



João Victor de Oliveira Miranda e Silva

**Pronomes em esquizofrenia: análise de textos escritos no
contexto de mídia social**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para
obtenção do grau de Mestre em Letras/Estudos
Estudos da Linguagem pelo Programa de
Pós-graduação em Estudos da Linguagem da
PUC-Rio.

Orientadora: Cilene Aparecida Nunes Rodrigues

Coorientador: Adriano Alonso Veloso

Rio de Janeiro
Março 2022



João Victor de Oliveira Miranda e Silva

**Pronomes em esquizofrenia: análise de textos escritos no
contexto de mídia social**

Dissertação apresentada como requisito parcial para
obtenção do grau de Mestre pelo Programa de
Pós-graduação em Estudos da Linguagem da
PUC-Rio. Aprovada pela Comissão Examinadora
abaixo:

Cilene Aparecida Nunes Rodrigues
Orientadora
Departamento de Letras – PUC-Rio

Adriano Alonso Veloso
Coorientador
UFMG

Rafael Moreno Ferro de Araújo
Resting Mind Clínica de Psiquiatria & Saúde

Eduardo Kenedy Nunes Areas
UFF

Rio de Janeiro, 21 de março de 2022.

Todos os direitos reservados. A reprodução, total ou parcial, do trabalho é proibida sem autorização da universidade, da autora e do orientador.

João Victor de Oliveira Miranda e Silva

Graduou-se em Letras - Bacharelado Bilíngue (Português e Inglês) pela PUC-Rio em 2018. Mestrando do curso Estudos da Linguagem pela PUC-Rio (2019). Atua no Grupo de Pesquisa em Teoria Sintática, projeto comparativo de estruturas linguísticas coordenado por Cilene Rodrigues, sob a linha "Referência Pronominal: Construção e Desconstrução", que versa sobre a estruturação da referência de pronomes pela gramática e sua desestruturação frente a distúrbios mentais do construto da esquizotipia, em especial a esquizofrenia.

Ficha Catalográfica

Silva, João Victor de Oliveira Miranda e

Pronomes em esquizofrenia : análise de textos escritos no contexto de mídia social / João Victor de Oliveira Miranda e Silva ; orientadora: Cilene Rodrigues ; co-orientador: Adriano Veloso. – 2022.
137 f. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Letras, 2022.
Inclui bibliografia

1. Letras – Teses. 2. Esquizofrenia. 3. Pronomes. 4. Truncamento sentencial. 5. Anomalia referencial. 6. Complexidade sintática. I. Rodrigues, Cilene. II. Veloso, Adriano. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Letras. IV. Título

CDD: 400

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Gostaria de agradecer, antes de tudo, à minha orientadora, Cilene Rodrigues, e ao meu co-orientador, Adriano Veloso, por terem sido pilares firmes e fortes durante todas as etapas desta dissertação. Através da nossa parceira, viabilizamos esta pesquisa, da qual tenho muito orgulho, e obtive uma experiência inimaginável, que foi passada a mim com muita responsabilidade. Obrigado!

Gostaria também de oferecer os meus mais sinceros agradecimentos a Sílvia Guerra, da Ciência da Computação (UFMG), que foi uma parceira de trabalho incrível e fundamental para esta pesquisa.

A todas as outras pessoas queridas envolvidas direta e indiretamente no processo desta pesquisa, seja por apoio acadêmico ou afetivo: muito obrigado, cada um tem uma parte importante na conclusão deste trabalho.

RESUMO

Silva, João Victor de Oliveira Miranda e; Rodrigues, Cilene (Orientadora); Veloso, Adriano (Coorientador). **Pronomes em esquizofrenia: análise de textos escritos no contexto de mídia social**. Rio de Janeiro, 2022. 137p. Dissertação de Mestrado - Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro.

A presente dissertação é um estudo teórico-experimental sobre linguagem na esquizofrenia, com foco no uso de pronomes em textos escritos por nativos do inglês, disponíveis na plataforma *Reddit*, com análise adicional da complexidade sintática no nível sentencial. A esquizofrenia é um distúrbio mental que se apresenta como uma heterogeneidade de disfunções neuropsicológicas, sendo uma de suas principais características disfunções linguísticas, em especial no uso e compreensão de expressões nominais e na redução da complexidade sintática. Foram coletados, analisados e comparados um total de 78 textos (39 de usuários com autodeclaração de esquizofrenia e 39 de controles pareados). Os textos coletados foram anotados sintaticamente e tratados estatisticamente considerando fatores relativos à complexidade sintática no nível sentencial e pronominal. Comparações intergrupais revelaram diferenças significativas no domínio sentencial e pronominal, com o grupo esquizofrenia apresentando menor volume de truncamento sentencial não-anômalo e maior volume de anomalias no emprego de pronomes plenos de terceira pessoa com leitura referencial. Estes resultados estão de acordo com as conclusões de estudos anteriores. Partindo do entendimento de que a Gramática é um mecanismo computacional cognitivo, as observações obtidas indicam que a esquizofrenia leva a uma redução na complexidade sintática de processos gramaticais responsáveis pela construção do significado.

PALAVRAS-CHAVE

Esquizofrenia; Pronomes; Truncamento sentencial; Anomalia referencial; Complexidade sintática.

ABSTRACT

Silva, João Victor de Oliveira Miranda e; Rodrigues, Cilene (Advisor); Veloso, Adriano (Co-advisor). **Pronouns in schizophrenia: written-texts analysis on social media**. Rio de Janeiro, 2022. 137p. Dissertação de Mestrado - Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro.

The present dissertation is a theoretical-experimental study on language in schizophrenia, focusing on the usage of pronouns in written texts produced by native speakers of English and published on *Reddit*, considering in addition syntactic complexity at the sentential level. Schizophrenia is a mental disorder characterized by a set of heterogeneous dysfunctions, with language-malfunctioning being one of its main characteristics. In particular, misuse and misinterpretation of nominal expressions and syntactic-complexity reduction are linguistic hallmarks of schizophrenia. 78 texts (39 from *Reddit* users with self-declaration of schizophrenia and 39 from paired controls) were automatically collected, manually tagged and statistically compared. The tagging process was based on syntactic complexity at the sentential and nominal level. Between-group statistical comparisons showed significant differences, with the schizophrenic group exhibiting less non-anomalous sentential truncations (ellipsis) and more anomalous 3Person referential pronouns. This is in accordance with previous independent studies. Assuming human Grammar to be a computational apparatus of the cognitive system, we interpret the obtained results as suggesting that schizophrenia leads to reductions on meaning-building derivational grammatical processes.

KEYWORDS

Schizophrenia; Pronouns; Sentential truncation; Referential anomaly; Syntactic complexity.

Sumário

1.	Introdução	11
1.1.	A Gramática como aparato cognitivo	11
1.2.	Gramática e esquizofrenia	15
1.3.	Objetivos, perguntas norteadoras e hipóteses de trabalho	17
1.4.	Metodologia	18
1.5.	Justificativa	19
1.6.	Organização da dissertação	20
2.	Esquizofrenia	21
2.1.	Caracterização geral da esquizofrenia	21
2.2.	Esquizofrenia: prejuízos cognitivos	24
2.3.	Esquizofrenia e linguagem	26
2.4.	Alterações linguísticas	28
2.4.1.	Distúrbios fonológicos, morfológicos e sintáticos	28
2.4.2.	Distúrbios semânticos e pragmáticos	33
2.5.	Esquizofrenia e memória de trabalho	37
3.	Pronomes em gramática típica e atípica	40
3.1.	Expressões nominais, referência e pronomes	40
3.2.1.	Pronomes: interface sintaxe-semântica	43
3.2.2.	Tipos de pronomes, estrutura sintática e interpretabilidade	46
3.2.3.	Pronomes em inglês	49
3.3.	Pronomes em gramáticas atípicas	52
3.3.1.	Envelhecimento atípico	52
3.3.2.	Desenvolvimento atípico	54
3.3.3.	Pronomes em esquizofrenia	56
3.4.	Pronomes na escrita	60
4.	Pronomes em esquizofrenia: análise de textos escritos no contexto de mídia social	64

4.1.	Método de coleta de dados	65
4.2.	Participantes	69
4.3.	Análise dos dados	70
4.3.1.	Parâmetros de análise	70
4.3.2.	Esquema de anotação morfossintática (etiquetagem)	71
4.3.3.	Exemplos da amostra final	73
4.3.4.	Análise estatística	75
4.4.	Resultados	76
4.5.	Discussão dos resultados	81
5.	Conclusão	87
6.	Referências bibliográficas	90
7.	Anexo 1 - <i>flairs</i> de autodeclaração (esquizofrenia)	104
8.	Anexo 2 - <i>flairs</i> excludentes	106
9.	Anexo 3 - lista lexical excludente (transtornos mentais)	111
10.	Anexo 4 - lista lexical (idade)	115
11.	Anexo 5 - lista lexical (sexo)	117
12.	Anexo 6 - lista lexical (nativos)	120

Lista de figuras e tabelas

Figura 1 - Estrutura do DP proposta por Abney (1987)	41
Figura 2 - Diferença estrutural - pronomes fortes (<i>Strong Pronoun</i>), pronomes fracos (<i>Weak Pronoun</i>) e clíticos (<i>Clitic</i>)	47
Figura 3 - Exemplo de <i>flair</i> (<i>Schizophrenia</i> em verde)	65
Figura 4 - Gráfico com distribuição (em quantidade numérica e percentual) dos textos que pertencem ao grupo de 100 a 500 palavras e textos que ultrapassaram o limite de 500 palavras	67
Figura 5 - Número de postagens coletadas por cada <i>subreddit</i> selecionado (grupo alvo)	68
Figura 6 - Número de postagens coletadas por cada <i>subreddit</i> selecionado (grupo controle)	68
Figura 7 - Comparações entre grupos da média de sentenças matrizes truncadas (MS+TRUNCTotal) e sentenças matrizes truncadas não-anômalas (MS+TRUNC-A)	76
Figura 8 - Comparações entre grupos da média de pronomes plenos referenciais de terceira pessoa anômalos (O+R-L3P+A)	78
Figura 9 - Comparações entre grupos da média de sentenças matrizes truncadas não-anômalas (MS+TRUNC-A)	79
Figura 10 - Comparações entre grupos da média de pronomes plenos referenciais de terceira pessoa anômalos (O+R-L3P+A)	81
Tabela 1 - Distribuição por sexo (grupo alvo e controle)	69
Tabela 2 - Distribuição por idade (grupo alvo e controle)	69
Tabela 3 - Média e desvio padrão por idades (grupo alvo e controle)	70
Tabela 4 - Distribuição por declarações de ansiedade e depressão (grupo alvo e controle)	70
Tabela 5 - Comparações Mann-Whitney <i>U</i> entre grupos (amostra total) para categorias do nível da sentença	76
Tabela 6 - Comparações de Mann-Whitney <i>U</i> entre grupos (amostra total) para categorias do nível pronominal	77

