investigação apresentado para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Terapia da Fala, realizada sob a orientação científica do Professor Doutor José Fonseca e coorientação científica da Doutora Ana Castro.

DECLARAÇÃO

O presente Relatório de Investigação descreve um estudo original, realizado sob a orientação científica do Professor Doutor José Fonseca e coorientação da Professora Doutora Ana Castro, e cumpre os requisitos necessários à obtenção do grau de Mestre em Terapia da Fala, outorgado pela Escola Superior de Saúde do Instituto Politécnico de Setúbal (ESS/IPS).

A redação do presente documento segue as normas da revista científica *Aphasiology* (cf. <https://www.tandfonline.com/journals/paph20>), sendo, nos casos omissos, complementada pelas normas de redação definidas no Manual do Mestrado em Terapia da Fala da ESS/IPS.

Apesar de adotar o formato exigido pela referida revista, o texto encontra-se integralmente redigido em português.

Setúbal 9 de junho de 2025

O Candidato:

Assinado por: **Susana Isabel Centeio Moreira de Sousa**
Num. de Identificação: 09568345
Data: 2025.06.10 08:29:53+01'00'

Os Orientadores

Assinado por: **JOSÉ MANUEL BORGES FONSECA**
Num. de Identificação: 06567141
Data: 2025.06.10 08:58:10+01'00'

AGRADECIMENTOS

Ao Professor Doutor José Fonseca expresso o mais profundo agradecimento pelo desafio, pelo incentivo, pela orientação rigorosa e pela disponibilidade constante com que me acompanhou em cada etapa deste trabalho.

À Professora Doutora Ana Castro, minha coorientadora, agradeço a atenção e o acompanhamento prestados ao longo do desenvolvimento deste estudo.

Gostaria também de manifestar o meu sincero agradecimento à Professora Dina Alves, coordenadora do Mestrado, pela sua constante disponibilidade, atenção e proximidade. A sua presença atenciosa e o seu compromisso com os alunos foram determinantes não apenas para a concretização deste trabalho, mas para a experiência formativa global que o mestrado representou para mim.

À minha colega e amiga Rita, obrigada por seres sempre uma querida, presente, atenta e parceira de verdade. O teu apoio, companheirismo e incentivo tornaram este percurso não só possível, mas também mais leve e cheio de significado.

À minha amiga e sócia Isabel, obrigada por me trazeres equilíbrio quando a balança teimava em pender para o lado errado.

À minha família, um agradecimento especial pelo amor, paciência e compreensão por tudo em que falhei. Aos meus filhos, por me lembrarem todos os dias do que é verdadeiramente importante. Ao meu querido marido, pelo carinho e apoio incondicional, mesmo nos momentos de maior exigência.

RESUMO

Este estudo teve como objetivo analisar o desempenho microlinguístico de adultos saudáveis numa tarefa de discurso, examinando o impacto do sexo, da idade e da escolaridade na fluência, na complexidade sintática e na diversidade lexical. Participaram 400 adultos, agrupados por idade (18–39, 40–64 e ≥ 65 anos) e por nível de escolaridade (4–12 anos e >12 anos).

Os participantes realizaram uma tarefa de descrição de imagem com base na figura “O Encantador de Histórias”. Foram recolhidas medidas de fluência, complexidade sintática e diversidade lexical, analisadas através de estatísticas descritivas e inferenciais, incluindo comparações entre grupos e modelos preditivos.

A escolaridade revelou-se um preditor significativo em todas as variáveis linguísticas. Os participantes com mais de 12 anos de escolaridade produziram discursos mais complexos e lexicalmente mais diversos. Curiosamente, os indivíduos com 65 anos ou mais apresentaram melhor desempenho em termos de complexidade sintática e diversidade lexical, mesmo após controlo da escolaridade.

Estes resultados evidenciam a influência de fatores sociocognitivos no desempenho linguístico e sustentam a utilização de tarefas narrativas, como a descrição da figura “O Encantador de Histórias”, na avaliação da linguagem em adultos.

Palavras-chave: microlinguística, tarefa narrativa, complexidade sintática, diversidade lexical, adultos saudáveis

ABSTRACT

This study aimed to analyze the microlinguistic performance of healthy adults in a discourse task, examining the impact of age and education on fluency, syntactic complexity, and lexical diversity. A total of 400 adults participated, grouped by age (18–39, 40–64, and ≥ 65 years) and education level (4–12 years and >12 years).

Participants completed a picture description task based on “The Storyteller”. Measures of fluency, syntactic complexity, and lexical diversity were collected and analyzed using descriptive and inferential statistics, including group comparisons and predictive models.

Education level emerged as a significant predictor across all linguistic variables. Participants with more than 12 years of education produced more complex and lexically diverse discourse. Interestingly, individuals aged 65 or older demonstrated higher syntactic complexity and lexical diversity, even after controlling for education.

These findings highlight the influence of sociocognitive factors on language performance and support the use of narrative tasks in adult language assessment.

Keywords: microlinguistics, narrative task, syntactic complexity, lexical diversity, healthy adults

ÍNDICE

INTRODUÇÃO	10
Linguagem	11
Microlinguística	14
Influência da idade, escolaridade, sexo e lateralidade na capacidade linguística.....	14
Alterações microlinguísticas no discurso afásico	21
Questões de Investigação/Objetivos.....	24
Questões de investigação	24
Objetivos	24
METODOLOGIA.....	25
Transcrição e segmentação	27
Procedimento de análise	30
Instrumentos para análise de discurso	36
Fiabilidade	37
Fiabilidade na divisão de enunciados.....	37
Fiabilidade na análise microlinguística	40
RESULTADOS	43
CONCLUSÃO	48
BIBLIOGRAFIA	50
ANEXOS	64
ANEXO A - Figura “O Encantador de Histórias”	65
ANEXO B - Critérios de Transcrição	66
APÊNDICES	67
APÊNDICE A - Fiabilidade.....	68
APÊNDICE B - Análise de dados.....	69

Lista de Abreviaturas e Siglas

BAAL-R – Bateria de Avaliação da Afasia de Lisboa – Revisada

BDAE – Boston Diagnostic Aphasia Examination

CIU – Correct Information Unit (Unidade de Informação Correta)

CoreLex – Core Lexicon (Léxico Nuclear)

COSMIN – Consensus-based Standards for the selection of health Measurement INstruments

DL – Diversidade Lexical

ELAN – EUDICO Linguistic Annotator

ICC – Intraclass Correlation Coefficient (Coeficiente de Correlação Intraclass)

MLU – Mean Length of Utterance (Comprimento Médio do Enunciado)

MMSE – Mini-Mental State Examination

MTLD – Measure of Textual Lexical Diversity

SCD – Subjective Cognitive Decline (Declínio Cognitivo Subjetivo)

SPSS – Statistical Package for the Social Sciences

TTR – Type-Token Ratio

T-Unit – Terminable Unit (Unidade Terminável)

WAB-R – Western Aphasia Battery – Revised

WPM – Words Per Minute (Palavras por Minuto)

INTRODUÇÃO

A linguagem é uma capacidade humana complexa que permite aos indivíduos comunicar, expressar pensamentos e interpretar o mundo que os rodeia. O discurso espontâneo, enquanto manifestação funcional da linguagem, constitui um dos domínios mais sensíveis à avaliação clínica e neuropsicológica, refletindo de forma direta as competências linguísticas e cognitivas do indivíduo.

A avaliação do discurso, especialmente em tarefas como a descrição de imagens, tem ganho destaque na prática clínica por proporcionar amostras controladas, mas ecológicas, que permitem analisar a estrutura do discurso em diferentes níveis, desde a seleção lexical até à complexidade sintática. Estas tarefas são amplamente utilizadas em contextos de avaliação de perturbações adquiridas da linguagem, como a afasia, e estão incluídas em baterias reconhecidas, como o Boston Diagnostic Aphasia Examination (BDAE) (Goodglass et al., 2001) e a Western Aphasia Battery-Revised (WAB-R) (Kertesz, 2006).

Apesar da ampla utilização de tarefas de descrição de imagem em contextos de avaliação linguística, continua a ser importante reunir dados normativos em Português Europeu, que considerem a diversidade da população adulta. A tarefa “O encantador de histórias”, prevista para integrar a versão revista da Bateria de Avaliação da Afasia de Lisboa (BAAL-R), surge neste contexto como uma proposta relevante.

A recolha de dados em falantes saudáveis visa apoiar a construção de critérios de referência, contribuindo para a adequação e sensibilidade dos instrumentos de avaliação disponíveis. A análise microlinguística, com especial enfoque na diversidade lexical e na complexidade sintática, constitui uma abordagem metodológica rigorosa e informativa. Através da segmentação do discurso em Unidades Termináveis (T-Units) e da identificação de estruturas coordenadas e subordinadas, torna-se possível caracterizar com maior precisão o

desempenho linguístico, bem como explorar o impacto de variáveis sociodemográficas como a idade, a escolaridade e o sexo.

Paralelamente, a literatura tem vindo a evidenciar que fatores como a escolaridade influenciam significativamente o desempenho linguístico, favorecendo a produção de frases mais complexas e de vocabulário mais diversificado (Kim & Kintz, 2024; Nippold et al., 2014). Já a idade pode estar associada a alterações na fluência e na estrutura do discurso, embora com efeitos atenuados pela reserva cognitiva e pelas práticas culturais e educacionais (Roll, 2015; Scott, 1988).

O presente estudo tem, assim, como objetivo principal descrever os padrões microlinguísticos do discurso descritivo em adultos saudáveis falantes de português europeu, com base na tarefa de descrição da imagem “O encantador de histórias”, e verificar a influência da idade, da escolaridade e do sexo nos parâmetros analisados.

Linguagem

A linguagem é concebida como um sistema semiótico através do qual os indivíduos constroem e interpretam significados, articulando as suas experiências do mundo exterior e interior (M. A. Halliday, 1978; Vygotsky, 1984).

Evidências provenientes da neuroimagem funcional e da eletrofisiologia demonstram que a linguagem resulta da ativação coordenada de redes cerebrais amplas e interconectadas. Para além das tradicionais áreas peri-sílvicas, estão envolvidas regiões do córtex pré-frontal, temporal anterior, parietal inferior, córtex cingulado e estruturas subcorticais. Estas redes não se limitam aos componentes linguísticos, como a fonologia, morfossintaxe e semântica, mas integram também outras funções cognitivas, como a memória de trabalho, o controlo executivo, a atenção e a cognição social (Hickok & Poeppel, 2007; Poeppel, 2014). Segundo Hagoort (2013), o processamento da linguagem afasta-se da visão clássica baseada numa divisão rígida entre as áreas de Broca e Wernicke, sendo suportado por uma rede distribuída

que interage com sistemas de memória, controlo executivo e cognição social. Esta perspetiva é partilhada por Fedorenko & Thompson-Schill, (2014), que propõem que o cérebro não opera com módulos isolados para cada componente da linguagem, mas sim com redes flexíveis e adaptativas, ajustadas às exigências comunicativas e cognitivas do momento.

Neste contexto Friederici et al. (2017), destaca o papel da interação entre redes especializadas, predominantemente no hemisfério esquerdo, e áreas associadas à cognição, como o córtex pré-frontal e o cíngulo. Estudos mais recentes, como o de (Sun & Moens, 2023), mostram que o cérebro processa a linguagem de forma distribuída e simultânea em vários níveis, fonológico, semântico e sintático, evidenciado pela correspondência entre padrões de ativação cerebral e representações de modelos computacionais de linguagem.

Zhang et al (2024), demonstram que o processamento linguístico envolve redes dinâmicas, sensíveis ao tipo de estímulo (lexical, frásico ou discursivo), enquanto. Ye et al. (2025), reforçam a natureza integrada do sistema linguístico, mostrando que o cérebro articula informação conceptual, emocional e sintática mesmo na ausência de linguagem explícita, quando por exemplo alguém observa uma imagem e, sem ouvir ou ler palavras, começa a evocar mentalmente frases para a descrever.

O uso da linguagem em discurso, decorre do conhecimento da língua, mas também da capacidade que cada interlocutor tem para pôr em prática esses conhecimentos, envolvendo a ativação e interação de múltiplos subsistemas cognitivos e linguísticos (Faria, 2003; Harris Wright & Capilouto, 2012).

As características linguísticas de uma população são influenciadas por fatores sociais, culturais e individuais. O comportamento linguístico de um grupo, num determinado teste, reflete de forma básica essas características (Radanovic et al., 2004).

A descrição de imagens é um método frequentemente utilizado para provocar discurso narrativo em estudos de linguagem, permitindo elicitar amostras de discurso estruturadas e

comparáveis entre indivíduos (Bryant et al., 2016; Cannizzaro & Coelho, 2013). Embora esta tarefa possa ser considerada artificial, apresenta vantagens significativas, tanto em investigação como na prática clínica, ao proporcionar um estímulo visual padronizado que orienta o discurso e minimiza variações temáticas (Cooper, 1990). Como salientam Boucher et al. (2022), a descrição de imagens, como a cena do piquenique da WAB-R (Kertesz, 2006) permite analisar de forma detalhada variáveis micro e macrolinguísticas, incluindo produtividade verbal, diversidade lexical, complexidade sintática, erros de fala, informatividade e eficiência comunicativa. Para além disso, esta abordagem facilita a recolha de dados normativos essenciais à comparação com populações clínicas, como indivíduos com afasia ou defeito cognitivo ligeiro (Kim & Kintz, 2024). De facto, a obtenção de amostras de discurso em tarefas de descrição de imagens contribui para o desenvolvimento de medidas de referência cultural e linguisticamente ajustadas, que são fundamentais para a avaliação de perturbações adquiridas da linguagem, especialmente em contextos de envelhecimento (Boucher et al., 2022).

A análise do discurso conversacional em afasia tem sido amplamente estudada através de bancos de dados como o AphasiaBank, que recolheu amostras padronizadas do discurso de indivíduos com afasia e neurotípicos. A análise desses discursos utiliza medidas como densidade proposicional, unidades de informação corretas (Correct Information Units) (CIUs) e léxico central (CoreLex), que permitem quantificar o impacto das alterações linguísticas na comunicação funcional (MacWhinney & Fromm, 2016).

Microlinguística

A microlinguística é uma abordagem analítica centrada na estrutura interna da língua, focando-se em unidades linguísticas como fonemas, morfemas, palavras e sintagmas, bem como nas suas relações e funções no discurso (Crystal, 2008; Fromkin et al., 2013).

No âmbito da análise do discurso, a microlinguística investiga os aspetos formais da produção linguística, incidindo em fenómenos como a morfossintaxe, o léxico, a fluência verbal e os erros linguísticos, incluindo disfluências, omissões e substituições (Coulthard, 2014). Esta abordagem distingue-se da macrolinguística, que examina o funcionamento global do discurso, abrangendo a coesão textual, a organização enunciativa, e os aspetos pragmáticos e contextuais da comunicação (Armstrong, 2000; Sherratt, 2007). A análise conjunta destes dois níveis permite uma compreensão mais completa das competências linguísticas dos indivíduos, especialmente em contextos de perturbações adquiridas da linguagem (Armstrong, 2000; Sherratt, 2007).

Dentro da análise do discurso, a microlinguística abrange medidas como a diversidade lexical, a complexidade sintática e a produção de unidades de informação corretas (CIUs), sendo estas variáveis essenciais para compreender a eficiência comunicativa dos falantes (Kim & Kintz, 2024)

A aplicação da microlinguística na avaliação da linguagem tem sido relevante em diversas áreas, incluindo a investigação clínica e a recolha de dados normativos para a criação de testes diagnósticos destinados à identificação de perturbações da linguagem, como a afasia (Paradis, 2004)

Influencia da idade, escolaridade, sexo e lateralidade na capacidade linguística

As capacidades linguísticas são influenciadas por múltiplos fatores, entre os quais o sexo, a idade, a escolaridade e a lateralidade. Estudos demonstram que esses fatores afetam diferentes domínios da linguagem, incluindo a fluência verbal, a complexidade sintática, a

diversidade lexical e o conteúdo informativo do discurso (Kim & Kintz, 2024; Pereira & Ortiz, 2022).

O envelhecimento impacta a produção do discurso de forma diferenciada. A complexidade sintática tem sido estudada como um possível marcador linguístico do envelhecimento e de patologias cognitivas (Kim & Kintz, 2024). Embora a memória semântica e o vocabulário se mantenham relativamente preservados, verifica-se uma redução no conteúdo informativo e uma maior dificuldade no acesso lexical com o envelhecimento (Pistono et al., 2017). Adultos mais velhos tendem a produzir menos Unidades de Informação Correta (UIC) e a apresentar um declínio na Diversidade Lexical (DL) ao longo da vida (Capilouto et al., 2016; Kim & Kintz, 2024).

Segundo Rofes et al. (2020), a complexidade sintática tende a diminuir com a idade, com uma progressiva redução do uso de estruturas subordinadas e um aumento de pausas e hesitações. No entanto, a produção de frases passivas e outras construções sintáticas mais exigentes do ponto de vista cognitivo sofre um impacto ainda mais acentuado em indivíduos com defeito cognitivo ligeiro, o que sugere que a sintaxe pode funcionar como um marcador precoce de declínio cognitivo (Marini et al., 2005; Stark, Alexander, et al., 2023). Embora se observem estas alterações na produção sintática e lexical, alguns aspetos da linguagem permanecem relativamente preservados. Por exemplo, adultos mais velhos tendem a produzir discursos mais organizados e coesos do que adultos mais jovens, possivelmente recorrendo a estratégias compensatórias para mitigar dificuldades no acesso lexical (Andreetta et al., 2012).

Esta capacidade de manter uma estrutura discursiva coesa, apesar das limitações microlinguísticas, poderá refletir mecanismos adaptativos de reorganização linguística (Armstrong, 2000).

Vários estudos têm demonstrado que a escolaridade exerce uma influência significativa no desempenho linguístico, afetando desde tarefas básicas de compreensão até

produções discursivas mais complexas. (Radanovic et al., 2004), destacam que indivíduos com menor escolaridade tendem a apresentar dificuldades em múltiplas dimensões da linguagem, incluindo compreensão, fluência e produção sintática. De forma complementar, Pereira & Ortiz (2022), sublinham que essa influência se manifesta particularmente em tarefas como nomeação, repetição e compreensão de frases complexas, podendo inclusive comprometer a validade das avaliações neuropsicológicas quando não se consideram os níveis educacionais. Estes autores observam ainda uma menor diversidade lexical e maior dificuldade na produção fluente por parte dos indivíduos com menos anos de escolaridade. Em consonância com estes achados, (Bayot & Gonçalves, 2016), reportaram que, tanto a idade como a escolaridade têm um impacto estatisticamente significativo na memória semântica e na evocação lexical.

A exposição contínua à leitura e à escrita ao longo da vida pode atenuar o impacto do envelhecimento nas capacidades linguísticas e cognitivas, funcionando como um fator de proteção contra o declínio cognitivo. Indivíduos com maior escolaridade e maior exposição à linguagem escrita apresentam melhor desempenho em tarefas de fluência verbal, evocação lexical e compreensão sintática, mesmo em idades avançadas (Bialystok, 2007; Malcorra et al., 2021; Pawlowski et al., 2012). O envolvimento em práticas de leitura e escrita parece contribuir para o fortalecimento da capacidade cognitiva, permitindo que o cérebro mantenha funções linguísticas e executivas por períodos mais longos, independentemente do nível educacional inicial (Branco et al., 2014; Tessaro et al., 2020).

A hipótese da reserva cognitiva sugere que atividades intelectualmente estimulantes, como a leitura e a escrita, podem criar redes neurais mais eficientes e resistentes ao declínio cognitivo (Snitz et al., 2009). Indivíduos que mantêm o hábito da leitura e da escrita não só apresentam menor impacto do envelhecimento na linguagem, como também demonstram

maior capacidade de adaptação a alterações cerebrais, reforçando a ideia de que a estimulação contínua pode mitigar efeitos neurodegenerativos (Cooper, 1990).

.Estudos sobre discurso normal ao longo da idade mostram que, embora o envelhecimento traga desafios na recuperação lexical, o contacto frequente com a leitura e a escrita ajuda a manter a diversidade lexical e a precisão do discurso (Cooper, 1990).

O estudo de Kim & Kintz (2024), destaca que, apesar de haver um declínio em unidades de informação corretas (CIUs) ao longo da vida, indivíduos mais escolarizados conseguem manter níveis mais altos de informatividade no discurso.

A adaptação de instrumentos de avaliação neuropsicológica à realidade educacional dos indivíduos é fundamental para garantir a precisão dos resultados e evitar vieses na interpretação do desempenho cognitivo (Goodglass et al., 2001; Radanovic et al., 2004). É por isso, essencial a utilização de testes que consideram a escolaridade na sua normatização, como o BDAE (Goodglass et al., 2001), amplamente reconhecido por incluir critérios de referência ajustados ao nível de escolaridade dos indivíduos.

(Rabinovich et al., 2020), referem diferenças entre os sexos na produção linguística, sobretudo ao nível da escolha lexical e da carga emocional associada ao discurso. Estes autores observaram que, embora homens e mulheres utilizem linguagem figurativa com maior carga emocional do que a linguagem literal, existem diferenças subtis na dimensão de dominância, ou seja, no grau de controlo ou poder transmitido pelas expressões utilizadas. Hossain e Samin (2022), ao analisarem discursos públicos, verificaram que oradores masculinos tendem a usar com maior frequência termos ligados a categorias linguísticas, psicológicas e sociais do que as oradoras femininas, o que sugere possíveis diferenças nas escolhas lexicais. No entanto, quando se analisam tarefas estruturadas, como a descrição de imagens, a investigação tem-se centrado essencialmente na influência da idade e da

escolaridade, sem considerar o sexo como variável de análise (Boucher et al., 2022; Kim & Kintz, 2024)

A lateralidade cerebral, entendida como a especialização funcional de um dos hemisférios cerebrais, constitui um fator relevante na investigação das capacidades linguísticas. Em indivíduos destros, a linguagem é predominantemente processada no hemisfério esquerdo, embora existam variações individuais na dominância hemisférica (MacWhinney & Fromm, 2016). Estas diferenças na lateralização podem influenciar o desempenho em tarefas linguísticas, tanto ao nível da produção como da compreensão, e assumem particular importância em populações clínicas, como as pessoas com afasia, em que a reorganização funcional do cérebro pode condicionar a recuperação das competências linguísticas (MacWhinney & Fromm, 2016).

No contexto do envelhecimento, a literatura aponta para uma maior ativação do hemisfério direito como mecanismo compensatório face à diminuição da eficiência linguística do hemisfério esquerdo, fenómeno conhecido como hipótese da compensação hemisférica (Kim & Kintz, 2024). Este mecanismo pode ser vantajoso para a manutenção das capacidades linguísticas, mas está, por vezes, associado a uma redução da especificidade lexical e da complexidade sintática (Cooper, 1990; Kim & Kintz, 2024).

Relativamente a populações clínicas, alguns autores sugerem que a lateralização da linguagem pode influenciar a recuperação após lesões cerebrais. Indivíduos com lateralização atípica, como alguns canhotos ou ambidestros, poderiam apresentar maior plasticidade neural e, consequentemente, um potencial acrescido de reorganização funcional das áreas linguísticas, o que poderia facilitar a recuperação em casos de afasia (Knecht, 2000; MacWhinney & Fromm, 2016; Szaflarski et al., 2002). No entanto, esta relação não é consensual, uma vez que outros autores não encontram diferenças significativas na recuperação da linguagem em função da lateralização, sugerindo que outros fatores, como a

extensão da lesão ou a idade, poderão ter maior impacto (Karbe et al., 1998; Kertesz et al., 1979; Tivarus et al., 2012). Binder et al. (2011), num estudo de neuroimagem evidenciaram que a lateralização atípica está associada a uma maior ativação bilateral durante tarefas linguísticas, o que, embora favoreça alguma compensação, pode resultar em produções linguísticas menos eficientes e menos específicas

Como é amplamente reconhecido, a linguagem é predominantemente processada no hemisfério esquerdo na maioria dos indivíduos destros (cerca de 90% a 95%), enquanto pessoas canhotas ou ambidexas apresentam maior variabilidade na lateralização, com aproximadamente 70% a processarem a linguagem no hemisfério esquerdo, 15% no hemisfério direito e os restantes 15% a apresentarem uma distribuição bilateral ou mais equilibrada (MacWhinney & Fromm, 2016). Esta variabilidade pode ter implicações relevantes no desempenho linguístico, especialmente em situações de lesão cerebral. Assim, indivíduos cuja lateralização da linguagem envolve ambos os hemisférios de forma mais equilibrada parecem demonstrar maior flexibilidade cognitiva na reorganização funcional da linguagem, o que poderá constituir uma vantagem em casos de defeito cognitivo ligeiro. No entanto, esta reorganização neural nem sempre se traduz em melhor desempenho, podendo associar-se a dificuldades na fluência verbal e na precisão lexical (Song et al., 2022).

Por outro lado, indivíduos com alta escolaridade e lateralização atípica podem demonstrar melhores mecanismos de compensação, o que reduz o impacto do envelhecimento ou de patologias neurológicas sobre a linguagem (Pereira & Ortiz, 2022).

A complexidade sintática refere-se à variedade e sofisticação das estruturas gramaticais utilizadas por um falante ou escritor. É um indicador central de maturidade linguística e tem sido amplamente estudada tanto em contextos de desenvolvimento como na idade adulta (Berman & Verhoeven, 2002; Scott, 1988a).

O conceito de T-Unit foi introduzido por (Hunt, 1970), sendo definido como uma oração principal com todas as suas dependências subordinadas. Esta unidade tem sido amplamente utilizada na análise de discurso espontâneo e escrito, por permitir uma quantificação objetiva da extensão e complexidade das produções linguísticas (Lu, 2011).

Além do número total de T-Units, investiga-se frequentemente a proporção de T-Units que incluem coordenação, subordinação ou ambas (estruturas mistas), o que oferece uma medida de profundidade sintática (Lu, 2011; Nippold et al., 2007).

Outras métricas frequentemente associadas incluem: o número médio de orações por T-Unit (densidade clausal), o comprimento médio das T-Units (palavras por T-Unit) e a diversidade de subordinadas (completivas, adverbiais, relativas). Estas medidas são consideradas sensíveis a variações individuais e contextuais (Beers & Nagy, 2009; Biber & Gray, 2011).

A escolaridade surge de forma consistente na literatura como um dos preditores mais fortes de complexidade sintática. Indivíduos com níveis mais elevados de escolaridade tendem a produzir mais T-Units, de maior comprimento e com estruturas mais complexas, tanto em tarefas orais como escritas (Nippold et al., 2005; Reno & McMaster, 2024).

Em falantes adultos, a diversidade sintática e lexical tende a aumentar com o nível de escolaridade, refletindo-se numa maior variedade de frases subordinadas e numa organização discursiva mais elaborada (Malcorra et al., 2022; Branco et al., 2014).

Ao longo do desenvolvimento, a complexidade sintática aumenta da infância até ao início da idade adulta. Em adultos saudáveis, porém, estudos mostram uma estabilidade relativa da sintaxe ao longo da vida adulta (Nippold et al., 2007). Há evidência de um declínio ligeiro nos índices de fluência e sin-taxa em idade avançada, como aumento de disfluências (Bortfeld et al., 2001), embora esse efeito seja frequentemente mitigado por variáveis como escolaridade e reserva cognitiva.

A faixa etária dos 45–59 anos tem sido descrita como um período de máximo desempenho em tarefas linguísticas exigentes (Nippold et al., 2014)

A complexidade sintática varia substancialmente consoante o tipo de tarefa. Tarefas narrativas ou descritivas (como a descrição de uma imagem) tendem a elicitar mais estruturas subordinadas do que tarefas conversacionais ou respostas curtas (Scott & Windsor, 2000). Tarefas cognitivamente exigentes, como justificação ou comparação, estão associadas a maior densidade sintática (Rioux & Thordardottir, 2023).

Alterações microlinguísticas no discurso afásico

As alterações microlinguísticas do discurso afásico correspondem a perturbações que ocorrem a nível lexical, morfossintático e fonológico, afetando a construção interna dos enunciados. Estas alterações resultam em produções discursivas marcadas por dificuldade de acesso lexical, simplificação sintática e erros morfológicos (Armstrong, 2000; Bryant et al., 2016). A relação entre déficits microlinguísticos e dificuldades macrolinguísticas é amplamente reconhecida, pois as dificuldades na formulação de frases comprometem a coerência global do discurso (Andreetta & Marini, 2015).

A fluência verbal em indivíduos com afasia pode estar comprometida de diferentes maneiras. Pacientes com afasia de Broca apresentam discurso telegráfico, com pausas prolongadas e dificuldade na formulação de frases completas, enquanto indivíduos com afasia de Wernicke podem produzir discurso fluente, mas desorganizado e com parafasias frequentes (Gordon, 2020). O estudo de (Bryant et al., 2019), indica que a análise de discurso pode ser usada para avaliar a qualidade da produção verbal, diferenciando os tipos de afasia com base na fluência e coesão textual.

As dificuldades de acesso lexical são um dos aspetos mais destacados na produção microlinguística de indivíduos com afasia, manifestando-se por anomias (falhas na evocação de palavras) e parafasias (substituições inapropriadas de palavras). Tais dificuldades refletem

uma redução da densidade lexical e da variedade de palavras utilizadas (Andreetta et al., 2012; Boyle, 2011). A riqueza lexical, medida por índices como o type-token ratio (TTR), encontra-se frequentemente diminuída, indicando menor diversidade de vocabulário (Stark et al., 2025).

Além das anomias, as parafasias semânticas (uso de palavras semanticamente relacionadas, mas incorretas) e as parafasias fonológicas (trocas de sons dentro da palavra) são frequentes e comprometem a precisão do discurso (Andreetta et al., 2012; Bryant et al., 2016). A ocorrência destas parafasias prejudica a inteligibilidade e a clareza do discurso, gerando produções confusas e desorganizadas a nível microlinguístico (Boyle, 2011).

Indivíduos com afasia apresentam também dificuldades ao nível da morfologia, evidenciadas por omissões ou substituições de morfemas flexionais, como marcadores de tempo, número e género (Boyle, 2011; Bryant et al., 2016). Estas falhas contribuem para a produção de frases agramaticais, e afetam a correção e a estrutura formal do discurso (Thompson & Bastiaanse, 2012).

O débito verbal também é significativamente afetado, verificando-se pausas frequentes, hesitações, autocorreções, reformulações e prolongamentos (Bryant et al., 2016; Stark, Bryant, et al., 2023a). Estas características contribuem para um discurso com ritmo reduzido e baixa velocidade de produção de palavras por minuto (WPM), comprometendo a naturalidade e a eficiência comunicativa (Berube et al., 2019). A presença de silêncios prolongados e interrupções reflete também o esforço cognitivo para o planeamento e a produção linguística (Armstrong, 2000).

Estudos evidenciam, nas afasias fluentes, a diminuição de utilização de palavras de conteúdo (substantivos, verbos, adjetivos), com um aumento proporcional de palavras de função (pronomes, preposições), refletindo um empobrecimento do conteúdo informacional (Stark et al., 2023). A razão nome/verbo e a proporção de palavras de conteúdo vs. palavras

funcionais são medidas microlinguísticas relevantes que evidenciam essas alterações (Berube et al., 2019).

Entre as manifestações microlinguísticas mais relevantes encontram-se as alterações de natureza sintática, com impacto direto na estruturação das frases e no encadeamento discursivo. De acordo com (Agmon et al., 2025), e (Alyahya, 2024), pessoas com afasia produzem discursos com uma significativa redução da complexidade sintática, evidenciada pela menor frequência de estruturas subordinadas e maior uso de frases simples ou com coordenação rudimentar. Estes padrões são particularmente marcados na afasia não fluente, onde a dificuldade na produção de morfemas funcionais e construção de dependências sintáticas complexas é acentuada.