Tabela 7 - Comparações de Mann-Whitney U entre grupos (amostra parcial) para categorias do nível da sentença	78
Tabela 8 - Comparações de Mann-Whitney U entre grupos (amostra parcial) para categorias do nível pronominal	79

1

Introdução

Esta dissertação compõe o projeto “Referência Pronominal: Construção e Desconstrução”, coordenado pela pesquisadora Cilene Rodrigues da PUC-Rio, e financiado pelo CNPq - Edital Humanidades/2018. O projeto em questão possui dois braços de investigação: um teórico, que busca, por meio da teoria da gramática, investigar a divisão de trabalho entre pronomes plenos e nulos em línguas orais e de sinais, estabelecendo generalizações translinguísticas sobre o comportamento sintático, semântico e pragmático desses itens lexicais; outro, de natureza interdisciplinar, investiga em que medida a estruturação sintática de pronomes se desestrutura em casos de transtornos mentais, em especial aqueles relativos ao construto da esquizotipia, particularmente a esquizofrenia. Nossa proposta de pesquisa se insere no segundo braço de investigação. Concomitantemente, a proposta é de caráter interdisciplinar, fazendo fronteira com a área da Ciência da Computação¹, por meio de parceria de pesquisa com o Laboratório de Inteligência Artificial do Departamento de Ciências da Computação da UFMG, coordenado pelo pesquisador Adriano Veloso, coorientador da presente investigação. O objetivo dessa parceria é o emprego de técnicas de aprendizado de máquina para coleta e análise de dados, nos possibilitando desenvolver um banco de dados linguísticos e um modelo qualitativo de *tagging* de estruturas linguísticas que pode ser utilizado no treinamento de algoritmos de aprendizado de máquina em processamento de Linguagem Natural.

1.1.

Gramática: um aparato cognitivo

A linguagem pode ser considerada um componente cognitivo humano, biologicamente inato. A partir desta concepção, ela pode ser interpretada como fruto do desenvolvimento de uma capacidade cognitiva computacional

¹ Importante ressaltar que, apesar de a Ciência da Computação trabalhar com um padrão de metodologia que envolve análise de dados em larga escala, em nosso trabalho utilizamos suas técnicas para coleta e filtragem de uma larga escala de dados, mas nossa amostra é composta por um n reduzido, porém robusto para investigações dentro do escopo da Linguística Formal.

compartilhada pelos seres humanos. Discussões nesta área apontam que a linguagem é uma característica particularmente marcante da espécie humana, diferenciando-a de outras espécies (Chomsky, 2005; Berwick & Chomsky, 2016; Chomsky, 2017).

Essa visão teórica acerca do fenômeno da linguagem compreende uma corrente de estudos denominada gerativismo, ou linguística gerativa, que teve início no final da década de 50 e procura elaborar um modelo teórico formal, abstrato, capaz de descrever e explicar o que é e como funciona a linguagem humana (Chomsky, 1965; Kenedy, 2008). De acordo com Kenedy (2008), o gerativismo entende que o dispositivo cognitivo linguístico da espécie humana é destinado a constituir o conhecimento linguístico de um falante acerca de determinada(s) língua(s), compreendendo, assim, a *Faculdade da Linguagem*. O gerativismo interpreta as línguas humanas não como um comportamento socialmente condicionado, mas como uma faculdade mental. A partir dessa concepção de base, o papel do teórico gerativista é constituir um modelo teórico capaz de descrever e explicar como se dá o funcionamento dessa Faculdade, ou seja, como ela funciona e a natureza do seu caráter geneticamente determinado, visto que as línguas do mundo diferem entre si.

Em termos de desenvolvimento do indivíduo, a linguagem compreende um estado mental inicial que engatilha o processo de aquisição de uma língua específica a partir de um conjunto de propriedades universais e invariantes sob as quais qualquer língua humana está submetida (Chomsky, 1981; 1986), denominado Gramática Universal (GU). A estas regras universais, então, associa-se um arranjo finito de parâmetros que marcam as diferenças entre as diversas línguas observadas no mundo. Esses parâmetros estão possivelmente associados aos diferentes feixes de traços das categorias funcionais que compõem o léxico humano (Chomsky, 1995). A fixação desses traços durante o processo de aquisição da linguagem marca o estágio mental final do processo de aquisição, a gramática na mente do falante adulto.

Esta teoria é denominada Teoria de Princípios e Parâmetros e representa um avanço nos estudos acerca do entendimento da linguagem humana, oferecendo uma explicação plausível que captura ambas as semelhanças e diferenças entre as línguas através de uma teoria formal, ou seja, oferece um modelo explicativo para análises empíricas que procuram resposta para o famoso “Problema de Platão”,

postulado a partir da observação de que crianças adquirem suas línguas maternas sem esforços e, ao mesmo tempo, de forma bem-sucedida (em condições normais) sob as condições de pobreza do estímulo. De acordo com Hauser et al. (2002), a grande variedade de sentenças que qualquer falante de uma língua pode produzir e entender tem uma implicação importante para o entendimento da linguagem humana a partir do fenômeno da aquisição. Uma criança é exposta a uma fração pequena de sentenças possíveis em sua língua, limitando, então, sua base de dados para a construção de uma versão geral daquela língua em sua mente. Nesse contexto, a tarefa da criança de aprender uma língua, baseando-se apenas no material linguístico encontrado em seu ambiente, parece árdua, mas, na verdade, mostra-se natural e, por padrão, sem esforço. Isso leva à hipótese de que existem disposições inatas para o sistema cognitivo encarregado da aquisição linguística. Essas características inatas restringem a variedade de possíveis gramáticas para que a criança consiga internalizar, precisamente, a sua língua materna. A este sistema, Hauser et al. (2002) denominam Gramática Universal e, portanto, apresentam uma abordagem ao Problema de Platão baseada na hipótese de que uma criança possui conhecimento inato gramatical devido à carga biológica que define a Gramática Universal. Assim, ao ser exposta a um dado linguístico primário (comportamento verbal de terceiros, por exemplo), desenvolve uma gramática particular, que é reconhecida em um determinado contexto cultural como uma determinada língua (Roberts, 2019).

Através dessa perspectiva o conhecimento de uma língua é entendido como a fixação de valores paramétricos deixados em aberto por um componente cognitivo inato da espécie humana e esse processo ocorre a partir da experiência determinada por um *input* linguístico primário. Propriedades não-universais que caracterizam a gama de diferenças que existem entre as línguas no mundo podem ser, por exemplo, o fenômeno do movimento *qu-* (*-wh*) e a falta de marcação morfológica de Caso. Esse conjunto de propriedades, então, deve ser internalizado a partir do *input* linguístico aos quais uma criança na fase de aquisição da linguagem é submetida no ambiente em que vive. Além dos parâmetros a serem fixados, o componente cognitivo (GU) faz com que todas as línguas humanas sejam regidas por princípios universais compartilhados, que de acordo com Roberts (2019) podem ser, por exemplo: a operação *Merge*, que concatena material linguístico de forma recursiva e computacional, gerando conjuntos de

expressões ordenadas e estruturadas; categorias funcionais, concebidas como categorias que integram o léxico (e.g. D, C e v) e que carregam informação acerca das propriedades gramaticalmente relevantes na língua e contribuem para a organização de uma estrutura sintática hierárquica na qual elementos de categorias lexicais (como nomes e verbos) são incorporados numa derivação linguística.

A linguagem não é um sistema desvinculado dos demais submódulos da cognição. Seu funcionamento e uso dependem da interação com outros processos cognitivos. De acordo com Hauser et al. (2002), a faculdade da linguagem possui duas facetas, uma que inclui um sistema combinatorial cuja principal função é gerar representações linguísticas recursivas e é específico da espécie humana (*Faculty of Language Narrow Sense* - FLN/ou gramática), e outra que engloba a primeira, composta por módulos encontrados em outras espécies (*Faculty of Language Broad Sense* - FLB). Estes módulos envolvem pelo menos dois sistemas internos ao organismo, denominados sistemas sensório-motor (S-M), que externaliza, via modalidade oral ou de sinais, expressões linguísticas formadas por processos combinatoriais recursivos, e conceptual-intencional (C-I), que conecta as mesmas expressões linguísticas ao pensamento formal, aos processos de referência, de raciocínio e de planejamento.

Para ter um melhor entendimento da relação dos sistemas S-M e C-I com a linguagem, pode-se pensar que o sistema S-M é responsável por lidar com som, ou seja, por captar informação fonológica, enquanto que o sistema C-I é responsável por lidar com significado através da interpretação semântica do que é captado. A gramática recebe a informação linguística a partir do sistema S-M e sua interpretação semântica envolve, primeiramente, a interpretação das partes que a compõem, segundo o Princípio da Composicionalidade (Pelletier, 2001), que define que o sentido de uma expressão complexa (uma sentença, por exemplo) é uma função dos sentidos das suas partes e da forma como estão combinadas, ou seja, o significado de uma frase é determinado pelos significados dos seus constituintes e pelo modo como estes estão combinados. Cabe à gramática, a partir do sistema C-I, construir a representação mental de um enunciado linguístico a partir da interpretação das unidades linguísticas presentes nele e sua organização. Assim, sistemas cognitivos atuam juntos na legibilidade das representações linguísticas, formando a Língua-I (Língua Interna/LI), composta por um grupo de submódulos (fonologia, morfologia, sintaxe, semântica

e pragmática) maturados através do processo de aquisição linguística, que realizam interações que conectam som/símbolo a significado de uma forma produtiva, tendo como *output* conjuntos de expressões ordenadas (Hauser et al., 2002).

Em resumo, a Língua-I compreende o conhecimento linguístico e é um sistema computacional de domínio específico responsável por produzir e processar informação linguística, sendo composto por um conjunto interno de submódulos conectados entre si através de interfaces gramaticais internas e, também, conectado a habilidades cognitivas externas (Hauser et al., 2002; Roberts, 2019). Ela representa a competência de um determinado falante, ou seja, é o conhecimento implícito que um falante nativo possui de sua língua adquirida através do processo de aquisição da linguagem, compreendendo um conjunto de regras interiorizadas pelo falante que permite a produção criativa e interpretação de infinitas sentenças em sua língua. Esse conhecimento linguístico internalizado (língua-I) é contraposto ao desempenho, caracterizado pelo uso real da língua em uma situação concreta, sujeito a imperfeições causadas por fatores externos ao conhecimento do sistema linguístico em si, como, por exemplo, questões de memória e atenção (Chomsky, 1965; Hauser et al., 2002; Chomsky, 2005). A competência, então, é o conhecimento da estrutura de uma determinada língua, enquanto que o desempenho é a manifestação desse conhecimento em uma determinada situação de uso.

A partir das ideias apresentadas, pode-se supor que, em casos em que a cognição humana é afetada, podem surgir anormalidades linguísticas associadas tanto a problemas de desempenho (devido a, por exemplo, prejuízos nos módulos cognitivos que interagem com o sistema computacional da linguagem) como a desestruturações no próprio sistema computacional da linguagem.

1.2.

Gramática e esquizofrenia

A esquizofrenia é um transtorno psicótico que está incluído no espectro da esquizofrenia e dos outros transtornos psicóticos e seus sintomas abrangem uma série de anormalidades neurocognitivas que levam a dificuldades amplas de habilidades cognitivas (APA, 2013). A literatura apresenta que o transtorno é

responsável por prejuízos amplos que compreendem baixa performance em testes de memória, quociente de inteligência (QI), função executiva e habilidades de atenção e outros (Goldberg et al., 1990; Censits et al., 1997; Torrey, 2002). Esses comprometimentos afetam a cognição, levando a complicações em atividades da “vida normal”, como reduções na competência social (Penn et al., 2008), capacidade de viver de forma independente e sucesso vocacional (Goldberg & Green, 1995).

Estudos acerca da comunicação de pacientes esquizofrênicos apontam que seu discurso é definido como desorganizado e difícil de acompanhar, exibindo associações confusas, tangencialidade, descarrilamento e incoerência (Andreasen, 1979a, b; Andreasen, 1986). A literatura na área apresenta alterações no discurso de esquizofrênicos que indicam anormalidades em diferentes níveis da linguagem: sintático, morfológico, semântico e pragmático (Covington, 2005; Kuperberg 2010; Walenski et al., 2010; Moro et al., 2015), tanto em produção como em compreensão. Esta caracterização aponta para a centralidade de marcadores linguísticos neste distúrbio, considerando não apenas questões de diagnóstico, mas também a caracterização e a análise sobre a cognição humana quando acometida por ele (Docherty et al., 1996; Hinzen & Rosselló, 2015; Hinzen, 2017).

A literatura na área demonstra diversos problemas com definitude referencial (Ditman & Kuperberg, 2011; Docherty et al. 2012; Hinzen, 2017), particularmente refletida no uso anômalo de pronomes. Dentre os relatos, as ocorrências de diferenças no discurso esquizofrênico mostram-se através de: uso acima do normal de pronomes pessoais (Watson et al., 2012; Fineberg et al., 2015), utilização de pronomes sem referência clara e sem antecedente claro (Harvey, 1983; Hinzen & Rosselló, 2015), maior quantidade de erros no uso de pronomes de terceira pessoa em comparação a pronomes de primeira e segunda pessoa (Sevilla et al., 2018; Chaves et al., 2020) e, também, maior erro de expressões definidas e pronomes (Çokal et al., 2018).

Outra característica do distúrbio que pode corroborar com déficits linguísticos é a memória de trabalho, que mostra-se prejudicada. Estudos mostram que a capacidade da memória de trabalho na esquizofrenia é reduzida, o que possui uma forte associação com tarefas cognitivas, especialmente aquelas envolvendo procedimentos linguísticos como, por exemplo, a atualização do plano de discurso (Docherty et al., 1996). Assim, falhas na comunicação podem ser

pensadas como déficits na memória de trabalho.

Nesta dissertação, investigamos a referência pronominal em textos escritos por esquizofrênicos, buscando realizar uma caracterização do uso pronominal no transtorno e suas anormalidades. Com isso, esperamos contribuir para o avanço no entendimento de processos gramaticais frente à esquizofrenia.

1.3.

Objetivos, perguntas norteadoras e hipóteses de trabalho

Esta dissertação tem os seguintes objetivos:

(a) Investigar a produção escrita de pacientes esquizofrênicos nativos do inglês, em mídias sociais, tendo em vista a complexidade estrutural, em particular a complexidade de pronomes. Isto é, descrever as peculiaridades de uso de pronomes em textos escritos em inglês por esquizofrênicos e analisá-las sob a ótica da Linguística Formal.

(b) Contribuir para a formação de um banco de dados que possa ser útil para estudos futuros sobre linguagem escrita em esquizofrenia.

(c) Realizar etiquetagem sintática manual dos textos analisados, baseado na teoria formal da Gramática, buscando contribuir para o desenvolvimento de um modelo de *tagging* refinado, útil no treinamento de algoritmos de aprendizado de máquina em Processamento de Linguagem Natural (NLP - *Natural Language Processing*), na área da Ciência da Computação.

A pesquisa parte das seguintes indagações:

(a) O transtorno da esquizofrenia afeta a gramática (língua-I), levando à redução da complexidade estrutural?

(b) Em que medida a forma (estrutura) e o significado (interpretação) dos pronomes são afetados?

(c) Essas alterações podem ser observadas na escrita?

Com base na literatura sobre esquizofrenia e linguagem e esquizofrenia e memória de trabalho (ver seção 2.5.), consideramos as seguintes hipóteses de trabalho:

- (a') A esquizofrenia leva a perdas de complexidade estrutural tanto no nível da sentença como no nível nominal.
- (b') Os pronomes são afetados tanto na forma como no significado.
- (c') Essas dificuldades podem ser observadas também na escrita, já que a escrita é uma tarefa cognitiva que induz aumento de memória de trabalho.

1.4.

Metodologia

Composição da amostra

Os dados foram captados por meio de um código de acesso a dados disponibilizados pela interface de programação de aplicações (API) *pushshift* (<https://pushshift.io/>), de textos escritos em inglês, anteriormente publicados (*postados*) na plataforma social *Reddit* (<http://reddit.com>). Participantes, grupos alvo e controle, foram captados a partir de autodeclaração de esquizofrenia (grupos alvo (presença) e controle (ausência)). Os grupos foram pareados por idade, sexo e diagnósticos adicionais (ansiedade e depressão). Filtragens automatizadas e manuais foram realizadas para obtenção de informações de *background* dos usuários e para descarte de participantes com transtornos mentais adicionais e/ou com postagens escassas ou com menos de 100 palavras. Foram selecionadas 78 postagens (uma para cada usuário) entre 100 palavras a 500 palavras, realizadas em *subreddits* fora do */r/schizophrenia*.

Análise morfossintática

Foi realizada etiquetagem (*tagging*) manual dos textos selecionados para compor a amostra, a partir dos seguintes parâmetros de análise: (a) número de palavras por postagem/texto; (b) tipo de sentença (matriz, subordinada, truncada); (c) tipo de truncamento em função de anomalia (+anômalo, - anômalo); (d) tipo de pronome em função da estrutura interna (pronomes fortes e fracos); (e) tipo de pronomes em função do traço de pessoa (+Locutor: pronomes de 1ª Pessoa e

2Pessoa, -Locutor: pronomes de 3Pessoa)²; (f) pronomes em função da referencialidade (+referencial, -referencial); (f) pronomes de 3Pessoa +referenciais em função de anomalias referenciais (+anômalo, -anômalo).

Análise estatística

O tratamento estatístico foi realizado com base nos parâmetros de análise estabelecidos (software SPSS (IBM)), com comparações não-paramétricas (Mann-Whitney *U* com valor de *P* fixo no nível .05) realizadas em duas etapas a fim de controlar a incidência de ansiedade e/ou depressão na amostra.

1.5.

Justificativa

A análise da produção escrita de pacientes esquizofrênicos mostra-se relevante para um maior entendimento do que pode ser considerado déficit geral de linguagem em esquizofrenia. Ressalta-se que análises dessas produções são ainda escassas, especialmente se comparadas a estudos baseados na produção oral.

A criação de um banco de dados acerca da produção textual em esquizofrenia se mostra de grande valia na viabilização de novos estudos que abordem a linguagem dentro do escopo do distúrbio.

A utilização de métodos da Ciência da Computação para captura de dados para análises formais é uma inovação que se mostra interessante e possivelmente produtiva. Dados coletados *in natura* por meio de técnicas de aprendizado de máquina podem contribuir para achados que não poderiam ser alcançados com a utilização de métodos experimentais ou de dados idealizados.

O desenvolvimento de um corpus qualitativo de etiquetagem de estruturas linguísticas mostra-se proveitoso para algoritmos de aprendizado de máquina em tarefas de NLP, visto que este tipo de dado como o que criamos é escasso na área.

Ainda, nosso trabalho promove a formação de equipe interdisciplinar no estudo de déficits linguísticos relativos a psicoses.

² O traço [+Locutor] se refere aos pronomes de primeira pessoa e segunda pessoa pois faz referência às entidades/indivíduos envolvidas na troca verbal. Pronomes de terceira pessoa são [-Locutor] pois não fazem referência a entidades diretamente envolvidas na troca verbal, mas sobre as quais o contexto verbal faz referência. Observe que não estamos considerando a informação de número, singular vs. plural. Esta informação não se mostra relevante no contexto em questão.

1.6.