Uma das alterações mais recorrentes no discurso afásico não fluente é a simplificação da estrutura sintática, traduzida em enunciados curtos, predominantemente simples, com frequente omissão de elementos obrigatórios, como pronomes, artigos e marcadores de tempo verbal (Andreetta et al., 2012; Stark et al., 2023). As pessoas com afasia não fluente apresentam uma redução significativa no comprimento médio dos enunciados (mean length of utterance - MLU), o que limita a quantidade de informação comunicada em cada enunciado (Bryant et al., 2016). Existe também uma redução na utilização de orações subordinadas, o que empobrece o discurso em termos de complexidade sintática e limita a expressão de relações lógicas e temporais (Stark et al., 2023).

Existem diferenças nos perfis de produção sintática em variantes clínicas de afasia. Na afasia agramática, observa-se uma drástica redução das T-Units e um uso muito limitado de subordinação. Em contraste, pessoas com afasia fluente, como a variante semântica ou logopénica, podem produzir frases mais longas, mas com erros morfossintáticos, trocas de palavras ou discurso vazio (Agmon et al., 2025; Pracar et al., 2025).

A fluência verbal em indivíduos com afasia pode estar comprometida de diferentes maneiras. Pacientes com afasia de Broca apresentam discurso telegráfico, com pausas prolongadas e dificuldade na formulação de frases completas, enquanto indivíduos com afasia de Wernicke podem produzir discurso fluente, mas desorganizado e com parafasias frequentes (Gordon, 2020). O estudo de (Stark, Bryant, et al., 2023b), indica que a análise de discurso pode ser usada para avaliar a qualidade da produção verbal, diferenciando os tipos de afasia com base na fluência e coesão textual.

A análise de discurso espontâneo tem sido uma ferramenta valiosa para identificar esses déficits e auxiliar no planejamento de intervenções terapêuticas (Bryant et al., 2016).

Questões de Investigação/Objetivos

Questões de investigação

1. Quais são os padrões microlinguísticos do discurso produzido por adultos saudáveis falantes de português europeu, na tarefa de descrição da imagem “O encantador de histórias”?
2. De que forma variáveis sociodemográficas como a idade, a escolaridade e o sexo influenciam os parâmetros microlinguísticos analisados (diversidade lexical, produção de classes gramaticais, complexidade sintática)?
3. Qual é a distribuição da complexidade sintática em termos de T-Units simples, coordenadas, subordinadas e mistas na população adulta saudável?

Objetivos

Objetivo geral:

Quantificar e descrever o desempenho de parâmetros microlinguísticos do discurso produzido por adultos saudáveis falantes de português europeu, obtido através da descrição oral da figura "O encantador de histórias", parte integrante da versão revista da Bateria de

Avaliação da Afasia de Lisboa (BAAL-R) , bem como verificar a influência da idade, sexo e escolaridade nesses parâmetros, contribuindo para a sua aferição na população portuguesa.

Objetivos específicos

1. Descrever os parâmetros microlinguísticos do discurso espontâneo, incluindo duração, débito verbal, diversidade lexical, distribuição gramatical e extensão média dos enunciados.
2. Analisar a complexidade sintática dos discursos com base na segmentação em T-Units.
3. Avaliar a influência da idade, escolaridade e sexo no desempenho microlinguístico e na complexidade sintática.
4. Contribuir para a construção de dados normativos do português europeu.

METODOLOGIA

Este estudo é retrospectivo quantitativo e correlacional.

Participantes

A amostra foi obtida por conveniência, estratégia que permite acesso prático a um conjunto de dados diversificado, facilitando a recolha em contextos ecológicos reais e melhorando a validade externa do estudo.

A amostra, composta por 400 sujeitos foi dividida em três grupos etários: 18–39 anos, 40–64 anos e 65 anos ou mais. Para efeitos de análise, os participantes foram organizados em três grupos etários: adultícia jovem, dos 18 aos 39 anos (Erikson, 1950); adultícia média, dos 40 aos 64 anos (Erikson, 1950; (Levinson, 1986); e adultícia tardia, a partir dos 65 anos (Erikson, 1950).

O grupo dos 18–39 anos integrou 104 participantes (26,0% da amostra), dos quais 56 eram do sexo feminino (53,8%) e 48 do sexo masculino (46,2%). No que respeita à

lateralidade, 97 participantes eram destros (93,3%) e 7 canhotos (6,7%). Relativamente à escolaridade, 26 indivíduos (25,0%) tinham entre 4 e 12 anos de escolaridade e 78 (75,0%) possuíam mais de 12 anos.

O grupo dos 40–64 anos incluiu 198 participantes (49,5%), com 122 mulheres (61,6%) e 76 homens (38,4%). A maioria era destra (182 participantes; 91,9%) e 16 eram canhotos (8,1%). Em termos de escolaridade, 106 participantes (53,5%) tinham entre 4 e 12 anos e 92 (46,5%) apresentavam mais de 12 anos de escolaridade.

O grupo dos 65 anos ou mais era composto por 98 participantes (24,5%), dos quais 64 eram do sexo feminino (65,3%) e 34 do sexo masculino (34,7%). Quanto à lateralidade, 93 participantes (94,9%) eram destros e 5 (5,1%) canhotos. No que diz respeito à escolaridade, 72 participantes (73,5%) tinham entre 4 e 12 anos de escolaridade e 26 (26,5%) tinham mais de 12 anos.

Entre os grupos etários não se verificaram diferenças estatisticamente significativas quanto ao sexo e à lateralidade. Contudo, identificaram-se diferenças significativas no número de anos de escolaridade, com os participantes mais jovens (18–39 anos) a apresentarem, em média, mais anos de escolaridade do que os de meia-idade (40–64 anos) e, especialmente, do que os participantes com 65 anos ou mais. (Ver Tabela B1 no Anexo B)

Os critérios para inclusão no estudo foram os seguintes:

- Sujeitos sem diagnóstico conhecido de patologias neurológicas ou psiquiátricas.
- Falantes nativos de Português Europeu, assegurando a homogeneidade linguística, essencial para a análise microlinguística (Armstrong, 2000).
- Idade igual ou superior a 18 anos, abrangendo adultos com maturidade linguística consolidada (Hillis, 2007).

- Variabilidade educacional; a amostra inclui participantes com níveis de escolaridade com 4 ou mais anos de escolaridade, permitindo a análise do impacto do contexto educacional nas competências linguísticas (Bryant et al. 2016).

- De ambos os sexos

Foram excluídos desta amostra participantes com sinais de comprometimento cognitivo, identificados através do Mini-Mental State Examination (MMSE) (Folstein et al., 1975).

Recolha de dados

As amostras discursivas foram recolhidas por terapeutas da fala e estudantes de terapia da fala entre 2020 e 2023, cobrindo várias regiões do país. Esta abordagem facilitou a obtenção de uma amostra representativa das práticas linguísticas naturais

Cada participante foi solicitado a descrever a imagem intitulada “O encantador de histórias” (ANEXO A), seguindo as instruções padronizadas: "Diga-me o que vê nesta imagem, o que se passa?"

As respostas foram gravadas em formato áudio, utilizando dispositivos adequados para garantir a clareza e a fidelidade sonora necessárias à posterior transcrição e análise.

As transcrições foram realizadas no software EUDICO Linguistic Annotator (ELAN, versão 6.7), desenvolvido pelo Instituto Max Planck de Psicolinguística (Max Planck Institute for Psycholinguistics, 2024).

Transcrição e segmentação

A escolha dos critérios para a divisão de enunciados no discurso espontâneo foi fundamentada em estudos linguísticos e pragmáticos que destacam a complexidade do discurso oral, caracterizado pela sua fluidez, falta de pontuação explícita e dependência de pistas contextuais. Autores como Couper-Kuhlen e Selting, (1996); Crystal, (2008) e (M. A.

K. Halliday & Matthiessen, 2013), enfatizam a necessidade de uma abordagem multidimensional para delimitar unidades discursivas. Assim, os critérios foram organizados hierarquicamente, começando por aspetos sintáticos e semânticos, considerados mais estáveis, seguidos por pistas prosódicas e, finalmente, fenómenos fonéticos, que refletem nuances adicionais do discurso.

A hierarquia proposta reflete a combinação de abordagens gramaticais tradicionais, que analisam a estrutura formal da linguagem, com métodos mais modernos, que incorporam a pragmática e a fonética. Essa integração permite captar tanto a estrutura linguística quanto as pistas contextuais e interacionais que emergem no discurso espontâneo (Brazil, 1997; Crystal, 2008).

1. Critérios sintático-semânticos

- Estrutura Frásica: A conclusão de uma estrutura gramatical completa (sujeito + predicado) é um indicador primário de um enunciado (Halliday e Matthiessen 2014).
- Conjunções: O uso de conjunções coordenativas e subordinativas indica a relação entre partes do discurso, auxiliando na identificação de limites entre enunciados (Leech et al., 1984)
- Unidade de Sentido e Referência: A introdução de novos referentes ou a mudança de tema sinaliza a transição entre enunciados (Crystal 2008).
- Início e Fim de Orações: Identificar o início e o fim de frases é uma prática consagrada na análise sintática e essencial para delimitar unidades discursivas (M. A. K. Halliday & Matthiessen, 2013).

2. Critérios Prosódicos

- Entoação: Mudanças na curva entoacional (subidas e descidas de tom) refletem a estrutura do discurso e ajudam a identificar transições entre enunciados (Brazil, 1997; Couper-Kuhlen & Selting, 1996).

- Pausas significativas, especialmente aquelas superiores a dois segundos, são indicativas de fronteiras discursivas, frequentemente marcando o fim de um pensamento ou ideia (Crystal, 2008).

- Ritmo e Velocidade: Alterações no ritmo ou na velocidade da fala são usadas para destacar ou separar partes do discurso (Couper-Kuhlen & Selting, 1996).

3. Critérios Fonéticos

- Fenómenos Paralinguísticos: Elementos como risos, suspiros, interjeições e onomatopeias frequentemente indicam transições discursivas ou ajudam a transmitir emoções e intenções comunicativas orais (Jefferson et al., 1984). Essa abordagem integrativa garante uma análise mais detalhada e confiável, essencial para estudos microlinguísticos e pragmáticos.

Adotou-se a abordagem naturalista para a transcrição, um método que reproduz minuciosamente o discurso tal como foi produzido, incluindo elementos verbais, não verbais e contextuais (Oliver et al., 2005). As regras de transcrição seguiram as orientações de Magro, (2007), com ênfase na fidelidade à oralidade e na identificação de marcadores linguísticos e paralinguísticos relevantes (ver Tabela B1 no Anexo B)

Para além da transcrição naturalista, procedeu-se ainda à anotação da glosa. A glosa, no contexto linguístico, refere-se a anotações interlineares que acompanham o texto principal, geralmente para esclarecer significados, fornecer traduções ou registar formas alternativas de expressão (Haspelmath, 2004).

No presente estudo, a glosa é registada como complemento do processo de transcrição naturalista, mas não é alvo de análise direta, dado que o foco é o registo dos constituintes gramaticais e das estruturas microlinguísticas.

Os 400 discursos foram divididos em enunciados por dois investigadores, sendo atribuídos 200 a cada um e posteriormente trocados.

Procedimento de análise

Posteriormente à anotação e segmentação, procedeu-se à análise microlinguística dos enunciados discursivos, seguindo os critérios abaixo descritos:

1. Quantificação e Variedade dos Constituintes Morfossintáticos:

Foram contabilizados os constituintes gramaticais presentes nos enunciados, incluindo nomes, verbos, locuções, conjunções, preposições, advérbios, numerais, adjetivos, artigos, determinantes e pronomes.

As contrações de preposições com artigos (ex.: "do", "na", "pelos") foram contabilizadas como preposições, de acordo com os critérios estabelecidos para manter a consistência na análise.

A decisão metodológica de considerar como equivalentes as formas reduzidas e normativas (como “tá” e “está”) assenta-se numa perspetiva funcional da linguagem, em que a função discursiva prevalece sobre a forma fonológica (Bagno, 1999; Bortoni-Ricardo & Silva, 2022). Além disso, à luz da conceção de gramática como sistema internalizado e não exclusivamente normativo (Bortoni-Ricardo & Silva, 2022), reconhece-se que a variação fonológica observada na fala espontânea não compromete a função nem o significado pretendido..

As interjeições têm gerado algum debate no campo da linguística, particularmente quanto ao seu papel na análise microlinguística. Alguns autores, como (Ameka, 1992), e (Leech, 2006), defendem que as interjeições são elementos essenciais do discurso, com funções que vão além do mero preenchimento emocional. Para esses autores, as interjeições são uma ponte entre a emoção e a linguagem, refletindo o estado emocional do falante e a dinâmica interativa do discurso.

Por outro lado, autores como Crystal (2008) e (Wilkins, 1992), consideram as interjeições como manifestações periféricas no sistema linguístico. Argumentam que não

seguem regras gramaticais tradicionais e, portanto, não pertencem ao núcleo da linguagem. Para estes autores, as interjeições são mais próximas da comunicação paralinguística do que de constituintes gramaticais, funcionando como respostas emocionais universais e não como elementos estruturais.

Considerando assim que as interjeições são elementos naturais do discurso, fornecedoras de pistas sobre a sua organização, considerou-se relevante a sua contabilização neste estudo como elementos microlinguísticos.

2. Extensão média dos enunciados:

Calculou-se a média de palavras por enunciado, fornecendo uma medida de complexidade discursiva. Esta métrica reflete a capacidade de organização linguística e o nível de detalhamento do discurso.

3. Débito Verbal:

Foi avaliada a fluência verbal por meio do débito verbal, definido como o número de palavras produzidas por minuto. Este indicador permitiu compreender o ritmo do discurso e a sua densidade informacional.

4. Rácio type-token:

Foi calculado o do rácio Type-Token (RTT), que compara o número de palavras únicas (types) com o número total de palavras (tokens) em cada discurso. Segundo Crystal (2008), um RTT mais elevado indica maior riqueza vocabular e flexibilidade linguística. O RTT, embora amplamente utilizado como medida de diversidade lexical, apresenta limitações reconhecidas, sobretudo em amostras de grande extensão, uma vez que tende a diminuir à medida que o número de palavras aumenta, independentemente da real riqueza vocabular do falante (Malvern et al., 2004a; McCarthy & Jarvis, 2010; McKee, 2000). Em estudos com populações clínicas, como adultos com afasia, o RTT mostrou-se incapaz de captar

eficazmente a perda de diversidade lexical, uma vez que tende a refletir mais a quantidade de discurso produzido do que a qualidade lexical (Fergadiotis et al., 2015).

Dada a grande dimensão da amostra ($n = 400$) e a extensão variável dos discursos produzidos, o uso do RTT revelou-se limitado, uma vez que esta medida é sensível ao comprimento do discurso e tende a subestimar a diversidade lexical em produções mais longas.

5. Número de palavras de conteúdo diferentes

Tendo em conta a limitação do RTT, particularmente em amostras extensas e variáveis, e a dificuldade em aplicar medidas mais sofisticadas como o índice D (vocd-D) (Malvern et al., 2004a) ou o MTLT (Measure of Textual Lexical Diversity) (McCarthy, 2005) sem recurso a software específico, optou-se neste estudo por utilizar o número de palavras de conteúdo únicas como medida de diversidade vocabular. Esta abordagem, já adotada em estudos anteriores (Campos, 2015; Colozzo et al., 2011; Owen & Leonard, 2002), permite uma avaliação direta e específica da variedade de vocabulário com carga semântica efetivamente utilizada pelos participantes.

A utilização de palavras de conteúdo únicas foca-se nas palavras que mais contribuem para a construção do significado do discurso, oferecendo assim uma perspetiva mais precisa da riqueza lexical, menos influenciada por palavras funcionais recorrentes.

Estudos recentes destacam, aliás, que a produção discursiva eficaz depende da capacidade de transmitir informações relevantes, o que pode ser avaliado através da frequência e diversidade de palavras de conteúdo utilizadas (Kim & Kintz, 2024).

A literatura sobre envelhecimento normal indica que, embora a fluência verbal e a complexidade sintática possam sofrer alterações ao longo da vida, a quantidade de palavras de conteúdo tende a manter-se relativamente estável, constituindo um marcador fiável da competência comunicativa (Cooper, 1990).