Organização da dissertação

No capítulo 2, apresentamos a esquizofrenia, considerando os impactos do distúrbio na cognição, particularmente na linguagem. No capítulo 3, oferecemos primeiramente uma visão geral das propriedades estruturais dos pronomes pessoais, com foco na correferência pronominal, levando em conta questões sintáticas, semânticas e pragmáticas. Na segunda parte, apresentamos as dificuldades com a referência pronominal observadas em mentes diferentes, especialmente na esquizofrenia. No capítulo 4, o estudo experimental conduzido por nós é apresentado e os resultados obtidos discutidos à luz da literatura sobre esquizofrenia e linguagem e da literatura sobre pronomes. O capítulo 5 é dedicado à conclusão da dissertação

2

Esquizofrenia

2.1.

Caracterização geral da esquizofrenia

Emil Kraepelin, em 1887, usou a descrição de demência precoce (*dementia praecox*) para designar o conjunto de sintomas associados à psicose. Posteriormente, o psiquiatra Eugen Bleuler (1911) propôs o termo *esquizofrenia*. Desde então, tem-se buscado compreender a esquizofrenia como uma desordem mental, o que levou a redefinições do termo frente aos desenvolvimentos e descobertas científicas. Atualmente, o distúrbio é caracterizado como pertencente ao espectro da esquizofrenia, um construto multifatorial representativo de um conjunto de traços de personalidade que predispõe o indivíduo ao desenvolvimento da esquizofrenia (APA, 2013).

Em relação à ontogenia, sabe-se que a esquizofrenia decorre de predisposições genéticas (Goldberg et al., 1995; Sitskoorn et al., 2004), porém existem evidências de que sua etiologia envolve concomitantemente fatores ambientais, sociais e psicológicos. Frente a essas questões, o diagnóstico da esquizofrenia ocorre através da observação da manifestação de diferentes sintomas clínicos e, também, através da avaliação do nível de déficit funcional apresentado pelo paciente (APA, 2013).

O distúrbio apresenta uma sintomatologia heterogênea, o que indica que um mesmo paciente pode apresentar uma multiplicidade de sintomas ao longo de diferentes fases, assim como pacientes diferentes podem apresentar sintomas diferentes (Kuperberg, 2010). Os sintomas da esquizofrenia podem manifestar-se como exacerbações de sensações, crenças e comportamentos (e.g. alucinações, delírios, desordem do pensamento formal), como reduções ou ausência de características, muitas vezes sociais, comuns em indivíduos neurotípicos (e.g. afeto constrito, anedonia, alogia e ausência de iniciativa) e também como déficits em habilidades cognitivas (e.g. memória, atenção, fluência verbal), como já foi apontado (Habtewold et al., 2019). Portanto, a esquizofrenia apresenta-se como uma heterogeneidade de disfunções neuropsicológicas e reflete anormalidades em múltiplos aspectos do comportamento humano (Kuperberg, 2010).

Atualmente, os principais critérios utilizados para o diagnóstico da esquizofrenia são fornecidos pela Associação Americana de Psiquiatria (*American Psychiatric Association - APA*) - *Manual de Diagnóstico e Estatística de Distúrbios Mentais* (DSM), que se encontra em sua quinta versão.

Apesar de ser possível diferenciar a esquizofrenia de outros distúrbios neurológicos (e.g. epilepsia e traumas no sistema nervoso central) através de exames laboratoriais e indícios físicos, a tarefa de realizar sua distinção de outros distúrbios psicóticos (e.g. transtorno bipolar) se mostra mais árdua e, por isso, exige uma análise multifatorial de critérios. De acordo com DSM-V (APA, 2013), para o diagnóstico da esquizofrenia, dois de cinco sintomas-chave de distúrbios psicóticos devem estar presentes no paciente por uma porção significativa de tempo, durante um período de pelo menos um mês. Tais sintomas são: alucinações, delírios, discurso desorganizado, comportamento desorganizado ou catatônico e sintomas negativos, como apatia e anedonia. Destes cinco, pelo menos um dos três primeiros sintomas (alucinações, delírios, discurso desorganizado) deve estar presente.

Alucinações, segundo o DSM-V, são experiências perceptivas que ocorrem sem um estímulo externo, ou seja, sem base sensorial. Embora possam ocorrer alucinações relativas a diferentes modalidades sensoriais, as de base auditiva são mais comuns na esquizofrenia. Essas percepções geralmente tomam a forma de vozes, familiares ou não, e geralmente são palavras, frases ou sentenças, podendo ser consideradas como formas linguísticas estruturadas, com conteúdo fonológico e semântico, caracterizadas, portanto, como experiências linguísticas (Hinzen & Rosselló, 2015).

Delírios, de acordo com o DSM-V, são crenças sem apoio na realidade externa. São diferentes de alucinações, pois as alucinações perceptivas fogem ao controle de quem as tem, enquanto os delírios são declarações ou proposições sobre a realidade ou sobre si mesmo que o indivíduo decide manter, tomando-as como verdadeiras. Podemos decidir em que acreditamos, mas não temos controle mental sobre experiências perceptivas. As crenças, neste caso, são fixas e consideradas por quem as possui como legitimamente verdadeiras. Elas devem ser linguisticamente articuladas, visto que uma crença como “eu vim parar na Terra através de uma bolha mística” (Hinzen & Rosselló, 2015, p. 2) não pode ser formada sem articulação linguística. As crenças que caracterizam delírio podem se

dar de diversas maneiras e tratarem dos mais diversos assuntos, desde a certeza de que uma catástrofe irá destruir o planeta (delírio niilista) à certeza de que se está sendo perseguido (delírio persecutório). De qualquer forma, o que todos os tipos de delírios possuem em comum é serem considerados bizarros e não verdadeiros por pessoas da mesma cultura e da mesma língua, caracterizando-se, assim, como proposições falsas.

A desorganização do pensamento (*Formal Thought Disorder* – FTD) é inferida por meio do discurso do paciente (APA, 2013) e, por isso, é muitas vezes também chamada de desorganização do discurso. De acordo com o DSM-V, a FTD se manifesta na linguagem, por meio de processos de descarrilamento (mudança súbita de um tópico para outro), de tangencialidade (respostas e perguntas que possuem uma relação oblíqua ou nenhuma relação com o assunto tratado) e de incoerência (discurso tão desorganizado que se mostra completamente ou quase incompreensível). A desorganização do pensamento pode se dar em um grau menor, sendo algo comum e inespecífico (discurso levemente desorganizado) ou maior, quando o discurso se torna constantemente incompreensível sem o sintoma ter sido desencadeado por outros fatores (e.g. uso de drogas ou déficits linguísticos). Assim, para que o sintoma seja considerado FTD deve-se considerar se é suficientemente grave a ponto de prejudicar de forma substancial a comunicação efetiva (APA, 2013).

A desorganização comportamental e a catatonia são marcadas por comportamentos julgados como estranhos e/ou bizarros por indivíduos saudáveis. A catatonia, por exemplo, apresenta uma alternância entre períodos de passividade e de negativismo e períodos de súbita excitação. A esquizofrenia leva também a sintomas negativos, com reduções nas reações emocionais e expressões faciais, no contato visual, nos movimentos corporais, na expressão emocional e confere um jeito de falar que aparenta ser desanimado. Ainda, a prosódia, que passa informações sobre o estado emotivo do falante, também pode se apresentar alterada na esquizofrenia. Pode haver ainda avolia, redução de atividades motivadas, autoiniciadas. Essas características podem somar-se à falta de interesse de participação em contextos sociais e verbais.

Em resumo, a esquizofrenia é um distúrbio com diversas nuances que afeta amplamente o comportamento humano. Desde os primeiros estudos acerca do distúrbio empreendidos por Kraepelin e, posteriormente, por Bleuler, os sintomas

da esquizofrenia e sua caracterização sofreram diversas mudanças e redefinições com o intuito de alcançar um melhor entendimento do transtorno. De qualquer modo, nunca foi questionada a conclusão de que os sintomas da esquizofrenia apresentam impactos cognitivos que afetam diretamente a linguagem.

2.2.

Esquizofrenia: prejuízos cognitivos

A literatura em esquizofrenia aponta para diversas anormalidades neurocognitivas, que variam desde um menor desempenho em testes de atividades mentais até diferenças neuro-estruturais relacionadas a déficits cognitivos (Goldberg & Green, 1995; Hinzen, 2017). Estudos relatam prejuízos cognitivos variados: déficits de aprendizagem, atenção, abstração, linguagem, velocidade psicomotora, velocidade de processamento, memória de trabalho, função executiva e quociente de inteligência (QI) (Goldberg & Green, 1995; Censits et al., 1997; Mohamed et al., 1999; Heaton et al., 2001; Torrey, 2002; Dickinson et al., 2004; Fioravanti et al., 2005).

A exemplo, a série de pesquisas de Goldberg et al. (1990) com gêmeos monozigóticos (16 pares) discordantes para esquizofrenia vs. gêmeos normais (7 pares pareados com o grupo alvo em idade e nível educacional) teve o intuito de controlar a variância genética em uma bateria de testes de diferentes medidas cognitivas (e.g. QI, aritmética, vocabulário e diversos outros). Os resultados indicaram performance mais baixa do grupo alvo em comparação ao grupo controle na maioria dos testes, apresentando diferenças severas em testes de vigilância, memória e formação de conceitos. Essas observações foram interpretadas como uma afetação maior na área do córtex frontotemporal. Goldberg et al. (1995), em um experimento semelhante, analisaram a performance de gêmeos não-esquizofrênicos (7 pares controle), gêmeos discordantes para a esquizofrenia (20 pares) e gêmeos concordantes para a esquizofrenia (8 pares), observando que, entre gêmeos discordantes, os indivíduos afetados pela esquizofrenia têm desempenho pior em diversos testes de memória, QI, decisão e tempo de resposta, aproximando-se, portanto, dos resultados obtidos para o grupo concordante de gêmeos. Em contraste, os indivíduos (dos gêmeos discordantes) não afetados não apresentam redução em nenhuma medida em comparação ao

grupo controle. Os autores concluem que os prejuízos cognitivos em esquizofrenia são estáveis apesar de diferenças contextuais (e.g. ambiente familiar) e se mostram salientes mesmo quando há controle da variância genética.

Os prejuízos cognitivos relativos à esquizofrenia se mostram até mesmo no período prodromal, que antecede a manifestação do distúrbio (Peuskens et al., 2005; Niendam et al., 2006; Monteiro & Louzã, 2007), e são marcantes em estágios iniciais da psicose (Goldberg et al. 1995; Censits et al. 1997; Gold et al., 1999; Mohamed et al., 1999; Bilder et al., 2000).

Torrey (2002) apresenta uma metanálise de 65 estudos sobre pacientes nunca tratados com medicação antipsicótica, com achados neuropsicológicos indicando anormalidades inerentes aos efeitos da esquizofrenia, e não relacionadas a medicamentos. Foram encontrados déficits de memória e aprendizagem, atenção, abstração, linguagem, funcionamento executivo e velocidade psicomotora. Em outra investigação, foram realizados testes com um grupo de alto risco para primeiro episódio de psicose (identificado através de critérios diagnósticos), mas que ainda não havia manifestado esquizofrenia. Os resultados indicam baixa performance em testes neuropsicológicos (velocidade de processamento, aprendizagem verbal, memória e velocidade motora) quando comparados aos resultados de controles (Niendam et al., 2006), reforçando, portanto, a presença de prejuízos cognitivos inerentes à esquizofrenia.

Censits et al. (1997) compararam resultados em testes neurocognitivos de pacientes esquizofrênicos em primeiro episódio de psicose e pacientes com esquizofrenia crônica. Os resultados obtidos foram semelhantes entre os dois grupos, ambos apresentaram baixa performance em testes de memória, função executiva e habilidades de atenção.

Em resumo, os comprometimentos neurocognitivos observados na desordem em discussão afetam a cognição, levando a complicações em atividades da “vida normal”, como reduções na competência social (Penn et al., 2008), capacidade de viver de forma independente e sucesso vocacional (Goldberg & Green, 1995).

2.3.

Esquizofrenia e linguagem

De todos os sintomas amplos da esquizofrenia, a FTD (desorganização do pensamento formal) é o que possui maior conexão direta com a linguagem (APA, 2013).

Andreasen (1979a, b) realizou uma série de estudos com 113 pacientes e, a partir disso, estabeleceu classificações para os tipos de FTD através de inferências discursivas. Isso levou, posteriormente, à criação de uma lista classificatória composta por elementos referentes a padrões percebidos no discurso de pacientes com FTD (Andreasen, 1986): pobreza discursiva (discurso caracterizado por poucas palavras); pobreza de conteúdo do discurso (vagueza discursiva, alogia, repetições); pressão de discurso (aumento excessivo na quantidade de discurso espontâneo, em comparação ao que é considerado normal ou comum socialmente); distração discursiva (dificuldade em manter o foco discursivo, perda de atenção); tangencialidade (respostas parciais ou irrelevantes a perguntas); descarrilamento (discurso solto, falta de associação entre ideias); incoerência (salada de palavras, discurso difícil de compreender); ilogicidade (discurso com falta de relações lógicas); associação sonora de palavras (escolha de palavras governada por sons, levando a dificuldade de entendimento); neologismos; aproximação de palavras (palavras já existentes usadas de formas novas e pouco convencionais, ou criação de novas palavras); circunstancialidade (discurso indireto e/ou pouco objetivo); perda de objetivo (dificuldade em levar uma sequência de pensamentos a uma conclusão lógica); perseverança (repetição persistente de palavras, ideias, etc.); ecolalia (padrão discursivo caracterizado pela repetição de palavras ou sentenças ouvidas); bloqueio (parada súbita no decorrer do discurso sem que uma ideia tenha sido completada); discurso pomposo ou excessivamente formal; autorreferência (foco excessivo em si mesmo, que pode ocorrer em situações pouco convencionais); e parafasia (substituição de uma palavra por outra ao tentar comunicar algo específico ou pronúncia errada de palavras).

Enfatiza-se que, apesar de nem todos os pacientes esquizofrênicos demonstrarem anomalias linguísticas patentes, investigações linguísticas no contexto do distúrbio são cruciais para o diagnóstico (Kuperberg, 2010). A

centralidade da linguagem justifica-se pelo fato de poder ser abordada diretamente. Por isso, diversos estudos focam na identificação dos componentes da gramática prejudicados na esquizofrenia (Covington, 2005; Ditman & Kuperberg, 2010; Kuperberg, 2010). Desde então, investigações acerca de alterações no comportamento verbal de esquizofrênicos indicam anormalidades em diferentes níveis da linguagem: sintático, morfológico, semântico e pragmático (Covington, 2005; Kuperberg 2010; Walenski et al., 2010, Moro et al., 2015), com investigações sobre produção e compreensão. Características marcantes da comunicação na esquizofrenia, observadas em estudos anteriores³, apontam que o discurso é governado por maior ênfase em traços fonológicos e semânticos (Chaika, 1990), apresenta maior dificuldade de entendimento de estruturas gramaticais (Moro et al., 2015), anomalias referenciais (Ditman & Kuperberg, 2010, Hinzen, 2017; Sevilla et al., 2018), redução na complexidade sintática (Çokal et al., 2018; Chaves, 2022), dificuldade de compreensão de expressões figurativas (e.g. ironia, provérbios, metáforas e expressões idiomáticas) (Mitchell & Crow, 2005; Rossetti et al., 2018), e produção de narrativas com falta de sequências lógicas de proposições e/ou ausência de informação (Docherty & Gottesman, 2000). Estas anomalias caracterizam a comunicação em esquizofrenia como estranha, indicando que a sintomatologia da esquizofrenia não envolve apenas questões cognitivas e linguísticas mais amplas, mas também questões gramaticais específicas, de ordem estrutural, aproximando a relação entre a forma e conteúdo na linguagem humana. Nessa linha de entendimento, a gramática não pode ser desvinculada da cognição humana, pois atua na construção de significado e na estruturação da forma do pensamento (Hinzen & Rosselló, 2015; Hinzen, 2017).

³ É importante ressaltar que, conforme informam Çokal et al. (2018), nem todas as investigações linguísticas acerca da esquizofrenia empreendem ou reportam medidas de QI, o que os autores apontam como medidas importantes para um entendimento mais aprofundado da esquizofrenia e seus sintomas (e.g. FTD). Moro et al. (2015) e Çokal et al. (2018), por exemplo, encontraram correlações significativas entre medidas de QI e performances em tarefas linguísticas. No estudo de Moro et al. (2015), medidas de QI total e taxas de acerto em tarefas de conhecimento sintático foram correlacionadas, enquanto que, no experimento de Çokal et al. (2018), foram encontradas correlações entre medidas de QI e complexidade sintática e anomalia referencial. Feitas essas observações, informamos que nem todos os estudos aprofundados por nós na seção 2.4. reportam medidas de QI, mas quando o fazem mencionamos em nossas descrições dos respectivos estudos. Além disso, percebe-se na literatura uma tendência ao pareamento de nível educacional como medida de controle. Pareamentos de nível educacional, quando reportados, também foram mencionados.

2.4.

Alterações linguísticas

2.4.1.

Distúrbios fonológicos, morfológicos e sintáticos

Estudos da produção linguística na esquizofrenia mostraram que problemas de forma, além da falta de marcadores estruturais e ordem sintática definida, diferenciam a linguagem esquizofrênica: em diversos estudos, em comparação a controles, grupos de pessoas afetadas por esquizofrenia apresentaram anomalias fonológicas (Clemmer, 1980; Cohen et al., 2014), maior produção de sentenças mais curtas e simples no nível sintático (DeLisi, 2001, Chaves, 2022), menor conhecimento estrutural (Moro et al., 2015) e mais erros morfosintáticos (Walenski et al., 2010). É evidente também dificuldades na estruturação de sintagmas nominais, especialmente pronomes (ver capítulo 3).

Clemmer (1980) investiga o discurso esquizofrênico através de tarefas de ler e reproduzir histórias. Os participantes liam em voz alta histórias de um parágrafo e depois as reproduziam com suas próprias palavras. O desempenho de 20 pacientes esquizofrênicos, com pelo menos o ensino médio concluído, foi comparado ao desempenho do grupo controle (20 participantes pareados ao grupo alvo em idade, raça, sexo e nível educacional). As tarefas envolveram duas histórias distintas, a primeira sem nenhum tipo de estranhamento linguístico e a segunda com peculiaridades gramaticais, como troca ou inversão inesperadas de palavras. As medidas analisadas foram: pausas silenciosas (duração, frequência e localidade), dicção (duração e comprimento de sílabas, grau de articulações) e hesitações (frequência e duração). A versão usual das histórias melhor discriminou os grupos, visto que o grupo controle apresentou resultados semelhantes aos do grupo esquizofrenia nas tarefas envolvendo as versões peculiares, devido às dificuldades consequentes das inversões e trocas inesperadas de palavras. Os resultados da tarefa de leitura indicam maior tempo de pausa silenciosa (nos dois tipos de história) e duração de sílabas, dicção mais lenta, fluência discursiva mais desconectada e maior número de hesitações em esquizofrênicos. Na tarefa de recontar as histórias, os esquizofrênicos apresentaram mais hesitações, diferenciando-se dos controles também em duração

silábica, frequência e localidade das pausas. A seguir, é apresentado um exemplo, encontrado em Clemmer (1980, p. 180), de um trecho, segundo o autor, discriminado como associado à esquizofrenia:

(1) *and the child (530 msec) ran (1470 msec) out in front (2100 msec) of the truck (1600 msec) and" ah (1400 msec).*

O autor aponta que o fragmento acima é associado à esquizofrenia pois possui pausas frequentes e longas, apontadas no trecho pela contagem em milissegundos (em parênteses).

Cohen et al. (2014) realizaram uma metanálise de 13 estudos, com um total de 480 pacientes esquizofrênicos e 326 controles. Os resultados obtidos revelam a produção de alogia e efeito de embotamento afetivo na esquizofrenia. Medidas linguísticas relacionadas a pausas também se mostraram relevantes.

Prejuízos morfossintáticos são relatados em Walenski et al. (2010). Em tarefa experimental envolvendo preenchimento de lacunas em sentenças do inglês com formas verbais no passado do indicativo, falantes esquizofrênicos (43), comparados a controles saudáveis (42 pareados em idade), apresentaram maior dificuldade com formas verbais regulares, em que o passado é formado pela sufixação de *-ed* à raiz verbal (e.g. *dance* → *danced*). Essa dificuldade mostrou-se atenuada em formas verbais irregulares (e.g. *run* → *ran*), as quais podem não envolver um processo derivacional, sendo antes formas lexicalizadas. Durante a investigação, os autores realizaram diversos controles sobre fatores específicos de a indivíduos: sexo dos indivíduos, anos de educação, QI e destreza manual. Os achados foram interpretados pelos pesquisadores como reflexo de redução de memória procedural em esquizofrenia. Foi encontrada também correlação positiva entre traços de FTD e erros na produção do passado regular.