A padronização da análise linguística em instrumentos de avaliação amplamente utilizados, como o *Boston Diagnostic Aphasia Examination – Third Edition* (Goodglass et al., 2001) e a *Western Aphasia Battery-Revised* (Kertesz, 2007), reforça a importância da quantificação das palavras de conteúdo para estabelecer parâmetros normativos e comparar diferentes grupos etários e níveis de escolaridade (Boucher et al., 2022). A riqueza lexical, entendida como um dos principais indicadores da competência linguística e da diversidade vocabular de um discurso (Segalowitz & Lane, 2000), exige, por isso, a definição rigorosa das categorias lexicais que constituem as palavras de conteúdo, distinguindo-as das palavras de função. Assim, para a análise da riqueza lexical no presente estudo, foi contabilizado o número de palavras de conteúdo diferentes (types), medida que permite avaliar a diversidade lexical informativa de cada discurso, excluindo repetições e refletindo a variedade real de unidades lexicais mobilizadas pelos falantes.

A opção por contabilizar apenas palavras de conteúdo diferentes visa captar a capacidade do falante de recorrer a um vocabulário diversificado, e não apenas a extensão ou fluência do discurso (Costa, 2008; Malvern et al., 2004a), sendo uma medida sensível na distinção entre populações com desenvolvimento típico e atípico da linguagem (Colozzo et al., 2011; Kover & Ellis Weismer, 2014).

A seleção das palavras de conteúdo para a análise da riqueza lexical não é pacífica, dada a controvérsia existente na literatura e a diversidade de opiniões entre autores. No presente trabalho, foram consideradas palavras de conteúdo os substantivos, os adjetivos e os verbos, por representarem unidades lexicais plenas de significado, fundamentais na construção do discurso, seguindo a perspectiva de autores como (SIM-SIM, 1998; Viana et al., 2017). Incluem-se também os numerais, enquanto subclasses dos substantivos com valor referencial, por acrescentarem informação relevante ao enunciado. Optou-se por excluir os

advérbios, dado o seu papel muitas vezes periférico na construção do conteúdo informativo central, e também as locuções.

Embora algumas locuções, como as adverbiais e verbais, possam ter um significado lexical forte, muitas outras desempenham funções estritamente gramaticais, atuando como conectores discursivos ou elementos estruturantes, sem contribuírem de forma significativa para a riqueza lexical do discurso. Vários autores referem que locuções prepositivas, conjuntivas e adverbiais funcionam, maioritariamente, como unidades fixas com valor gramatical ou discursivo, utilizadas para estabelecer relações lógicas, temporais ou modais entre segmentos do discurso, sem acrescentar novo conteúdo informativo. A análise do conteúdo lexical deve privilegiar as palavras portadoras de significado conceptual, como defende (SIM-SIM, 1998), o que reforça a exclusão de locuções com função meramente gramatical na contabilização da riqueza lexical. Estudos anteriores sugerem que a análise da riqueza lexical é mais precisa quando se baseia em unidades lexicais individuais, garantindo maior rigor na distinção entre palavras de conteúdo e palavras funcionais (Nation & Nation, 2001). Dessa forma, ao excluir as locuções, pretende-se assegurar maior consistência metodológica e fidelidade na medição da riqueza lexical dos discursos analisados.

Apesar dos advérbios poderem, em determinados contextos, transportar informação semântica, tanto (Silva, 2008), quanto Costa, (2008), destacam a natureza predominantemente funcional dessa categoria, o que justifica a sua exclusão do conjunto de palavras de conteúdo em determinadas análises linguísticas, nomeadamente nas que visam avaliar a carga informativa lexical do discurso.

Segundo Silva (2008), o advérbio é, antes de mais, uma palavra com função de modificação, funcionando como circunstante ou elemento modificador do verbo, do adjetivo ou da frase, o que o aproxima das palavras de função. Também Costa (1998), sublinha a heterogeneidade da classe dos advérbios, distinguindo entre aqueles que introduzem algum

conteúdo semântico (como os de modo ou lugar) e outros que funcionam como operadores discursivos e gramaticais, modulando polaridade, foco ou modalidade, sem acrescentar novo conteúdo. Esta perspetiva é partilhada por autores como Vilela (2001), que inclui os advérbios no grupo das palavras invariáveis, com função predominantemente gramatical, e SIM-SIM (1998), que, no contexto do desenvolvimento lexical, privilegia substantivos, adjetivos e verbos como principais categorias portadoras de conteúdo.

6. T-Units

Tendo em conta o objetivo de caracterizar em detalhe a produção linguística dos participantes, tornou-se relevante incluir na análise microlinguística a avaliação da complexidade sintática. Esta dimensão diz respeito ao modo como os falantes estruturam as suas ideias através de orações simples ou compostas, recorrendo à coordenação e/ou subordinação (Nippold et al., 2005; Rioux & Thordardottir, 2023).

Para avaliar esta dimensão, utiliza-se a unidade de análise designada T-Unit (terminable unit), definida como uma oração principal com todas as orações subordinadas ou estruturas não oracionais a ela associadas ((Hunt, 1970). A T-Unit representa uma unidade sintática e comunicativa completa, sendo particularmente adequada para o estudo do discurso oral (Lintunen & Mäkilä, 2014).

A segmentação das T. Units será realizada manualmente. Considera-se T-Unit apenas a unidade que contém pelo menos uma oração com verbo conjugado em forma finita (Nippold et al., 2005; Scott, 1988). Expressões incompletas, como interjeições ou nomes isolados, serão excluídas da contagem.

A divisão entre duas T-Units ocorre quando há duas ou mais orações principais independentes, ligadas por coordenação ou separadas por pausas marcadas com mudança sintática ou de tópico. No entanto, estruturas como reformulações, parênteses explicativos ou

inserções com função de esclarecimento devem ser integradas na mesma T-Unit, desde que dependam da mesma oração principal (Scott, 1988b; Xu & Casal, 2023).

Será tido em conta que, em discurso espontâneo uma única T-Unit pode, em alguns casos, prolongar-se por dois enunciados, especialmente quando há hesitações, repetições ou reformulações pelo falante, desde que a estrutura mantenha uma única oração principal e as suas subordinadas (Scott, 1988).

As T-Units identificadas serão classificadas quanto ao tipo de organização sintática, seguindo critérios definidos em estudos anteriores (Nippold et al., 2005; Rioux & Thordardottir, 2023; Xu, 2023):

- ✓ T-Units simples: apenas uma oração principal, sem coordenação nem subordinação;
- ✓ T-Units com coordenação: ligação de duas ou mais orações principais através de conjunções como e, mas, ou;
- ✓ T-Units com subordinação: presença de pelo menos uma oração subordinada, introduzida por conjunções como porque, quando, embora;
- ✓ T-Units mistas: combinação de orações coordenadas e subordinadas dentro da mesma unidade.

Esta classificação qualitativa permite analisar não apenas o número de unidades, mas também o tipo de complexidade estrutural presente, oferecendo uma visão mais abrangente da competência sintática dos participantes (Gardner-Neblett & Alvarez, 2024; Rioux & Thordardottir, 2023).

Instrumentos para análise de discurso

Para a análise discursiva recorreu-se ao software ELAN (Max Planck Institute for Psycholinguistics, 2024), amplamente reconhecido por sua eficiência na anotação e segmentação de discursos orais. O ELAN permitiu a anotação precisa de enunciados e a

contagem automatizada dos constituintes morfossintáticos. Posteriormente foi feita a exportação de dados para cálculo de métricas descritivas pelo IBM SPSS Statistics (Versão 29) (IBM, 2024), como extensão média de enunciados, débito verbal e TTR.

Os resultados da análise microlinguística contribuirão para a compreensão detalhada das características linguísticas e comunicativas da amostra avaliada. Esses dados serão essenciais para a construção de parâmetros normativos e para a avaliação de populações com alterações linguísticas.

Fiabilidade

A avaliação da fiabilidade interjuízes e intrajuízes foi realizada com base nos critérios da Consensus-based Standards for the selection of health Measurement INstruments (COSMIN) Tool (Mokkink et al., 2020), recomendada para estudos que avaliam propriedades de instrumentos de medida. Esta ferramenta orientou a análise da qualidade metodológica da segmentação em enunciados e das medidas microlinguísticas (número de palavras e diversidade lexical). A aplicação de testes estatísticos, como o Coeficiente de Correlação Intraclass (ICC), o Alfa de Cronbach e a correlação de Spearman, aliada à estabilidade das condições de análise e ao cegamento dos juízes, garantiu uma avaliação robusta e altamente fiável, com todos os critérios da COSMIN cumpridos.

Fiabilidade na divisão de enunciados

Concordância Interjuízes

Como parte da metodologia para a análise microlinguística conduzida neste estudo, foi realizada uma avaliação de concordância interjuízes com o objetivo de validar a fiabilidade do processo de segmentação de discursos em enunciados.

A análise de concordância foi conduzida por dois terapeutas da fala, experientes em análise linguística e clínica, seguindo os critérios metodológicos previamente estabelecidos, uma vez que este estudo integra uma investigação mais ampla.

Foram analisados 40 discursos transcritos, extraídos de uma amostra representativa da população-alvo do estudo principal.

Os dois terapeutas da fala realizaram, de forma independente, a segmentação dos discursos em enunciados, seguindo os critérios previamente estabelecidos.

Ambos os juízes participaram de várias sessões iniciais, na qual os critérios foram discutidos em detalhe, e exemplos práticos foram analisados para garantir a uniformidade do processo.

Seguidamente foram aplicadas as seguintes métricas estatísticas:

Porcentagem de Concordância Absoluta: Foi calculada a proporção de discursos em que os juízes concordaram plenamente na segmentação. Os resultados mostraram que 87,5% dos discursos (35 casos) apresentaram concordância total, enquanto 12,5% (5 casos) tiveram pelo menos um enunciado discordante.

Coefficiente de Correlação Intraclass (ICC): Utilizando um modelo de efeitos mistos bidirecional, o ICC foi calculado para medir a consistência entre os juízes:

Medidas únicas: ICC = 0,981 (IC 95%: 0,965 - 0,990);

Medidas médias: ICC = 0,990 (IC 95%: 0,982 - 0,995). Ambos os resultados apresentaram significância estatística ($p < 0,001$), indicando um elevado grau de confiabilidade.

Adicionalmente, o Alfa de Cronbach foi calculado para avaliar a consistência interna do protocolo de segmentação, alcançando um valor de 0,990, o que demonstra excelente confiabilidade e robustez metodológica.

Os resultados da avaliação de concordância interjuízes validaram a metodologia empregue na segmentação dos discursos, permitindo a continuidade da análise microlinguística no estudo principal.

Concordância Intrajuízes

Como parte da validação do protocolo de segmentação em enunciados, realizou-se uma análise de concordância intrajuízes para avaliar a consistência das decisões tomadas pela mesma juíza (Susana) em dois momentos distintos. Esta análise incluiu abordagens quantitativa, considerando o número de enunciados segmentados, e qualitativa, focando na concordância dos locais de divisão.

A mesma juíza participou de ambas as etapas, realizadas em momentos separados por um intervalo de tempo de 3 meses, suficiente para minimizar o viés de memória.

Foram analisados 40 discursos transcritos, representativos da população-alvo do estudo principal. Esses discursos foram segmentados pela juíza em dois momentos distintos, seguindo os critérios previamente definidos.

Os resultados de ambos os momentos foram comparados por meio de análises quantitativa e qualitativa.

Análise Quantitativa

Foi realizada a comparação do número de enunciados divididos pela juíza nos dois momentos. Este método visou identificar a estabilidade no número de segmentações realizadas e foi complementado pela avaliação estatística de consistência.

Análise Qualitativa

Além da comparação numérica, procedeu-se a uma análise qualitativa, categorizando os discursos de acordo com a concordância ou discordância na segmentação:

1. Concordância em todos os enunciados: Os discursos em que a segmentação coincidiu integralmente entre os dois momentos foram contabilizados. Observou-se concordância total em 82,5% dos discursos (33 casos).

2. Discordância em pelo menos um enunciado: Os discursos em que houve diferenças na segmentação foram classificados nesta categoria, abrangendo 17,5% dos casos (7 discursos).

Para reforçar a avaliação quantitativa, utilizou-se o ICC, que demonstrou alta consistência intrajuízes (Ver tabela A1 no Apêndice A):

- Medidas únicas: ICC = 0,989, intervalo de confiança de 95% [0,977 - 0,995];
- Medidas médias: ICC = 0,995, intervalo de confiança de 95% [0,988 - 0,997].

Ambos os valores apresentaram significância estatística ($p < 0,001$).

O Alfa de Cronbach foi calculado para avaliar a consistência interna das segmentações realizadas, com um valor de 0,995, indicando excelente confiabilidade.

Os resultados da análise intrajuízes, tanto quantitativa quanto qualitativa, validaram a robustez e a estabilidade do protocolo de segmentação em enunciados.

Fiabilidade na análise microlinguística

Para garantir a fidelidade da análise, foram selecionadas como medidas de referência o número total de palavras (tokens) e o número de palavras diferentes (types), métricas amplamente utilizadas na investigação sobre desenvolvimento lexical e variabilidade vocabular (Malvern et al., 2004b; Templin, 1957).

Essas medidas foram aplicadas tanto na concordância intrajuízes quanto na concordância interjuízes. Para o cálculo da concordância, foi acordado que 10% do número total de análises seriam verificadas, correspondendo a um total de 40 análises discursivas.

A avaliação foi realizada em dois momentos distintos, com um intervalo de três meses entre as análises, a fim de testar a estabilidade das contagens ao longo do tempo. Para aferir a consistência intrajuiz, foram comparadas 40 amostras discursivas analisadas pelo mesmo avaliador.

A análise de concordância intrajuízes, realizada através ICC, demonstrou uma elevada fiabilidade nas medições entre os dois momentos avaliados.

Para o total de palavras, o ICC (2,1) foi de $r = 1.000$, com um intervalo de confiança entre 0.999 e 1.000, indicando uma concordância praticamente perfeita e uma total estabilidade nas medições.

Para o total de palavras diferentes, o ICC (2,1) foi de $r = 0.969$, com um intervalo de confiança entre 0.943 e 0.984, refletindo uma concordância ainda muito elevada, embora ligeiramente inferior à verificada para o total de palavras. Essa diferença pode estar associada a pequenas variações na identificação das palavras únicas entre os momentos analisados.

Em ambos os casos, os valores foram estatisticamente significativos ($p < 0.001$), confirmando que a correlação não ocorreu por acaso. No geral, os resultados indicam que a metodologia utilizada para a segmentação e quantificação das palavras nos discursos é altamente fiável e consistente ao longo do tempo, garantindo uma avaliação estável e reprodutível.

A correlação de Spearman complementa esta análise ao verificar se os valores do 1º e 2º momento seguem um padrão semelhante, ou seja, se as unidades avaliadas que apresentam valores mais altos num momento também tendem a manter essa posição no outro momento.

A correlação entre o total de palavras no 1º momento e o total de palavras no 2º momento é $R = 0,998$ ($p < 0,001$).

A correlação entre o número de palavras diferentes no 1º momento e o número de palavras diferentes no 2º momento é $R = 0,968$ ($p < 0,001$).

Esse valor, embora ligeiramente inferior ao da primeira tabela, ainda representa uma correlação muito forte, indicando alta estabilidade na contagem de palavras únicas entre os momentos.

Esses valores indicam que eventuais variações na segmentação seguiram um padrão sistemático, reforçando a confiabilidade do método aplicado.

O intervalo de três meses entre as análises representa um desafio metodológico, pois pode levar a variações na segmentação por diversos fatores. No entanto, a correlação extremamente alta sugere que os critérios aplicados foram consistentes e reproduzíveis, reforçando a confiabilidade do processo de segmentação de enunciados (Krippendorff, 2018).

Os resultados indicam uma altíssima concordância intrajuízes, tanto na análise dos valores totais quanto nas diferenças entre os momentos, sustentando a reprodutibilidade dos critérios utilizados na segmentação dos enunciados. O coeficiente de Spearman acima de 0,95 sugere que o método utilizado é altamente confiável para futuras aplicações.

Com o objetivo de assegurar a consistência da análise microlinguística, procedeu-se à avaliação da fiabilidade interjuízes nas medidas de número total de palavras (tokens) e de palavras diferentes (types), com base numa amostra de discursos analisados independentemente por dois juízes. A concordância entre juízes foi avaliada através da correlação de Spearman.