DeLisi et al. (1997), em estudo sobre anomalias no nível neuro-estrutural, realizaram tarefas linguísticas envolvendo linguagem receptiva e expressiva (e.g. identificar palavras após ouvi-las apontando para imagens e fala livre por 6 minutos), memória (e.g. recontar uma história após 30 minutos de escuta). Foram realizadas testagens de inteligência verbal e atenção verbal procurando investigar se resultados nessas tarefas estão associados a resultados em tarefas linguísticas e

testagens de anatomia cerebral. Os resultados obtidos pelos pesquisadores também apontam para maior número de erros de morfologia em esquizofrênicos (e.g. problemas com formação do plural e com marcação de tempo passado), com resultados inconclusivos sobre as associações testadas. DeLisi (2001) analisou 29 pacientes diagnosticados como crônicos para esquizofrenia, 9 pacientes que apresentaram primeiro episódio esquizofreniforme, 12 controles. 11 famílias com familiares esquizofrênicos também participaram do experimento (20 participantes esquizofrênicos ou transtorno esquizoafetivo e 13 participantes não-psicóticos). Para a análise, solilóquios orais foram gravados a partir de uma tarefa que consistia em imaginar e descrever o que ocorre em uma série de 7 imagens. A partir disso, a sintaxe de cada participante foi acessada de acordo com 5 características: (1) número de sentenças com ordem errada de palavras; (2) número de elementos nulos; (3) número de sentenças com argumentos verbais ausentes; (4) número de sentenças subordinadas e coordenadas; e (5) número de erros gramaticais. A partir de análises estatísticas, DeLisi observou que pacientes com esquizofrenia, em comparação a controles, apresentaram redução na complexidade sintática (sentenças subordinadas e coordenadas), menor número de palavras por solilóquio, mais conteúdo inapropriado e menor coesão entre sentenças.

Quanto ao componente sintático, Özcan et al. (2017), comparando 50 pacientes esquizofrênicos com 50 controles saudáveis (idade, sexo e nível de educação pareados) em 4 tarefas (narração de imagens, contação de histórias baseadas em sequências de imagens, produção discursiva semi-estruturada e discurso livre), observaram redução da complexidade sintática na esquizofrenia: maior número de sentenças simples e sentenças mais curtas. Essas observações estão de acordo com os resultados de outros estudos relatando déficits tanto na compreensão quanto na produção de estruturas sintáticas complexas (e.g. subordinação de sentenças). Essa redução sintática pode estar relacionada à redução de memória de trabalho (Condray et al., 1996; Condray et al., 2002).

Da mesma forma, Condray et al. (2002) relatam que pacientes esquizofrênicos demonstram dificuldades em tarefas envolvendo construções sintaticamente complexas como, por exemplo, ordenar complementos verbais, objetos diretos e indiretos, dentro de estruturas argumentais e entender as relações lógicas entre termos de sentenças relativas. Os pesquisadores realizaram controles

para distúrbios neurológicos, QI verbal, habilidade de leitura e abuso de substâncias. Esquizofrênicos (32, sexo masculino), em comparação a controles (22, sexo masculino), apresentam dificuldades na compreensão de sentenças relativas, com erros na atribuição de papéis sintáticos.

Moro et al. (2015) avaliaram o conhecimento sintático de falantes italianos com esquizofrenia (58) em comparação a controles (30). Foram apresentadas visualmente, a cada participante, um total de 150 sentenças em italiano com distribuição igual entre sentenças bem formadas e sentenças com violações gramaticais (90 sentenças) e semânticas (60 sentenças). Os erros sintáticos envolviam violação na derivação de interrogativas –qu (violação de movimento cíclico e efeitos de superioridade), violação da localidade de clíticos em sentenças afirmativas (posição do referente em violação ao princípio B (Chomsky, 1981)), erros na atribuição de foco contrastivo com inversão de ordem. A análise do grau de acerto em detecções de violações sintáticas sugere performance diminuída em esquizofrênicos, o que não foi observado na detecção de anomalias semânticas. De acordo com a investigação, esquizofrênicos possuem maiores dificuldades em identificar violações sintáticas como exemplificado na sentença a seguir, retirada do conjunto de estímulos experimentais de Moro et al. Note que há aqui uma violação de ilha, já que o pronome interrogativo é extraído de dentro de uma oração adjunta:

(2) Chi gli scrivi prima di incontrare?

“Who do you write to him before meeting?”

(Moro et al., 2015, p. 151)

Além disso, medidas de QI foram reportadas como correlacionadas a taxas de acerto de anomalias sintáticas e negativamente correlacionadas a taxas de acerto de sentenças curtas corretas, levando os autores a sugerirem que habilidades cognitivas globais podem estar associadas aos resultados dos testes.

Chaves (2022), analisando narrativas orais de experiências oníricas (sonhos) e de vigília, produzidas por pacientes esquizofrênicos falantes de português brasileiro, relata maior volume de sentenças matrizes em esquizofrênicos, comparados a controles. Relata também menor volume de sentenças truncadas não

anômalas. Esses resultados indicam, portanto, preferência, na esquizofrenia, por sentenças com menor complexidade sintática (matrizes em oposição a subordinadas) e dificuldade de realizar truncamentos gramaticais (elisão) no domínio da sentença. A seguir apresentamos uma transcrição analisada pela autora, apresentada em seu capítulo de descrição experimental e modificada aqui para conter apenas as marcações de estruturas sintáticas:

(3) [MS fica aquela sensação estranha, um sonho estranho, sabe? [MS um outro amigo estava cismado um pouco.] [STRUNC *pro* estava] Isso. Não, não [MSTRUNC *pro* foi] [MS *pro* pararam [ES quando eu acordo, sabe?] [MS aí elas param] [MSTRUNC imagem, doutora] [MS *pro* tem imagem] [MS Vixe, *pro* já sonhei demais com isso, em tantos lugares.] [STRUNC uns lugares estranhos] [MSTRUNC cismado] [MSTRUNC *pro* era] Não. Não, não. [MSTRUNC *pro* era estrada] [MS *pro* parecida como se *pro* fosse um bairro [ESRC que eu já morei] sabe?]

Nas marcações da autora, por exemplo, é possível perceber um número maior de sentenças matrizes (MS) do que sentenças subordinadas (ES). Além disso, os resultados de Chaves indicam ainda maior volume de pronomes nulos em esquizofrenia. Voltaremos a este achado na segunda parte do capítulo 3, quando discutiremos pronomes em esquizofrenia.

Çokal et al. (2018) relatam diferenças significativas na complexidade sintática de pacientes esquizofrênicos. Foi analisado o uso de expressões nominais e a complexidade sintática de 30 participantes com esquizofrenia (15 com +FTD e 15 -FTD), 15 parentes de primeiro grau de esquizofrênicos e 15 neurotípicos (controle) em uma tarefa de descrição de um recorte de quadrinho. Os autores reportam medidas de QI, indicando que pacientes esquizofrênicos (com ou sem FTD) apresentaram medidas de QI semelhantes, mas menores do que as apresentadas por grupos não-clínicos, o que os autores apontam estar de acordo com investigações anteriores. As amostras de discurso foram coletadas enquanto os participantes contavam uma história retratada por 8 quadrinhos, que foram transcritas para análise de variáveis linguísticas (anomalia de anáfora de terceira pessoa, anomalia referencial geral, anomalia referencial em expressão nominal definida, tipos de expressão nominal, complexidade sintática e erros sintáticos).

O grupo esquizofrenia +FTD produziu mais anomalias referenciais do que controles e também produziu a menor quantidade de expressões nominais definidas (e.g. pronomes, expressões nominais dêiticas), enquanto parentes de primeiro grau de esquizofrênicos produziram a maior quantidade de expressões nominais definidas, o que também os distingue dos controles. Em esquizofrenia +FTD, observou-se ainda maior quantidade de anomalias referenciais em expressões nominais definidas (e.g. referências vagas, uso indevido da terceira pessoa e etc.) e a menor taxa de complexidade sintática (e.g. menor número de sentenças subordinadas). As taxas de produção de expressões nominais definidas distinguiram os dois subgrupos de esquizofrenia (+FTD, -FTD), com -FTD produzindo mais expressões nominais definidas. Os autores apresentam, como exemplo, o seguinte trecho para apresentar anomalia referencial, em que se torna difícil distinguir as relações ambíguas construídas através de *he/him*:

(4) ‘There is a man phoning a man and he is making an appointment to come and have dinner with him’

(Çokal et al., 2018, p. 6)

Os resultados foram interpretados pelos autores como evidência de que os sintomas da esquizofrenia podem ser identificados em padrões gramaticais, com redução de estrutura sintática, e também que medidas neurocognitivas (e.g. QI) podem estar associadas a medidas linguísticas. Voltaremos às dificuldades gramaticais em esquizofrenia, com foco em expressões nominais (pronomes), na segunda parte do capítulo 3.

2.4.2.

Distúrbios semânticos e pragmáticos

A literatura aponta que esquizofrênicos possuem dificuldade de compreensão de expressões figurativas como ironia, provérbios, metáforas e expressões idiomáticas (Mitchell & Crow, 2005; Rossetti et al., 2018) e produção de narrativas com falta de sequências lógicas de proposições e ausência de informação (Docherty & Gottesman, 2000). Rossetti et al. (2018), por exemplo, realizaram uma metanálise focada na compreensão de metáforas. Os resultados

apontam uma preferência na esquizofrenia pelo sentido literal (denotativo). Foram detectados também problemas de memória semântica e de memória de trabalho (ver seção 2.5).

Champagne-Lavau & Stip (2010) testaram 20 pacientes esquizofrênicos e 20 controles (pareados em idade, anos de educação e QI), todos destros e falantes nativos de francês, para três habilidades: pragmática, Teoria da Mente e funções executivas. Os testes de pragmática envolveram entendimento de metáforas (solicitação para explicar metáforas idiomáticas vs. não-idiomáticas) e pedidos indiretos (explicar enunciados após serem apresentados contextos que influenciam a interpretação literal vs. não-literal). As observações obtidas indicam dificuldades na interpretação de metáforas (e.g. *mon ami a le coeur gros* ('my friend has a heavy heart')) e de pedidos indiretos em esquizofrênicos, correlacionando-se com a falta de flexibilidade discursiva e menor performance em Teoria da Mente.

Docherty & Gottesman (2000) realizaram entrevistas com gêmeos monozigóticos e dizigóticos (diferenças individuais foram controladas para níveis de psicopatologia geral e não houveram grandes disparidades de sexo, idade e situação socioeconômica), discordantes para esquizofrenia. Os participantes esquizofrênicos apresentaram maiores níveis de problemas linguísticos em comparação aos pares não afetados. A medida discriminatória foi o número de referências ausentes, ou seja, quando um DP é utilizado sem que um possível antecedente seja recuperável no discurso, como apresentado pelos autores:

(5) I like to work all right. Some of **those shops** were filthy. I liked **the bakeries**, some of **the shops** are clean. (No prior mention of any shops or bakeries).

(Docherty & Gottesman, 2000, p. 3)

Parola et al., (2020), investigando também questões pragmáticas, utilizaram metodologia de aprendizado de máquina com aplicação de rede Bayesiana para caracterizar a relação entre cognição, Teoria da Mente e habilidades pragmáticas em esquizofrenia em comparação a controles saudáveis. Um dos objetivos foi identificar as habilidades cognitivas e pragmáticas mais informativas para a discriminação da esquizofrenia. Para isso, aplicaram avaliações de diferentes

performances pragmáticas (linguística, extralinguística, paralinguística, contextual e conversacional, Teoria da Mente, funções cognitivas e outras) e inteligência geral. A população esquizofrênica e controle foram pareadas em sexo, idade, nível de educação e passaram por testes neuropsicológicos (*Mini-Mental State Examination* (MMSE), *Token Test* e *Denomination scale of the Aechener Aphasia test* (AAT)) visando controlar disparidades de resultados. Assim, o classificador de rede Bayesiana foi utilizado para discriminar entre esquizofrênicos e controles, concluindo que, de todas as avaliações, apenas habilidades linguísticas relativas à pragmática se mostram como diferenciadoras diretas entre esquizofrênicos e controles. Tais habilidades pragmáticas envolveram testes de expressão gestual, ironias orais e gestuais, atos de comunicação direta e indireta, entendimento de sentido (literal vs. não-literal), e participação adequada em conversações.

O discurso de pacientes esquizofrênicos também é apontado como governado por associações léxico-semânticas (Chaika, 1990; Titone et al., 2000; Ditman & Kuperberg, 2010). Investigações semânticas observaram que, na esquizofrenia, a interpretação semântica é influenciada de forma inapropriada por estímulos anteriores a partir de falhas de inibição do efeito de *priming* semântico, causando anormalidades na produção e compreensão de significado (Gernsbacher et al., 1999; Kuperberg et al., 2009; Ditman & Kuperberg, 2010). Ao realizar uma metanálise de estudos acerca das associações de palavras observadas no discurso esquizofrênico, Ditman & Kuperberg (2010) apontam para incoerências baseadas na preocupação com o significado semântico⁴. Considere o exemplo a seguir, onde o conteúdo semântico de uma palavra determina o aparecimento de outras palavras (e.g. *wise* → *nowise*; *whys* → *wherefores*; *fours* → *fives* → *numbers*), não necessariamente conectadas estruturalmente.

(6) “If you think you are being wise to send me a bill for money I have already paid, I am in nowise going to do so unless I get the whys and wherefores

⁴ A preferência pelo significado semântico em detrimento do significado pragmático parece ser uma característica da linguagem na esquizotipia de modo geral. Chaves & Rodrigues (2020), no prelo) mostram, em estudos sobre interpretação de expressões nominais definidas em português brasileiro, que falantes com maior soma de traços esquizotípicos negativos, em particular afeto constrito e percepção alterada, têm preferência por atribuir leituras de indivíduo a expressões nominais definidas, suprimindo leituras de classe. Essa preferência é marcante quando o contexto discursivo reforça a leitura de indivíduo.

from you to me. But where the fours have been, then fives will be, and other numbers and calculations and accounts to your no-account...”

(Ditman & Kuperberg, 2010, p. 3)

Ditman & Kuperberg apontam que essas associações semânticas se tornam mais salientes em *input* com textos maiores, sugerindo, portanto, que este tipo de associação aumenta com o incremento linguístico.

No nível das expressões nominais é patente uma preferência por expressões indefinidas e a ocorrência de expressões definidas anômalas. Anomalias referenciais e falta de definitude são consideradas características clássicas em falhas de comunicação em esquizofrenia (Hinzen & Rosselló 2015, Hinzen 2017; Sevilla et al., 2018). Sevilla et al. (2018) investigam os efeitos de FTD na linguagem. Foram coletadas narrativas (orais) de 40 pacientes esquizofrênicos (20 +FTD, 20 -FTD), e de 14 controles saudáveis pareados em QI pré-mórbido. As medidas linguísticas avaliadas foram: captura de anomalias referenciais, expressões nominais definidas e indefinidas, pronomes, parafasia, violações de restrições semânticas (em que o significado lexical das palavras aponta que o evento não pode ocorrer), número de sentenças subordinadas, número de relações gramaticais (contadas através da identificação de itens gramaticais) e erros gramaticais formais (violações sintáticas). 4 dessas variáveis apresentaram diferenças entre os grupos de participantes: expressões nominais definidas, expressões nominais de terceira pessoa, pronomes e parafasias. Todas elas apontam para o uso anômalo de expressões nominais referenciais (e.g. uso de *a girl* em contexto em que *the girl* é o sintagma apropriado), em que expressões nominais definidas e pronomes foram mais afetados do que expressões nominais indefinidas. Essas observações estão de acordo com as ideias de Hinzen (2017) e Hinzen & Rosselló (2015), que argumentam que, se a linguagem é o configurador do pensamento humano, fenótipos cognitivos e linguísticos devem se relacionar. Zimmerer et al. (2017) também discutem a ancoragem dêitica na esquizofrenia, argumentando que há ruptura dessa ancoragem, que envolve estruturas de expressões nominais plenas e pronomes.

2.5.

Esquizofrenia e memória de trabalho

Observações correlacionando falhas em tarefas cognitivas, incluindo tarefas linguísticas, com falhas em memória de trabalho são abundantes na literatura sobre esquizofrenia, além de associações com déficits de controle executivo e de Teoria da Mente (Perry et al. 2001; Bowie & Harvey, 2006; Park & Gooding, 2014).⁵

A memória de trabalho é um componente de memória temporária associado a processos relativos à aquisição, armazenagem, recuperação e manipulação de informação (Perry et al., 2001; Lezak et al., 2004; Kuperberg, 2010).

Condray et al. (1996) relatam correlação entre baixa complexidade gramatical, dificuldade no processamento de sentenças complexas vs. sentenças simples, e redução de memória de trabalho verbal. Condray et al. (2002) testam se a sintaxe receptiva é prejudicada em esquizofrenia e, para isso, realizaram um teste de acerto de compreensão entre 32 pacientes (sexo masculino) diagnosticados com esquizofrenia e 22 controles (sexo masculino e sem diagnóstico de transtorno psiquiátrico) que consiste em responder questões do tipo *qu-*, em língua inglesa (“Who did X? Who was done X?”), sobre informações em sentenças do tipo matriz e relativas apresentadas oralmente. Os resultados foram testados em função de possíveis relações com resultados de funções cognitivas através de análise de regressão linear com entradas simultâneas de variáveis (uso do programa SPSS) a partir de resultados de testes de associação entre linguagem e compreensão e habilidade intelectual geral (WAIS-R), com uso de testes de correlação (Pearson *r*). Resultados demonstraram que o grupo de pacientes esquizofrênicos apresentou performance caracterizada por menor taxa de acertos (média 14% menor) na tarefa de compreensão sintática do que controles, levando a crer que a sintaxe receptiva é prejudicada em esquizofrenia e, conseqüentemente, esquizofrênicos possuem redução de compreensão de sentenças gramaticalmente complexas. Igualmente importante, os autores observaram que os resultados do teste de compreensão estão correlacionados com o funcionamento em domínios da memória, observação que os autores interpretam

⁵ Para associação entre FTD e memória de trabalho, ver Nestor et al. (1998).

como associada ao entendimento de que a sintaxe receptiva envolve múltiplos componentes funcionais e, portanto, um aumento de complexidade sintática pode acarretar em aumento de carga de funções cognitivas que envolvem a memória de trabalho, levando a performances prejudicadas de compreensão em esquizofrenia.

Conclusão semelhante é apresentada em Bagner et al. (2003), que também encontraram forte correlação entre baixa complexidade sintática e falhas de memória de trabalho na esquizofrenia. Na mesma linha, Docherty et al. (1996) e Docherty (2012) associam falhas na construção da referência linguística a problemas de memória de trabalho. Forbes et al. (2009) realizaram uma meta-análise de 187 estudos sobre esquizofrenia e cognição, apontando para forte correlação entre baixa pontuação em tarefas de memória de trabalho e baixa performance em tarefas fonológicas e viso-espaciais. Ainda, Becker et al. (2012), em estudo sobre performance em tarefas cognitivas na esquizofrenia, encontram associações entre alogia e baixa memória de trabalho verbal. Em contrapartida, Moro et al. (2015) (seção 2.4.1. acima) buscaram por correlações entre resultados obtidos em testes linguísticos e pontuações em teste de memória de trabalho (*n-back task*), não encontrando nenhum resultado significativo.

A memória de trabalho parece estar também implicada em atividades de escrita. Pesquisa realizada com crianças italianas (de Vita et al., 2021) indica que a memória de trabalho é um recurso relevante no desenvolvimento e domínio da escrita: crianças com melhores resultados em testes de memória de trabalho apresentam melhor performance em tarefas relativas ao registro escrito. Igualmente, as observações de Kellog et al. (2007) também associam memória de trabalho verbal com habilidade de escrita em alunos universitários.