Os resultados revelaram uma correlação muito forte e estatisticamente significativa entre os dois avaliadores, tanto na contagem de tokens como de types. Para os tokens, a correlação de Spearman foi de $\rho = 0,982$, com um valor-p inferior a 0,001, indicando uma elevada consistência na contabilização do número total de palavras. Relativamente aos types, a correlação foi de $\rho = 0,979$, também com um valor-p inferior a 0,001, demonstrando igualmente uma concordância quase perfeita na identificação das palavras únicas utilizadas por cada participante.

Estes resultados suportam a elevada fiabilidade do protocolo de análise microlinguística, garantindo que a diversidade e a extensão lexical dos discursos foram avaliadas de forma coerente e sistemática pelos dois juízes. A solidez desta etapa

metodológica reforça a validade dos dados obtidos e a robustez das conclusões que deles advêm.

RESULTADOS

Analisaram-se 400 discursos relativos à descrição da imagem “O Encantador de Histórias”, cujos resultados podem ser consultados na Tabela B2 no Apêndice B.

A análise por grupos de escolaridade revelou diferenças significativas no desempenho microlinguístico dos participantes (Tabela B4 no Apêndice B). Os indivíduos com escolaridade superior a 12 anos apresentaram um desempenho superior em diversos indicadores linguísticos. Observou-se uma maior diversidade lexical, evidenciada pelo uso mais alargado de nomes, verbos e palavras de conteúdo diferentes. A extensão média dos enunciados também foi superior neste grupo, sugerindo uma maior capacidade discursiva. No que respeita à complexidade sintática, destacou-se um aumento no número de T-Units com coordenação, indicando um uso mais frequente de estruturas coordenadas. Adicionalmente, verificou-se um uso mais frequente de preposições. O débito verbal revelou uma tendência para ser mais elevado nos participantes com maior escolaridade.

Verificaram-se diferenças estatisticamente significativas entre os grupos etários em várias dimensões do desempenho linguístico (Tabela B5 no Apêndice B). No plano da fluência, os resultados indicaram diferenças no número de enunciados produzidos e no débito verbal, com uma tendência para valores mais baixos nos participantes de idade mais avançada.

A extensão média dos enunciados também variou com a idade, sendo mais elevada no grupo mais jovem, mais baixa nos indivíduos de meia-idade e intermédia nos participantes mais velhos. No domínio lexical, observaram-se diferenças significativas no número de nomes e de verbos diferentes utilizados. Por outro lado, não se verificaram diferenças

estatisticamente significativas nas variáveis relacionadas com a complexidade sintática. Em particular, as medidas associadas às T-units, incluindo T-units simples, com coordenação, com subordinação, mistas e o total de T-units complexas, não diferiram significativamente entre os grupos etários. Estes dados sugerem que a idade influencia sobretudo a fluência e a diversidade lexical básica, sem afetar de forma evidente a produção de estruturas sintáticas complexas.

A análise post hoc de Tukey (Tabela B6 no Apêndice B) permitiu identificar diferenças estatisticamente significativas entre os grupos etários em variáveis associadas à fluência e à extensão do discurso. Os participantes mais jovens (18–39 anos) apresentaram um débito verbal superior ao dos grupos de meia-idade (40–64 anos) e mais velhos (+65). Verificou-se também que os jovens produziram enunciados significativamente mais longos do que os adultos de meia-idade. Em relação ao número total de enunciados, os participantes mais velhos produziram mais enunciados do que os dois grupos mais jovens. Estes resultados sugerem que a idade influencia sobretudo a velocidade e a extensão da produção discursiva, sem afetar de forma significativa a complexidade sintática ou a diversidade lexical geral.

Com o objetivo de identificar os preditores da complexidade sintática e da diversidade lexical, foram realizadas regressões lineares múltiplas (Tabelas B7 no Apêndice B). Como variáveis dependentes consideraram-se, respetivamente, o número total de T-units complexas e o total de palavras de conteúdo diferentes. As variáveis independentes incluídas foram o grupo de escolaridade e o grupo etário, por se tratar de fatores com suporte teórico na literatura e por apresentarem efeitos relevantes nas análises anteriores. Sexo e lateralidade foram excluídos por não apresentarem contributo estatisticamente significativo nem fundamentação teórica robusta neste contexto.

Na regressão linear com a complexidade sintática (número de T-Units complexas) como variável dependente, a escolaridade revelou-se um preditor significativo: os

participantes com mais de 12 anos de escolaridade produziram, em média, mais T-Units complexas ($B = 0.76$, $p = .002$). O grupo etário dos 65 anos ou mais também apresentou um efeito positivo significativo ($B = 0.97$, $p = .006$), sugerindo que a idade avançada, quando controlada pela escolaridade, está associada a uma maior complexidade sintática. O modelo explicou 4,8% da variância total ($R^2 = .048$).

Na segunda regressão, centrada na diversidade lexical — medida pelo número de palavras de conteúdo diferentes —, a escolaridade voltou a destacar-se como preditor significativo: os participantes com mais de 12 anos de escolaridade produziram, em média, mais palavras diferentes do que os restantes ($B = 4.40$, $p < .001$). O grupo etário dos 65 anos ou mais apresentou igualmente um efeito positivo significativo ($B = 3.39$, $p = .033$), sugerindo que a idade avançada está associada a uma maior diversidade lexical, mesmo quando controlada a escolaridade. Por contraste, os participantes entre os 40 e os 64 anos revelaram menor diversidade lexical relativamente ao grupo mais jovem ($B = -2.68$, $p = .028$). O modelo explicou 7,3% da variância observada ($R^2 = .073$).

Estes resultados revelam que tanto a escolaridade como a idade influenciam positivamente o desempenho linguístico, embora de formas distintas e, por vezes, não detetadas pelas análises de variância simples. A regressão múltipla permitiu evidenciar efeitos cruzados e revelar contributos independentes de cada variável, nomeadamente no caso da idade, cujo impacto positivo em algumas dimensões (como a diversidade lexical e a complexidade sintática) apenas se tornou visível após controlo pela escolaridade.

Ainda assim, importa destacar que os valores de R^2 obtidos foram baixos em ambos os modelos ($R^2 = .048$ e $R^2 = .073$), o que indica tamanhos de efeito globais reduzidos. Ou seja, embora a escolaridade se afirme como o preditor mais robusto e consistente, o seu impacto estatístico explica apenas uma fração da variância observada, o que é expectável dada a complexidade multifatorial da produção linguística em discurso espontâneo. Assim, a

relevância dos efeitos identificados deve ser interpretada com cautela, valorizando-se sobretudo o seu significado teórico e clínico, mais do que o seu peso estatístico isolado.

DISCUSSÃO

O presente estudo teve como principal objetivo descrever os parâmetros microlinguísticos do discurso espontâneo de adultos saudáveis falantes de português europeu, com base na tarefa de descrição da imagem “O encantador de histórias”. A análise centrou-se em variáveis como a diversidade lexical, o débito verbal, a extensão média dos enunciados, a produção de classes gramaticais e a complexidade sintática, medida por T-Units. Adicionalmente, procurou-se avaliar de que forma variáveis sociodemográficas, nomeadamente a idade, a escolaridade, o sexo e a lateralidade, influenciam o desempenho linguístico dos participantes.

Os resultados revelaram um impacto significativo da escolaridade em diversos parâmetros do discurso. Indivíduos com escolaridade superior a 12 anos produziram discursos mais extensos, com maior número de palavras de conteúdo diferentes, e exibiram maior complexidade sintática, nomeadamente através de estruturas coordenadas. Estes dados estão em consonância com a literatura, que identifica a escolaridade como um dos principais preditores do desempenho linguístico, influenciando a fluência, a variedade lexical e a construção sintática (Nippold et al., 2005; Radanovic et al., 2004; Reno & McMaster, 2024). O contacto continuado com práticas de leitura e escrita, habitualmente associado a níveis educacionais mais elevados, contribui para o desenvolvimento e manutenção de competências linguísticas complexas, funcionando como fator protetor contra o declínio cognitivo (Branco et al., 2014; Pawlowski et al., 2012).

A idade mostrou impacto em variáveis relacionadas com a fluência e a extensão do discurso. Participantes mais jovens (18–39 anos) apresentaram um débito verbal mais elevado

e produziram mais enunciados do que os grupos de meia-idade (40–64 anos) e adultos mais velhos (65+). A extensão média dos enunciados também foi superior neste grupo. Estes achados são consistentes com investigações que apontam para um declínio moderado da fluência verbal com o envelhecimento, refletindo alterações no acesso lexical e na velocidade de processamento (Marini et al., 2005; Kim & Kintz, 2024; Boyle, 2011). No entanto, não se verificaram diferenças significativas entre os grupos etários quanto à complexidade sintática, medida através das T-Units. Este resultado sugere que, apesar da idade afetar o ritmo e a extensão do discurso, a competência sintática se mantém estável ao longo da vida adulta, em concordância com autores como (Nippold et al., 2014; Roll, 2015).

Quanto à diversidade lexical, o número de palavras de conteúdo diferentes variou significativamente com a escolaridade, mas não com a idade. Este dado reforça a ideia de que a diversidade vocabular depende mais do contexto educacional e da exposição à linguagem escrita do que do envelhecimento por si só (Malcorra et al., 2021). A opção metodológica de contabilizar palavras de conteúdo diferentes, em detrimento do rácio type-token, revelou-se adequada, considerando as limitações reconhecidas desta medida em amostras extensas (Fergadiotis et al., 2015).

O sexo e a lateralidade não mostraram influência estatisticamente significativa nas variáveis microlinguísticas analisadas, o que corrobora estudos que apontam para uma reduzida expressão destas variáveis em tarefas estruturadas como a descrição de imagens (Boucher et al., 2022; Kim & Kintz, 2024).

Importa referir que não se observaram diferenças significativas entre os grupos etários quanto ao sexo e à lateralidade, assegurando homogeneidade nesses aspetos. No entanto, a escolaridade variou significativamente entre grupos, com os mais jovens a apresentarem, em média, mais anos de escolaridade. Esta diferença poderá introduzir um viés na interpretação dos efeitos da idade, uma vez que o nível educativo está fortemente associado ao desempenho

linguístico. Ainda assim, os efeitos das variáveis sociodemográficas foram analisados de forma independente, o que permite isolar o contributo de cada variável e reforça a validade interna dos resultados.

Para além das análises comparativas, a aplicação de regressões lineares múltiplas permitiu aprofundar a compreensão dos fatores que influenciam o desempenho microlinguístico. A escolaridade destacou-se como o preditor mais robusto tanto da diversidade lexical como da complexidade sintática, confirmando estudos prévios (Nippold et al., 2005; Reno & McMaster, 2024). A idade, embora sem efeitos significativos nas análises de variância, revelou um contributo positivo quando controlada pela escolaridade, sugerindo que adultos mais velhos com mais anos de escolaridade mantêm um desempenho linguístico elevado, em linha com a hipótese da reserva cognitiva (Bialystok, 2007; Cooper, 1990). Estes dados reforçam a importância de considerar os efeitos cruzados entre variáveis sociodemográficas na análise do discurso adulto.

Este trabalho fornece contributos relevantes para a construção de parâmetros normativos em Português Europeu e para a futura aferição da tarefa “O encantador de histórias” no âmbito da BAAL-R. O uso de uma tarefa de descrição de imagem padronizada, aplicada a uma amostra ampla e diversificada, permite estabelecer valores de referência ajustados à realidade sociocultural portuguesa e sensíveis às variações individuais esperadas em idade adulta.

CONCLUSÃO

O presente estudo contribuiu para a caracterização detalhada do desempenho microlinguístico em adultos saudáveis falantes de português europeu, com base na tarefa de descrição da imagem “O encantador de histórias”. A análise de 400 discursos permitiu

identificar padrões linguísticos robustos, úteis para a construção de dados normativos no âmbito da avaliação do discurso.

Os resultados evidenciaram um efeito marcante da escolaridade na diversidade lexical e na complexidade sintática, confirmando a sua importância como preditor central das competências linguísticas. Indivíduos com escolaridade superior a 12 anos apresentaram produções mais elaboradas, maior riqueza vocabular e estruturas sintáticas mais complexas. Este efeito foi consistente com a literatura e reforçado pelos modelos de regressão múltipla, que demonstraram que a escolaridade mantém um impacto significativo mesmo quando controladas outras variáveis. A idade revelou influência sobretudo na fluência e extensão do discurso, sendo os participantes mais jovens os que apresentaram maior débito verbal. No entanto, a competência sintática manteve-se estável ao longo da vida adulta, o que corrobora a preservação de estruturas linguísticas em falantes saudáveis.

A análise de regressão permitiu ainda observar que a idade, quando considerada em interação com a escolaridade, pode associar-se a um desempenho linguístico elevado, sugerindo que a reserva cognitiva decorrente de uma maior bagagem educativa contribui para mitigar os efeitos do envelhecimento. Estes achados sublinham a importância de considerar os efeitos cruzados entre variáveis sociodemográficas na interpretação dos dados linguísticos.

A ausência de impacto do sexo e da lateralidade nas variáveis analisadas está em linha com investigações anteriores que apontam para um peso reduzido destas variáveis em tarefas estruturadas. A metodologia utilizada, com segmentação rigorosa dos enunciados e análise microlinguística criteriosa, revelou-se fiável e adequada aos objetivos propostos.

No seu conjunto, os dados obtidos oferecem um contributo relevante para a aferição da tarefa “O encantador de histórias”, a integrar na versão revista da Bateria de Avaliação da Afasia de Lisboa (BAAL-R). A definição de parâmetros normativos baseados numa amostra ampla e diversa assegura maior sensibilidade e precisão na distinção entre padrões

linguísticos típicos e atípicos, reforçando o valor desta tarefa como ferramenta de avaliação clínica e investigação em linguagem adulta

BIBLIOGRAFIA

- Agmon, G., Cho, S., Ash, S., Cousins, K. A., Blennow, K., Zetterberg, H., Shaw, L. M., Pradhan, S., Kim, Y. D., & Liberman, M. Y. (2025). Automatic quantification of syntactic complexity in natural spontaneous speech of people with primary progressive aphasia. *Aphasiology*, 1–22.
- Alyahya, R. S. W. (2024). The public awareness and knowledge of aphasia in Saudi Arabia. *Aphasiology*, 38(8), 1377–1389. <https://doi.org/10.1080/02687038.2023.2287262>
- Ameka, F. (1992). Interjections: The universal yet neglected part of speech. *Journal of pragmatics*, 18(2–3), 101–118.
- Andreetta, S., Cantagallo, A., & Marini, A. (2012). Narrative discourse in anomic aphasia. *Neuropsychologia*, 50(8), 1787–1793. <https://doi.org/10.1016/j.neuropsychologia.2012.04.003>
- Andreetta, S., & Marini, A. (2015). The effect of lexical deficits on narrative disturbances in fluent aphasia. *Aphasiology*, 29(6), 705–723. <https://doi.org/10.1080/02687038.2014.979394>
- Armstrong, E. (2000). Aphasic discourse analysis: The story so far. *Aphasiology*, 14(9), 875–892. <https://doi.org/10.1080/02687030050127685>

- Bagno, M. (1999). *Preconceito lingüístico: O que é, como se faz*. Edições Loyola.
- Bayot, R., & Gonçalves, T. (2016). Multilingual author profiling using word embedding averages and svms. *2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA)*, 382–386.
- Beers, S. F., & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and writing*, 22, 185–200.
- Berman, R., & Verhoeven, L. (2002). Cross-linguistic perspectives on the development of text-production abilities: Speech and writing. *Written Language & Literacy*, 5(1), 1–43.
- Berube, S., Nonnemacher, J., Demsky, C., Glenn, S., Saxena, S., Wright, A., Tippet, D. C., & Hillis, A. E. (2019). Stealing Cookies in the Twenty-First Century: Measures of Spoken Narrative in Healthy Versus Speakers With Aphasia. *American Journal of Speech-Language Pathology*, 28(1S), 321–329. https://doi.org/10.1044/2018_AJSLP-17-0131
- Bialystok, E. (2007). Acquisition of literacy in bilingual children: A framework for research. *Language learning*, 57, 45–77.
- Biber, D., & Gray, B. (2011). Grammatical change in the noun phrase: The influence of written language use. *English Language & Linguistics*, 15(2), 223–250.
- Binder, J. R., Gross, W. L., Allendorfer, J. B., Bonilha, L., Chapin, J., Edwards, J. C., Grabowski, T. J., Langfitt, J. T., Loring, D. W., Lowe, M. J., Koenig, K., Morgan, P. S., Ojemann, J. G., Rorden, C., Szaflarski, J. P., Tivarus, M. E., & Weaver, K. E. (2011). Mapping anterior temporal lobe language areas with fMRI: A multicenter normative study. *NeuroImage*, 54(2), 1465–1475. <https://doi.org/10.1016/j.neuroimage.2010.09.048>

- Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender. *Language and Speech*, 44(2), 123–147.
<https://doi.org/10.1177/00238309010440020101>
- Bortoni-Ricardo, S. M., & Silva, K. A. da. (2022). Sociolinguística educacional: Uma entrevista com Stella Maris Bortoni-Ricardo. *Linguagem em (Dis) curso*, 22(1), 219–231.
- Boucher, J., Brisebois, A., Slegers, A., Courson, M., Désilets-Barnabé, M., Chouinard, A.-M., Gbeglo, V., Marcotte, K., & Brambati, S. M. (2022). Picture description of the Western Aphasia Battery picnic scene: Reference data for the French Canadian population. *American journal of speech-language pathology*, 31(1), 257–270.
- Boyle, M. (2011). Discourse treatment for word retrieval impairment in aphasia: The story so far. *Aphasiology*, 25(11), 1308–1326. <https://doi.org/10.1080/02687038.2011.596185>
- Branco, L. D., Cotrena, C., Pereira, N., Kochhann, R., & Fonseca, R. P. (2014). Verbal and visuospatial executive functions in healthy elderly: The impact of education and frequency of reading and writing. *Dementia & Neuropsychologia*, 8(2), 155–161.
<https://doi.org/10.1590/S1980-57642014DN82000011>
- Brazil, D. (1997). *The communicative value of intonation in English book*. Cambridge University Press.
- Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics & Phonetics*, 30(7), 489–518.
<https://doi.org/10.3109/02699206.2016.1145740>
- Bryant, L., Ferguson, A., Valentine, M., & Spencer, E. (2019). Implementation of discourse analysis in aphasia: Investigating the feasibility of a Knowledge-to-Action

- intervention. *Aphasiology*, 33(1), 31–57.
- <https://doi.org/10.1080/02687038.2018.1454886>
- Campos, R. J. P. M. de. (2015). *Rácio Type-Token e D como indicadores de desenvolvimento linguístico no português europeu*.
- Cannizzaro, M. S., & Coelho, C. A. (2013). Analysis of Narrative Discourse Structure as an Ecologically Relevant Measure of Executive Function in Adults. *Journal of Psycholinguistic Research*, 42(6), 527–549. <https://doi.org/10.1007/s10936-012-9231-5>
- Capilouto, G. J., Wright, H. H., & Maddy, K. M. (2016). Microlinguistic processes that contribute to the ability to relay main events: Influence of age. *Aging, Neuropsychology, and Cognition*, 23(4), 445–463.
- <https://doi.org/10.1080/13825585.2015.1118006>
- Colozzo, P., Gillam, R. B., Wood, M., Schnell, R. D., & Johnston, J. R. (2011). Content and form in the narratives of children with specific language impairment. *Journal of speech, language, and hearing research*, 54(6), 1609–1627.
- Cooper, P. V. (1990). Discourse Production and Normal Aging: Performance on Oral Picture Description Tasks. *Journal of Gerontology*, 45(5), P210–P214.
- <https://doi.org/10.1093/geronj/45.5.P210>
- Costa, J. (2008). O advérbio em português europeu. (*No Title*).
- Coulthard, M. (2014). Powerful evidence for the defence: An exercise in forensic discourse analysis. Em *Language and the Law* (pp. 414–427). Routledge.
- Couper-Kuhlen, E., & Selting, M. (1996). *Prosody in conversation: Interactional studies* (Número 12). Cambridge University Press.
- Crystal, D. (2008). *A dictionary of linguistics and phonetics (The Language Library)*. John Wiley & Sons Incorporated.

- Faria, I. H. (2003). O uso da linguagem. Em *Gramática da língua portuguesa* (7ª). Caminho.
- Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in cognitive sciences*, 18(3), 120–126.
- Fergadiotis, G., Wright, H. H., & Green, S. B. (2015). Psychometric evaluation of lexical diversity indices: Assessing length effects. *Journal of Speech, Language, and Hearing Research*, 58(3), 840–852.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, 12(3), 189–198.
- Friederici, A. D., Chomsky, N., Berwick, R. C., Moro, A., & Bolhuis, J. J. (2017). Language, mind and brain. *Nature human behaviour*, 1(10), 713–722.
- Fromkin, V., Rodman, R., & Hyams, N. (2013). *An Introduction to Language*. Boston: Cengage.
- Gardner-Neblett, N., & Alvarez, D. L. (2024). Sharing Stories Versus Explaining Facts: Comparing African American Children’s Microstructure Performance Across Fictional Narrative, Informational, and Procedural Discourse. *Journal of Speech, Language, and Hearing Research*, 67(11), 4431–4445.
- Goodglass, H., Kaplan, E., & Weintraub, S. (2001). *BDAE: The Boston diagnostic aphasia examination*. Lippincott Williams & Wilkins Philadelphia, PA.
- Gordon, J. K. (2020). Factor Analysis of Spontaneous Speech in Aphasia. *Journal of Speech, Language & Hearing Research*, 63(12), 4127–4127–4147. Academic Search Complete. https://doi.org/10.1044/2020_JSLHR-20-00340
- Hagoort, P. (2013). MUC (memory, unification, control) and beyond. *Frontiers in psychology*, 4, 416.

- Halliday, M. A. (1978). Ideas about language. *Arts: The Journal of the Sydney University Arts Association*, 11.
- Halliday, M. A. K., & Matthiessen, C. M. (2013). *Halliday's introduction to functional grammar*. Routledge.
- Harris Wright, H., & Capilouto, G. J. (2012). Considering a multi-level approach to understanding maintenance of global coherence in adults with aphasia. *Aphasiology*, 26(5), 656–672. <https://doi.org/10.1080/02687038.2012.676855>
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews neuroscience*, 8(5), 393–402.
- Hillis, A. E. (2007). Aphasia: Progress in the last quarter of a century. *Neurology*, 69(2), 200–213. <https://doi.org/10.1212/01.wnl.0000265600.69385.6f>
- Hossain, M. Z., & Samin, A. M. (2022). *Analysis of Male and Female Speakers' Word Choices in Public Speeches* (Versão 1). arXiv. <https://doi.org/10.48550/ARXIV.2211.06366>
- Hunt, K. W. (1970). Do Sentences in the Second Language Grow like Those in the First? *TESOL Quarterly*, 4(3), 195. <https://doi.org/10.2307/3585720>
- Jefferson, G., Atkinson, J. M., & Heritage, J. (1984). *Structures of social interaction*. Cambridge University Press New York.
- Karbe, H., Thiel, A., Weber-Luxenburger, G., Herholz, K., Kessler, J., & Heiss, W.-D. (1998). Brain Plasticity in Poststroke Aphasia: What Is the Contribution of the Right Hemisphere? *Brain and Language*, 64(2), 215–230. <https://doi.org/10.1006/brln.1998.1961>
- Kertesz, A. (2006). Western aphasia battery–revised (WAB-R). *Austin, TX: Pro-Ed*, 10.
- Kertesz, A. (2007). *Western aphasia battery—Revised*.

- Kertesz, A., Harlock, W., & Coates, R. (1979). Computer tomographic localization, lesion size, and prognosis in aphasia and nonverbal impairment. *Brain and Language*, 8(1), 34–50. [https://doi.org/10.1016/0093-934X\(79\)90038-5](https://doi.org/10.1016/0093-934X(79)90038-5)
- Kim, H., & Kintz, S. (2024). Discourse Production Across the Adult Lifespan: Microlinguistic Processes. *Topics in Cognitive Science*.
- Knecht, S. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain*, 123(12), 2512–2518. <https://doi.org/10.1093/brain/123.12.2512>
- Kover, S. T., & Ellis Weismer, S. (2014). Lexical characteristics of expressive vocabulary in toddlers with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*, 57(4), 1428–1441.
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications.
- Leech, G. (2006). *Glossary of English grammar*. Edinburgh University Press.
- Leech, G., Hoogenraad, R., Birch, D., & Birch, D. (1984). English grammar for today: A new introduction. Em *RELC Journal* (Vol. 15, Número 2, pp. 95–100). Sage Publications Sage CA: Thousand Oaks, CA.
- Levinson, D. J. (1986). A conception of adult development. *American Psychologist*, 41(1), 3–13.
- Lintunen, P., & Mäkilä, M. (2014). Measuring syntactic complexity in spoken and written learner language: Comparing the incomparable? *Research in Language*, 12(4), 377–399.
- Lu, X. (2011). A Corpus-Based Evaluation of Syntactic Complexity Measures as Indices of College-Level ESL Writers' Language Development. *TESOL Quarterly*, 45(1), 36–62. <https://doi.org/10.5054/tq.2011.240859>

- MacWhinney, B., & Fromm, D. (2016). AphasiaBank as BigData. *Seminars in Speech & Language, 37*(1), 10–22. Complementary Index. <https://doi.org/10.1055/s-0036-1571357>
- Magro, C. (2007). *Clíticos: Variações sobre o tema*. Universidade de Lisboa (Portugal).
- Malcorra, B. L., Wilson, M. A., & Hübner, L. C. (2021). Avaliação da produção discursiva oral no envelhecimento e sua relação com escolaridade e hábitos de leitura e escrita: Uma revisão sistemática. *Estudos Interdisciplinares sobre o Envelhecimento, 26*(2).
- Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004a). *Lexical diversity and language development*. Springer.
- Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004b). *Lexical diversity and language development*. Springer.
- Marini, A., Carlomagno, S., Caltagirone, C., & Nocentini, U. (2005). The role played by the right hemisphere in the organization of complex textual structures. *Brain and language, 93*(1), 46–54.
- McCarthy, P. M. (2005). *An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual, lexical diversity (MTLD)*. The University of Memphis.
- McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior research methods, 42*(2), 381–392.
- McKee, G. (2000). Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing, 15*(3), 323–338. <https://doi.org/10.1093/llc/15.3.323>
- Mokkink, L. B., Boers, M., Van Der Vleuten, C. P. M., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C., & Terwee, C. B. (2020). COSMIN Risk of Bias tool to assess the

- quality of studies on reliability or measurement error of outcome measurement instruments: A Delphi study. *BMC medical research methodology*, 20, 1–13.
- Nation, I. S., & Nation, I. S. P. (2001). *Learning vocabulary in another language* (Vol. 10). Cambridge university press Cambridge.
- Nippold, M. A., Cramond, P. M., & Hayward-Mayhew, C. (2014). Spoken language production in adults: Examining age-related differences in syntactic complexity. *Clinical linguistics & phonetics*, 28(3), 195–207.
- Nippold, M. A., Hesketh, L. J., Duthie, J. K., & Mansfield, T. C. (2005). Conversational versus expository discourse. *Journal of Speech, Language, and Hearing Research*, 48(5), 1048–1064.
- Nippold, M. A., Mansfield, T. C., & Billow, J. L. (2007). Peer Conflict Explanations in Children, Adolescents, and Adults: Examining the Development of Complex Syntax. *American Journal of Speech-Language Pathology*, 16(2), 179–188.
[https://doi.org/10.1044/1058-0360\(2007/022\)](https://doi.org/10.1044/1058-0360(2007/022))
- Oliver, D. G., Serovich, J. M., & Mason, T. L. (2005). Constraints and Opportunities with Interview Transcription: Towards Reflection in Qualitative Research. *Social Forces*, 84(2), 1273–1289. <https://doi.org/10.1353/sof.2006.0023>
- Owen, A. J., & Leonard, L. B. (2002). Lexical diversity in the spontaneous speech of children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 45(5), 927–937.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism*.
- Pawlowski, J., Remor, E., de Mattos Pimenta Parente, M. A., de Salles, J. F., Fonseca, R. P., & Bandeira, D. R. (2012). The influence of reading and writing habits associated with education on the neuropsychological performance of Brazilian adults. *Reading and Writing*, 25, 2275–2289.

- Pereira, A., & Ortiz, K. Z. (2022a). Language skills differences between adults without formal education and low formal education. *Psicologia: Reflexão e Crítica*, 35(1), 4.
<https://doi.org/10.1186/s41155-021-00205-9>
- Pereira, A., & Ortiz, K. Z. (2022b). Language skills differences between adults without formal education and low formal education. *Psicologia: Reflexão e Crítica*, 35, 4.
- Pistono, A., Pariente, J., Bézy, C., Pastor, J., Tran, T. M., Renard, A., Fossard, M., Nespoulous, J.-L., & Jucla, M. (2017). Inter-individual variability in discourse informativeness in elderly populations. *Clinical linguistics & phonetics*, 31(5), 391–408.
- Poeppel, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Current opinion in neurobiology*, 28, 142–149.
- Pracar, A. L., Biondo, N., Dronkers, N. F., & Ivanova, M. V. (2025). The neuroanatomy of Broca's aphasia. *Frontiers in Language Sciences*, 4, 1496209.
<https://doi.org/10.3389/flang.2025.1496209>
- Rabinovich, E., Gonen, H., & Stevenson, S. (2020). Pick a Fight or Bite your Tongue: Investigation of Gender Differences in Idiomatic Language Usage. *Proceedings of the 28th International Conference on Computational Linguistics*, 5181–5192.
<https://doi.org/10.18653/v1/2020.coling-main.454>
- Radanovic, M., Mansur, L. L., & Scaff, M. (2004). Normative data for the Brazilian population in the Boston Diagnostic Aphasia Examination: Influence of schooling. *Brazilian Journal of Medical and Biological Research*, 37(11), 1731–1738.
<https://doi.org/10.1590/S0100-879X2004001100019>
- Reno, E. A., & McMaster, K. L. (2024). Measuring linguistic growth in sentence-level writing curriculum-based measures: Exploring complementary scoring methods. *Language, speech, and hearing services in schools*, 55(2), 529–544.

- Rioux, E. J., & Thordardottir, E. (2023). Bilingual Adolescents' Complex Syntax Production in Both Languages Across Conversational and Expository Contexts. *Journal of Speech, Language, and Hearing Research*, 66(8), 2733–2749.
https://doi.org/10.1044/2023_JSLHR-22-00193
- Rofes, A., de Aguiar, V., Jonkers, R., Oh, S. J., DeDe, G., & Sung, J. E. (2020). What drives task performance during animal fluency in people with Alzheimer's disease? *Frontiers in psychology*, 11, 1485.
- Roll, M. (2015). A neurolinguistic study of South Swedish word accents: Electrical brain potentials in nouns and verbs. *Nordic Journal of Linguistics*, 38(2), 149–162.
- Scott, C. M. (1988a). Producing complex sentences. *Topics in language disorders*, 8(2), 44–62.
- Scott, C. M. (1988b). Producing complex sentences. *Topics in language disorders*, 8(2), 44–62.
- Scott, C. M., & Windsor, J. (2000). General Language Performance Measures in Spoken and Written Narrative and Expository Discourse of School-Age Children With Language Learning Disabilities. *Journal of Speech, Language, and Hearing Research*, 43(2), 324–339. <https://doi.org/10.1044/jslhr.4302.324>
- Segalowitz, S. J., & Lane, K. C. (2000). Lexical access of function versus content words. *Brain and language*, 75(3), 376–389.
- Sherratt, S. (sem data). Syntactic Complexity in Spoken Discourse: The Effect of Age and Socioeconomic Status. *Available at SSRN 5009888*.
- Sherratt, S. (2007). Multi-level discourse analysis: A feasible approach. *Aphasiology*, 21(3–4), 375–393.
- Silva, A. A. L. V. da. (2008). *Estatuto sintáctico doa" advérbios": Função e classe*.

- SIM-SIM, I. (1998). Do uso da linguagem à consciência linguística. *SIM-SIM, I. Desenvolvimento da linguagem. Lisboa: Universidade aberta*, 225–236.
- Snitz, B. E., Unverzagt, F. W., Chang, C.-C. H., Bilt, J. V., Gao, S., Saxton, J., Hall, K. S., & Ganguli, M. (2009). Effects of age, gender, education and race on two tests of language ability in community-based older adults. *International Psychogeriatrics*, 21(6), 1051–1062. <https://doi.org/10.1017/S1041610209990214>
- Song, H. S., Hwang, E., Yoo, S. A., Jeon, Y. S., Choi, S. R., Lee, Y., & Sung, J. E. (2022). Efficacy of a daily life based cognitive-linguistic training program: Evidence from randomized controlled trials for healthy middle-aged adults. *Korean Speech-Language & Hearing Association*, 31(3), 91–107.
- Stark, B. C., Alexander, J. M., Hittson, A., Doub, A., Igleheart, M., Streander, T., & Jewell, E. (2023). Test–Retest Reliability of Microlinguistic Information Derived From Spoken Discourse in Persons With Chronic Aphasia. *Journal of Speech, Language, and Hearing Research*, 66(7), 2316–2345. https://doi.org/10.1044/2023_JSLHR-22-00266
- Stark, B. C., Bryant, L., Themistocleous, C., Den Ouden, D.-B., & Roberts, A. C. (2023a). Best practice guidelines for reporting spoken discourse in aphasia and neurogenic communication disorders. *Aphasiology*, 37(5), 761–784. <https://doi.org/10.1080/02687038.2022.2039372>
- Stark, B. C., Bryant, L., Themistocleous, C., Den Ouden, D.-B., & Roberts, A. C. (2023b). Best practice guidelines for reporting spoken discourse in aphasia and neurogenic communication disorders. *Aphasiology*, 37(5), 761–784. <https://doi.org/10.1080/02687038.2022.2039372>
- Stark, B. C., Dalton, S. G., & Lanzi, A. M. (2025). Access to context-specific lexical-semantic information during discourse tasks differentiates speakers with latent

- aphasia, mild cognitive impairment, and cognitively healthy adults. *Frontiers in Human Neuroscience*, 18, 1500735. <https://doi.org/10.3389/fnhum.2024.1500735>
- Sun, J., & Moens, M.-F. (2023). Fine-tuned vs. Prompt-tuned supervised representations: Which better account for brain language representations? *arXiv preprint arXiv:2310.01854*.
- Szaflarski, J. P., Binder, J. R., Possing, E. T., McKiernan, K. A., Ward, B. D., & Hammeke, T. A. (2002). Language lateralization in left-handed and ambidextrous people: fMRI data. *Neurology*, 59(2), 238–244. <https://doi.org/10.1212/WNL.59.2.238>
- Templin, M. C. (1957). *Certain language skills in children; their development and interrelationships*.
- Tessaro, B., Hermes-Pereira, A., Schilling, L. P., Fonseca, R. P., Kochhann, R., & Hübner, L. C. (2020). Verbal fluency in Alzheimer’s disease and mild cognitive impairment in individuals with low educational level and its relationship with reading and writing habits. *Dementia & Neuropsychologia*, 14(3), 300–307. <https://doi.org/10.1590/1980-57642020dn14-030011>
- Thompson, C. K., & Bastiaanse, R. (2012). Introduction to agrammatism. Em *Perspectives on agrammatism* (pp. 1–16). Psychology Press.
- Tivarus, M. E., Starling, S. J., Newport, E. L., & Langfitt, J. T. (2012). Homotopic language reorganization in the right hemisphere after early left hemisphere injury. *Brain and Language*, 123(1), 1–10. <https://doi.org/10.1016/j.bandl.2012.06.006>
- Viana, F. L., Silva, C., Ribeiro, I., & Cadime, I. M. D. (2017). *Instrumentos de avaliação da linguagem: Uma perspectiva global*.
- Vygotsky, L. S. (1984). A formação social da mente. *São Paulo*, 3.
- Wilkins, D. P. (1992). Interjections as deictics. *Journal of Pragmatics*, 18(2–3), 119–158.

- Xu, Y., & Casal, J. E. (2023). Navigating complexity in plain English: A longitudinal analysis of syntactic and lexical complexity development in L2 legal writing. *Journal of Second Language Writing*, 62, 101059.
- Ye, Z., Ai, Q., Liu, Y., de Rijke, M., Zhang, M., Lioma, C., & Ruotsalo, T. (2025). Generative language reconstruction from brain recordings. *Communications Biology*, 8(1), 346.
- Zhang, Y., Wang, S., Dong, X., Yu, J., & Zong, C. (2024). Navigating Brain Language Representations: A Comparative Analysis of Neural Language Models and Psychologically Plausible Models. *arXiv preprint arXiv:2404.19364*.

ANEXOS

ANEXO A

Figura “O Encantador de Histórias”



ANEXO B

Critérios de Transcrição

Tabela B1

Critérios de Transcrição

Situação na Gravação	Regra de Transcrição	Exemplo
O que a pessoa diz não é audível	Colocar a informação entre parêntesis	(segmento de texto inaudível)
Não se percebe o que a pessoa diz	Colocar a informação entre parêntesis	Sim, (segmento de texto incompreensível), mas são azuis.
Incerteza sobre o que foi dito	Usar parêntesis delimitados por pontos de interrogação	?(capacidade)?
Erro na palavra dita pelo participante	Transcrever a palavra exata e, entre barras, a proposta	Vida/(via)/
Interrupção ou corte numa palavra	Colocar um travessão no final	Respond-
Sons não verbais	Indicar entre parêntesis	(risos)
Silêncios ou pausas significativas	Usar reticências delimitadas por parêntesis	(...)
Interjeições	Transcrever como na fala	Hmm, Mm, Ah
Formas contraídas ou truncadas	Transcrever conforme o participante diz	Pra (para a); Tá (está)
Palavras repetidas ou reformuladas	Colocar o excesso entre aspas	Vejo «uma» um candelabro «com» com as velas.
Divisão de Enunciados	Cada enunciado em uma nova linha	

APÊNDICES

APENDICE A

Fiabilidade

Tabela A1

Coeficiente de Correlação Intraclass

Medida	Correlação Intraclass	Intervalo de Confiança 95% (Limite Inferior - Limite Superior)	Valor	Teste F com Valor True0	df1	df2	Sig
Medidas Únicas	.989**	.977 - .995	215.000	39	39	39	<.001
Medidas Médias	.995***	.988 - .997	215.000	39	39	39	<.001

APÊNDICE B
Análise de dados

Tabela B1.

Análise da Variância (ANOVA) dos Anos de Escolaridade por Grupo Etário

Grupo Etário	Média (DP)	n
18–39 anos	14.15 (2.29)	142
40–64 anos	12.93 (3.88)	181
65+ anos	8.55 (4.59)	77

ANOVA unidirecional: $F(2, 397) = 63.70, p < .001$.

Tabela B2

Estatística Descritiva

Variável	N	Média	Desvio Padrão	Mínimo	Máximo
Número Enunciados	400	7.9	3.9	2.0	31.0
Extensão Média Enunciado	400	9.3	2.5	3.9	21.5
Tempo Discurso Segundos	400	33.2	18.7	6.0	114.0
Tempo Discurso Minutos	400	.6	.3	.1	1.9
Débito	400	2.3	.6	.8	7.2
Nomes	400	16.9	7.9	4.0	48.0
Nomes Diferentes	400	14.3	5.7	1.0	39.0
Verbos	400	14.4	8.7	1.0	72.0
Verbos Diferentes	400	9.8	4.8	1.0	32.0
Preposições	400	11.9	6.3	1.0	42.0
Preposições Diferentes	400	6.9	3.0	1.0	16.0
Advérbios	400	4.2	4.0	.0	24.0
Advérbios Diferentes	400	3.2	2.5	.0	13.0
Locuções	400	1.6	1.6	.0	10.0
Locuções Diferentes	400	1.5	1.4	.0	8.0
Conjunções	400	4.5	3.4	.0	20.0
Conjunções Diferentes	400	2.5	1.6	.0	9.0
Numerais	400	.1	.5	.0	6.0
Numerais Diferentes	400	.1	.4	.0	5.0
Interjeições	400	1.7	1.9	.0	12.0
Interjeições Diferentes	400	1.3	2.1	.0	34.0
Adjectivos	400	2.0	2.2	.0	13.0
Adjectivos Diferentes	400	1.8	1.9	.0	15.0
Artigos	400	10.8	5.3	.0	38.0
Artigos Diferentes	400	4.0	1.2	1.0	15.0
Pronomes	400	3.2	3.8	.0	35.0
Pronomes Diferentes	400	2.1	1.8	.0	10.0
Determinantes	400	.5	.9	.0	6.0
Determinantes diferentes	400	.4	.8	.0	5.0
Total	400	71.8	37.6	15.0	261.0
Total diferentes	400	46.4	18.8	14.0	116.0
Type token Ratio	400	.7	.1	.4	1.2
Total Palavras Conteúdo Diferentes	400	26.0	11.0	7.0	67.0
T-units simples	400	4.0	2.5	.0	20.0
T-units coordenadas	400	.8	1.0	.0	6.0
T-units subordinadas	400	1.5	1.5	.0	9.0
T-units mistas	400	.5	.8	.0	9.0
Total T-units	400	6.0	4.1	.0	26.0
Total T-Units complexas	400	2.8	2.4	.0	12.0

Tabela B3

Dados normativos por grupo de escolaridade

4-12 anos de Escolaridade

Nomes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	16.23	6.89	5.89	8.0	11.75	15.0	19.5	28.4
40-64	14.53	6.78	4.36	8.0	11.0	13.0	17.25	23.2
65+	18.1	8.56	5.25	10.0	12.0	15.0	21.0	31.0

Nomes Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	14.14	5.39	6.05	8.0	10.0	13.0	17.25	23.4
40-64	12.59	4.92	5.22	7.0	10.0	12.0	15.0	19.1
65+	14.57	5.7	6.03	8.0	11.0	13.0	17.0	21.0

Verbos

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	12.25	5.48	4.03	5.6	9.0	10.5	15.25	18.0
40-64	12.0	9.37	-2.05	4.9	7.0	10.0	14.0	20.1
65+	16.72	9.24	2.87	8.0	11.0	14.0	20.0	29.0

Verbos Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	8.89	3.3	3.94	5.3	6.75	8.5	11.0	13.0
40-64	8.28	4.63	1.34	4.0	5.0	7.0	10.0	13.1
65+	10.48	5.02	2.94	5.0	7.0	10.0	13.0	15.0

Adjetivos

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	1.89	1.97	-1.06	0.0	0.75	1.0	2.25	5.0
40-64	1.39	1.97	-1.57	0.0	0.0	1.0	2.0	3.0
65+	2.0	2.21	-1.32	0.0	0.0	2.0	3.0	4.0

Adjetivos Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	1.84	1.92	-1.03	0.0	0.75	1.0	2.25	4.7
40-64	1.28	1.63	-1.17	0.0	0.0	1.0	2.0	3.0
65+	1.79	1.78	-0.88	0.0	0.0	1.0	3.0	4.0

Advérbios

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	3.25	2.87	-1.06	0.0	1.0	3.0	5.0	7.0
40-64	3.53	4.4	-3.08	0.0	1.0	2.0	5.0	8.1
65+	5.43	4.09	-0.71	1.0	2.0	4.0	8.0	11.0

Advérbios Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	2.57	2.07	-0.54	0.0	1.0	2.0	4.0	5.0

40–64	2.64	2.56	-1.21	0.0	1.0	2.0	4.0	6.0
65+	3.72	2.51	-0.04	1.0	2.0	3.0	5.0	7.0

Pronomes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	2.43	3.24	-2.42	0.0	1.0	2.0	3.0	4.0
40–64	2.73	3.6	-2.67	0.0	0.0	2.0	4.0	6.1
65+	3.95	3.55	-1.37	0.0	1.0	3.0	6.0	8.0

Pronomes Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	1.5	1.09	-0.13	0.0	1.0	1.5	2.0	3.0
40–64	1.88	1.9	-0.97	0.0	0.0	1.0	3.0	4.0
65+	2.44	1.77	-0.21	0.0	1.0	2.0	3.0	4.0

Preposições

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	11.43	5.67	2.93	6.0	7.0	10.0	14.0	19.0
40–64	9.91	5.52	1.63	4.0	6.0	9.0	13.0	16.0
65+	12.11	5.84	3.35	5.0	9.0	10.0	15.0	21.0

Preposições Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	6.7	2.63	2.76	4.0	5.0	6.0	8.25	10.7
40–64	6.09	2.75	1.97	3.0	4.0	5.0	8.0	10.0
65+	6.92	2.62	2.98	4.0	5.0	7.0	9.0	10.0

Conjunções

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	3.7	2.73	-0.39	1.0	2.0	3.0	5.0	7.7
40–64	3.89	3.21	-0.92	1.0	2.0	3.0	5.0	7.1
65+	5.28	3.34	0.27	1.0	3.0	5.0	7.0	9.0

Conjunções Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	2.25	1.4	0.15	1.0	1.0	2.0	3.0	4.0
40–64	2.29	1.49	0.05	1.0	1.0	2.0	3.0	4.0
65+	2.75	1.53	0.45	1.0	2.0	3.0	4.0	5.0

Interjeições

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	1.66	2.02	-1.37	0.0	0.0	1.0	2.0	3.7
40–64	1.28	1.67	-1.23	0.0	0.0	1.0	2.0	3.1
65+	1.72	1.57	-0.64	0.0	1.0	1.0	3.0	3.0

Interjeições Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	1.16	1.2	-0.64	0.0	0.0	1.0	2.0	2.7
40–64	1.03	1.27	-0.88	0.0	0.0	1.0	2.0	3.0
65+	1.51	1.29	-0.42	0.0	1.0	1.0	2.0	3.0

Artigos

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	10.09	4.02	4.07	6.0	7.75	10.0	11.25	16.7
40-64	9.45	4.84	2.18	5.0	6.0	8.5	11.0	14.0
65+	12.15	6.26	2.76	6.0	8.0	10.0	15.0	20.0

Artigos Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	4.02	1.89	1.19	3.0	3.0	4.0	4.0	5.0
40-64	3.87	1.07	2.27	3.0	3.0	4.0	5.0	5.0
65+	4.07	1.05	2.5	3.0	3.0	4.0	5.0	5.0

Determinantes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	0.45	0.73	-0.64	0.00	0.00	0.00	1.00	1.00
40-64	0.35	0.64	-0.61	0.00	0.00	0.00	1.00	1.00
65+	0.44	0.65	-0.53	0.00	0.00	0.00	1.00	1.00

Determinantes Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	0.43	0.70	-0.61	0.00	0.00	0.00	1.00	1.00
40-64	0.33	0.62	-0.60	0.00	0.00	0.00	1.00	1.00
65+	0.43	0.59	-0.46	0.00	0.00	0.00	1.00	1.00

Numerais

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	0.14	0.51	-0.63	0.00	0.00	0.00	0.00	0.00
40-64	0.05	0.22	-0.28	0.00	0.00	0.00	0.00	0.00
65+	0.28	0.97	-1.17	0.00	0.00	0.00	0.00	1.00

Numerais Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	0.11	0.39	-0.47	0.00	0.00	0.00	0.00	0.00
40-64	0.05	0.22	-0.28	0.00	0.00	0.00	0.00	0.00
65+	0.25	0.83	-1.00	0.00	0.00	0.00	0.00	1.00

Locuções

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	1.66	1.35	-0.36	0.00	1.00	1.00	2.00	3.70
40-64	1.54	1.47	-0.67	0.00	0.75	1.00	2.00	3.00
65+	1.85	1.97	-1.11	0.00	1.00	1.00	2.00	5.00

Locuções Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	1.41	1.13	-0.28	0.00	1.00	1.00	2.00	2.00
40-64	1.42	1.27	-0.49	0.00	0.75	1.00	2.00	3.00
65+	1.61	1.68	-0.91	0.00	1.00	1.00	2.00	3.00

Total Palavras

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	65.18	27.4	24.08	36.3	45.0	57.5	80.25	112.5

40–64	60.65	37.56	4.31	28.0	39.0	52.5	69.5	100.3
65+	80.03	39.0	21.53	41.0	53.0	72.0	93.0	133.0

Total palavras diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	43.61	15.69	20.08	26.0	33.5	40.0	52.5	65.1
40–64	40.33	17.93	13.44	21.0	27.75	36.5	50.0	63.0
65+	48.92	17.9	22.07	31.0	38.0	47.0	56.0	70.0

Numero Enunciados

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	6.82	2.59	2.93	4.0	5.0	7.0	8.0	10.0
40–64	7.09	3.94	1.18	4.0	5.0	6.0	8.0	12.0
65+	9.23	4.5	2.48	5.0	6.0	9.0	11.0	17.0

Tempo Discurso Segundos

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	29.27	14.25	7.9	12.9	19.75	26.0	36.5	46.4
40–64	28.8	18.4	1.21	14.12	17.0	23.0	35.25	48.2
65+	40.42	20.59	9.53	20.0	25.0	38.0	50.0	65.0

Débito

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	2.43	0.84	1.17	1.66	2.0	2.24	2.56	3.12
40–64	2.22	0.62	1.28	1.6	1.86	2.17	2.53	2.81
65+	2.07	0.58	1.21	1.45	1.81	2.0	2.33	2.72

Total Palavras Conteúdo Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	24.98	9.57	10.63	14.6	17.75	23.0	29.75	39.0
40–64	22.2	9.78	7.53	12.0	16.0	20.0	26.0	33.1
65+	27.08	11.1	10.44	16.0	20.0	26.0	31.0	41.0

T-units simples

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	4.02	2.11	0.86	2.0	3.0	4.0	5.0	6.7
40–64	3.61	2.43	-0.04	1.0	2.0	3.0	4.25	6.1
65+	4.3	2.57	0.44	1.0	3.0	4.0	5.0	8.0

T-units com coordenação

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	0.59	0.73	-0.5	0.0	0.0	0.0	1.0	2.0
40–64	0.57	0.81	-0.64	0.0	0.0	0.0	1.0	2.0
65+	0.93	0.95	-0.49	0.0	0.0	1.0	1.0	2.0

T-units com subordinação

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	1.57	1.17	-0.19	0.0	1.0	1.5	2.0	3.0
40–64	1.22	1.31	-0.74	0.0	0.0	1.0	2.0	3.0
65+	1.72	1.88	-1.1	0.0	0.0	1.0	3.0	4.0

T-units mistas

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	0.25	0.53	-0.55	0.0	0.0	0.0	0.0	1.0
40–64	0.47	1.07	-1.13	0.0	0.0	0.0	1.0	1.0
65+	0.52	0.62	-0.41	0.0	0.0	0.0	1.0	1.0

Total T-units

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	5.14	2.5	1.38	2.3	3.75	5.0	6.0	8.7
40–64	5.11	3.97	-0.85	1.9	3.0	4.0	6.0	9.1
65+	6.44	4.2	0.14	3.0	4.0	5.0	8.0	11.0

Total T-Units complexas

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	2.41	1.65	-0.06	0.3	1.0	2.0	3.0	4.7
40–64	2.26	2.26	-1.13	0.0	1.0	2.0	3.0	5.0
65+	3.18	2.54	-0.63	1.0	1.0	3.0	4.0	6.0

+ 12 anos Escolaridade

Nomes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	18.81	8.55	5.98	10.0	13.0	17.0	22.0	29.7
40–64	16.65	8.1	4.51	7.0	10.0	17.0	21.0	27.0
65+	21.62	6.53	11.83	14.0	18.75	20.5	23.5	29.5

Nomes Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	15.93	6.23	6.58	9.0	12.0	15.0	19.0	22.7
40–64	13.65	5.69	5.13	6.0	9.0	13.0	17.0	21.0
65+	18.12	4.86	10.84	14.0	15.75	16.5	18.5	24.5

Verbos

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	15.85	8.57	3.0	6.3	10.0	14.5	20.0	25.7
40–64	14.4	8.39	1.81	6.0	8.0	12.0	18.0	26.0
65+	19.88	7.37	8.83	11.0	16.5	20.0	25.0	25.5

Verbos Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	10.71	5.14	3.0	5.0	7.0	10.0	13.0	17.0
40–64	10.01	4.88	2.69	5.0	6.0	9.0	12.0	17.0
65+	12.62	4.51	5.85	6.0	9.75	13.5	15.0	18.0

Adjetivos

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	2.32	2.45	-1.35	0.0	1.0	1.5	3.0	5.0
40–64	1.99	1.65	-0.49	0.0	1.0	2.0	3.0	4.0
65+	4.06	3.02	-0.47	1.0	2.0	3.5	5.25	6.5

Adjetivos Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	2.02	1.96	-0.92	0.0	1.0	1.0	3.0	5.0
40–64	1.85	1.49	-0.39	0.0	1.0	2.0	3.0	4.0
65+	4.19	3.43	-0.96	1.0	2.0	3.5	5.25	6.5

Advérbios

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	3.77	3.69	-1.77	0.0	1.0	3.0	5.0	7.7
40–64	4.79	4.05	-1.29	1.0	2.0	4.0	6.0	11.0
65+	6.94	3.55	1.61	1.5	4.75	8.0	9.25	11.0

Advérbios Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	3.01	2.37	-0.54	0.0	1.0	2.5	4.0	6.0
40–64	3.63	2.71	-0.43	1.0	2.0	3.0	5.0	8.0
65+	5.0	2.22	1.67	1.5	4.5	5.0	6.25	7.5

Pronomes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	3.36	3.4	-1.74	0.0	1.0	2.0	5.0	8.0
40-64	2.9	2.94	-1.51	0.0	1.0	2.0	4.0	7.0
65+	7.0	8.33	-5.5	1.0	3.0	4.5	9.25	11.5

Pronomes Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	2.18	1.98	-0.79	0.0	1.0	2.0	3.0	4.7
40-64	2.16	1.72	-0.42	0.0	1.0	2.0	3.0	4.0
65+	3.25	2.32	-0.24	0.0	2.0	3.0	5.0	6.0

Preposições

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	13.74	7.14	3.04	7.0	9.0	12.0	16.0	21.0
40-64	11.69	6.21	2.38	4.0	7.0	11.0	15.0	20.0
65+	14.31	5.53	6.01	9.0	10.0	13.0	16.5	23.5

Preposições Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	7.9	3.12	3.22	4.0	5.25	7.5	9.0	12.0
40-64	6.57	3.32	1.58	3.0	4.0	6.0	9.0	11.0
65+	7.5	2.66	3.51	4.0	5.75	8.0	9.0	10.5

Conjunções

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	4.56	3.49	-0.68	1.0	2.0	4.0	6.0	9.0
40-64	4.69	3.7	-0.86	1.0	2.0	3.0	6.0	9.0
65+	6.81	3.23	1.97	3.5	4.0	7.0	8.5	10.5

Conjunções Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	2.48	1.55	0.15	1.0	1.0	2.0	3.0	4.7
40-64	2.58	1.6	0.17	1.0	2.0	2.0	3.0	5.0
65+	3.69	2.06	0.6	1.5	2.0	3.0	5.0	6.5

Interjeições

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	1.7	1.59	-0.69	0.0	0.25	1.0	3.0	4.0
40-64	1.93	2.63	-2.02	0.0	0.0	1.0	3.0	5.0
65+	2.5	2.19	-0.79	0.0	1.0	2.0	4.25	5.0

Interjeições Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	1.33	1.19	-0.45	0.0	0.25	1.0	2.0	3.0
40-64	1.27	1.41	-0.84	0.0	0.0	1.0	2.0	3.0
65+	3.81	8.18	-8.46	0.0	1.0	2.0	2.5	4.5

Artigos

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	11.65	5.67	3.14	5.3	8.0	10.0	14.0	19.7
40-64	10.35	4.67	3.33	5.0	7.0	10.0	12.0	17.0
65+	13.0	5.4	4.89	5.5	10.75	14.0	15.0	19.0

Artigos Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	4.0	0.95	2.57	3.0	3.0	4.0	5.0	5.0
40-64	4.04	1.08	2.42	3.0	3.0	4.0	5.0	6.0
65+	4.25	1.13	2.56	3.0	3.75	4.0	5.0	5.5

Determinantes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	0.33	0.87	-0.98	0.00	0.00	0.00	0.00	1.00
40-64	0.65	1.21	-1.15	0.00	0.00	0.00	1.00	2.00
65+	0.75	0.77	-0.41	0.00	0.00	1.00	1.00	2.00

Determinantes Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	0.32	0.81	-0.89	0.00	0.00	0.00	0.00	1.00
40-64	0.59	1.01	-0.92	0.00	0.00	0.00	1.00	2.00
65+	0.75	0.77	-0.41	0.00	0.00	1.00	1.00	2.00

Numerais

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	0.16	0.40	-0.43	0.00	0.00	0.00	0.00	1.00
40-64	0.07	0.26	-0.32	0.00	0.00	0.00	0.00	0.00
65+	0.12	0.34	-0.39	0.00	0.00	0.00	0.00	0.50

Numerais Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	0.16	0.40	-0.43	0.00	0.00	0.00	0.00	1.00
40-64	0.09	0.28	-0.34	0.00	0.00	0.00	0.00	0.00
65+	0.12	0.34	-0.39	0.00	0.00	0.00	0.00	0.50

Locuções

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	1.71	1.68	-0.80	0.00	1.00	1.00	3.00	4.00
40-64	1.47	1.45	-0.71	0.00	0.00	1.00	2.00	4.00
65+	2.06	1.81	-0.65	0.50	1.00	1.00	3.25	4.50

Locuções Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	1.51	1.40	-0.58	0.00	1.00	1.00	2.00	3.00
40-64	1.35	1.31	-0.61	0.00	0.00	1.00	2.00	3.00
65+	1.69	1.40	-0.41	0.50	1.00	1.00	2.25	3.50

Total Palavras

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	77.97	39.07	19.36	36.3	54.25	69.0	94.5	134.8
40-64	71.58	36.57	16.73	26.0	48.0	69.0	91.0	129.0
65+	99.06	30.6	53.17	65.0	84.5	95.0	113.25	136.5

Total Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	50.04	19.37	20.99	28.0	35.25	45.5	59.0	77.2
40-64	46.44	18.82	18.22	21.0	35.0	45.0	57.0	73.0
65+	63.31	17.97	36.36	46.5	51.75	59.5	70.5	89.0

Numero Enunciados

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	8.01	3.58	2.65	4.0	6.0	7.0	9.75	13.0
40-64	7.89	3.98	1.92	4.0	5.0	7.0	10.0	13.0
65+	10.0	3.46	4.8	6.5	7.0	10.5	12.25	13.5

Tempo Discurso Segundos

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	32.61	18.09	5.48	14.3	21.0	29.0	39.0	54.7
40-64	33.49	18.73	5.39	14.0	21.0	30.0	41.0	54.0
65+	48.69	13.31	28.73	35.0	38.0	47.0	58.0	68.0

Débito

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	2.49	0.68	1.48	1.96	2.13	2.38	2.74	2.91
40-64	2.22	0.47	1.52	1.73	1.9	2.21	2.46	2.86
65+	2.06	0.48	1.34	1.43	1.77	2.12	2.29	2.63

Total Palavras Conteúdo Diferentes

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	28.82	11.91	10.95	15.3	21.0	27.0	34.75	40.0
40-64	25.6	10.27	10.2	12.0	19.0	25.0	32.0	38.0
65+	35.06	10.18	19.79	25.5	29.0	33.5	39.5	47.5

T-units simples

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	4.11	2.32	0.63	1.0	2.0	4.0	6.0	7.0
40-64	3.84	3.08	-0.79	1.0	2.0	3.0	5.0	8.0
65+	5.25	2.65	1.28	2.5	3.0	5.0	7.0	9.5

T-units com coordenação

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18-39	1.0	1.35	-1.03	0.0	0.0	1.0	1.0	2.7
40-64	0.7	1.01	-0.8	0.0	0.0	0.0	1.0	2.0
65+	1.25	0.93	-0.15	0.0	1.0	1.0	2.0	2.5

T-units com subordinação

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	1.63	1.41	-0.49	0.0	1.0	1.0	2.0	3.7
40–64	1.54	1.45	-0.63	0.0	0.0	1.0	2.0	4.0
65+	2.5	1.83	-0.24	0.5	1.0	2.5	4.0	5.0

T-units mistas

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	0.54	0.79	-0.64	0.0	0.0	0.0	1.0	2.0
40–64	0.48	0.81	-0.73	0.0	0.0	0.0	1.0	1.0
65+	0.94	1.06	-0.66	0.0	0.0	1.0	1.0	2.0

Total T-units

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	6.51	4.0	0.51	3.0	4.0	6.0	8.0	10.7
40–64	5.78	4.3	-0.68	2.0	3.0	5.0	8.0	10.0
65+	9.44	5.54	1.13	5.0	6.0	7.5	10.25	18.0

Total T-Units complexas

Grupo Etário	Média	DP	Média - 1.5DP	P10	P25	P50	P75	P90
18–39	3.17	2.43	-0.48	0.3	1.25	2.5	4.75	6.0
40–64	2.73	2.41	-0.88	0.0	1.0	2.0	4.0	6.0
65+	4.69	2.87	0.38	1.5	3.0	4.5	6.25	8.0

Tabela B4

Resultados do t-teste por Grupo de Escolaridade

Variável Linguística	<i>t</i>	<i>P</i>
Extensão Média Enunciado	-2.599	0.010
Nomes	-2.842	0.005
Verbos	-2.488	0.013
Preposições	-3.262	0.001
Interjeições	-1.974	0.049
Adjectivos	-3.053	0.002
Total	-2.661	0.008
Total Palavras Conteúdo Diferentes	-3.53	< .001
T-units coordenadas	-2.165	0.031
Total T-units	-2.311	0.021
Total T-Units complexas	-2.401	0.017

Tabela B5

Resultados da ANOVA (variáveis com diferenças entre grupos etários)

Variável Linguística	Grupos Etários Comparados	F (gl)	<i>p</i>
Número Enunciados	Grupo 1, Grupo 2, Grupo 3	7.57 (2, 397)	< .001
Extensão média dos enunciados	Grupo 1, Grupo 2, Grupo 3	4.45 (2, 397)	0.012
Débito verbal	Grupo 1, Grupo 2, Grupo 3	11.87 (2, 397)	< .001
Nomes diferentes	Grupo 1, Grupo 2, Grupo 3	6.26 (2, 397)	0.002
Verbos diferentes	Grupo 1, Grupo 2, Grupo 3	6.81 (2, 397)	0.001

Tabela B6

Comparações Post-hoc de Tukey entre Grupos Etários para Variáveis Linguísticas com Diferenças Significativas

Variável	Comparação de Grupos	Diferença Média	p	IC 95%
Número Enunciados	40–64 vs 65+	-1,94	< .001	[-3,16 ; -0,72]
Número Enunciados	18–39 vs 65+	-1,79	.003	[-3,06 ; -0,52]
Extensão Média Enunciado	18–39 vs 40–64	0,83	.009	[0,17 ; 1,49]
Débito	18–39 vs 65+	0,40	< .001	[0,19 ; 0,61]
Débito	18–39 vs 40–64	0,25	.001	[0,09 ; 0,42]
Nomes Diferentes	18–39 vs 40–64	2,17	.002	[0,68 ; 3,65]
Nomes Diferentes	40–64 vs 65+	-2,25	.010	[-4,04 ; -0,45]
Verbos Diferentes	40–64 vs 65+	-1,87	.013	[-3,40 ; -0,33]
Verbos Diferentes	65+ vs 40–64	1,87	.013	[0,33 ; 3,40]

Tabela B7

Resultados das Regressões Lineares Múltiplas por Grupos (Escolaridade e Idade)

Variável dependente: T-Units Complexas

Preditor	Coeficiente (β)	Erro padrão (SE)	p
Grupo Escolaridade (2)	0.76	0.24	0.002
Grupo Etário número (2)	-0.24	0.27	0.370
Grupo Etário número (3)	0.97	0.35	0.006

$R^2 = 0.048$

Variável dependente: Palavras de Conteúdo Diferentes

Preditor	Coeficiente (β)	Erro padrão (SE)	p
Grupo Escolaridade (2)	4.40	1.12	< 0.001
Grupo Etário número (2)	-2.68	1.21	0.028
Grupo Etário número (3)	3.39	1.58	0.033

$R^2 = 0.073$