Em consonância com as conclusões apresentadas acima, Thomas et al. (1993) apresentam resultados de pesquisa sobre textos escritos (Teste de Hunt) por esquizofrênicos, comparados a maníacos e controles saudáveis (i.e. sem transtorno mental). Não houve diferença entre os grupos em relação à complexidade sintática. No entanto, esquizofrênicos e maníacos apresentaram maior número de erros sintáticos e maior dificuldade na integração de informação. Os autores sugerem que as dificuldades linguísticas encontradas refletem déficits de memória de trabalho.

No capítulo 3 (seção 3.4.), apresentaremos resultados de pesquisas sobre língua escrita em esquizofrenia, indicando que problemas com a referência pronominal são observados também no registro escrito.

Essas observações são enfaticamente relevantes para nossa pesquisa, já que a nossa amostra é composta de textos escritos. Como detalharemos no capítulo 4 (ver também capítulo 1), partimos de hipóteses de que a esquizofrenia leva à redução da complexidade sintática, observada no domínio sentencial (preferência por sentenças simples e maior volume de truncamento sentencial) e, conseqüentemente, a dificuldades na construção da referência pronominal. Essas hipóteses se baseiam nas observações descritas acima, sobre baixa complexidade gramatical na esquizofrenia, esquizofrenia e redução de memória de trabalho, e sobre o papel da memória de trabalho na linguagem escrita.

3

Pronomes em gramática típica e atípica

3.1.

Expressões nominais, referência e pronomes

Expressões nominais (sintagmas determinantes em português ou *determiner phrases* em inglês - DPs) denotam entidades ou conjunto de entidades, animadas ou inanimadas, definidas ou não, específicas ou não, podendo, portanto, ser referenciais (quando denotam uma entidade definida), ou não-referenciais (quando denotam conjuntos de entidades ou entidades indefinidas, ou quando não denotam, como ocorre com pronomes expletivos) (Heim, 1982; Heim & Kratzer, 1998).

Na estrutura sintática, DPs ocupam posições argumentais (e.g. argumento externo e interno de um predicado verbal) e posições não argumentais, como a posição de tópico vinculada a uma posição argumental. Expressões nominais não denotativas (expletivos) ocupam a posição de sujeito sentencial, não funcionando semanticamente e sintaticamente como argumento de um predicado.

Importante para o nosso trabalho, DPs podem ser identificados pela possibilidade de substituição por um pronome, como em (7). Portanto, pronomes também são considerados DPs, pois apresentam a mesma distribuição sintática e as mesmas funções semânticas que DPs plenos.

- (7) a. **Esse empresário** comanda as **duas empresas**
 b. **Ele as** comanda
 c. **Ele** comanda **elas**

A definição formal de expressões nominais como DPs leva em consideração que toda expressão nominal possui, além do nome (N), uma categoria funcional (ou um conjunto de categorias funcionais) que encabeça a estrutura sintática, dominando a categoria N (Szabolcsi, 1982; Abney, 1987; Szabolcsi, 1994; Adger, 2003; Salzmann, 2020). A hipótese do DP foi formalmente introduzida por Abney (1987) e se tornou o modelo padrão para

análise sintática dentro da Teoria Gerativa, especialmente no Programa Minimalista (Chomsky, 1995). Assumiremos, portanto, que a estrutura sintática de uma expressão nominal é composta por pelo menos um exemplar da categoria lexical N e pelo item funcional D⁶, como representado na figura abaixo (porém ver discussão em nota de rodapé sobre a estrutura de pronomes, a qual deixa claro que a estrutura interna de DPs é mais complexa).

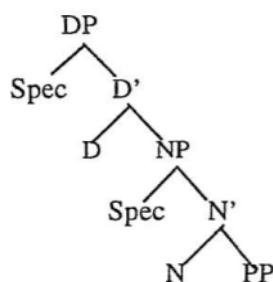


Figura 1. Estrutura do DP proposta por Abney (1987).

A proposta de Abney seguiu as ideias de Chomsky (1986) de que sentenças contêm uma projeção funcional de flexão (*IP - Inflection Phrase*). Abney propôs que expressões nominais, assim como sentenças, também projetam uma categoria funcional, o DP, aproximando, portanto, a estrutura nominal da estrutura sentencial. Dentre as propriedades oracionais que também são encontradas em sintagmas nominais, podemos observar, por exemplo, presença de concordância morfológica com traços *Phi*, marcação de tempo em algumas línguas, e restrições sobre extração de movimento qu-.

Abney propôs que o DP é a categoria funcional máxima projetada pelo substantivo lexical e que hospeda elementos da categoria dos determinantes como, por exemplo, artigos. O autor argumentou que determinantes, no domínio

⁶ É importante ressaltar que Salzmann (2020) e outros estudos como Bruening (2009) e Müller (2007) apontam que os argumentos presentes em Abney e estudos afins acerca da hipótese do DP residem em premissas teóricas que perderam força nos estudos sintáticos: apoiam-se em argumentos dentro do contexto da Teoria de Regência e Ligação e suas suposições acerca da estrutura dos sintagmas na época, assim como em paralelismos entre DP e CP. Algumas dessas premissas, segundo Salzmann, podem ser revisadas e pensadas em favor da categoria N, como pode ser visto na hipótese de assimetria do NP proposta por Bruening (em nominais como *too many/*much people*, em inglês, a categoria N se sobrepõe à categoria D e é o substantivo que dita a forma dos elementos dentro do sintagma). Entretanto, como esses estudos não são conclusivos (o autor, por exemplo, indica que há argumentos tanto para o encabeçamento por N quanto por D), assumimos a hipótese do DP.

nominal, são análogos aos modais no domínio oracional. Isso leva a hipóteses mais abrangentes de que toda expressão nominal definida possui o status de DP, o qual é constituído de categorias lexicais (NP) e categorias funcionais (Longobardi, 2001; 2005).⁷

Dessa forma, a hipótese do DP tem sido relevante para o entendimento de propriedades semânticas como, a exemplo, dêixis, definitude, especificidade e referencialidade (Abney, 1987; Knud, 1994; Chomsky, 1995; Lyons, 1999; Bernstein, 2008). A construção e interpretação da referência dependem da estrutura sintática do DP, ou seja, são dependentes dos processos combinatórios da gramática. Isto pode ser evidenciado na observação de que uma expressão nominal como “o cachorro” pode funcionar referencialmente de uma forma que apenas a palavra “cachorro” como um lexema ou conceito geral não pode. O item lexical puro “cachorro” não pode distinguir traços referenciais como, por exemplo, definitude, instanciado pela categoria D. Assim sendo, podemos dizer que a referência a indivíduos e objetos se dá através da construção da estrutura sintática de DPs, onde itens lexicais se combinam com itens funcionais (Bernstein, 2008; Longobardi, 2001; 2005). Nessa visão, portanto, a referencialidade depende da estrutura sintática e, dentro dessa perspectiva, DPs indefinidos e definidos diferem-se em estrutura sintática. A literatura aponta que DPs indefinidos possuem menos robustez estrutural (i.e., menos camadas estruturais) do que DPs definidos (Zamparelli, 2004).

Pode-se, portanto, concluir que as propriedades semânticas de expressões nominais refletem a estrutura sintática subjacente, pressupondo que expressões

⁷ Observa-se, em trabalhos como Bernstein (2008), que estudos acerca do DP chegaram ao entendimento de que DP projeta outras categorias internas além daquelas apresentadas em Abney (1987). Bernstein expõe que estudos acerca da estrutura oracional evidenciaram projeções funcionais correspondentes a concordância, tempo e negação, e que isso levou a paralelismos acerca de projeções funcionais internas ao DP e à sentença. A partir disso, novas propostas foram feitas sobre as projeções internas do DP que correspondem a, por exemplo, número (e.g. singular e plural), gênero e caso, assim como uma categoria funcional projetada para elementos quantificadores. Cinque (1994) propõe uma hierarquia universal de adjetivos através de uma estrutura funcional articulada, enquanto Ritter (1991, 1993) propõe uma projeção funcional correspondente à marcação de singular e plural de substantivos, denominada pela autora NumP, como complemento de D em seu estudo sobre a língua hebraica moderna. Picallo (1991, 2007) reivindica que gênero projeta uma categoria funcional dentro do DP, que a autora denominou GenP, localizada entre NP e NumP e refletindo o fato de que gênero é expresso diretamente sobre o radical de substantivos e número é expresso fora desse domínio. Danon (2001, 2010) argumenta que DPs possuem traços formais de definitude, argumento embasado, por exemplo, na presença do marcador de objeto *et* em hebraico, que atrela Caso à definitude. Reconhecemos os trabalhos teóricos citados, mas, dado o escopo e os objetivos da nossa pesquisa, não abordamos em detalhes a estrutura interna de DP.

nominais são compostas por categorias lexicais (e.g. N) combinadas com categorias funcionais (e.g. D).

3.2.1.

Pronomes: interface sintaxe-semântica

Os pronomes pessoais formam, em todas as línguas, um pequeno conjunto de elementos funcionais, universais (encontrados em todas as línguas), utilizados pela gramática para estabelecer referência e coerência discursiva. Seu uso é homogêneo na linguagem humana como substitutos de DPs plenos já mencionados no discurso (informação velha). Pronomes são estruturas de DP sem um determinante explícito (Longobardi, 1994; 2001; 2005), como exemplificado em (8) e (9) abaixo, retirados de Bernstein (2008, p. 1249). (8) é um DP definido com presença de um determinante. (9) também é um DP definido, mas sem presença de um determinante. Ou seja, pronomes e determinantes estão em distribuição complementar, o que nos leva, portanto, à conclusão de que pronomes ocupam o núcleo de D na estrutura sintática (Longobardi, 1994).

- (8) a. the linguist(s)
- b. [DP [D' the [NP [N' linguists]]]]
- (9) a. (*the) we linguists
- b. (*the) you linguists
- c. [DP [D' we/you [NP [N' linguists]]]] ⁸

A saliência de uma determinada entidade no contexto discursivo influencia sua possibilidade de ser representada por uma forma pronominal: quanto mais saliente for, maior a chance. Isso ocorre para evitar a repetição de nomes (Ariel, 1990; Gordon et al., 1993; Almor et al., 1999; Gordon et al., 2004; Saab, 2014). Repetições excessivas, quando não devidamente licenciadas pelo contexto discursivo, geram estranhamento pragmático ou agramaticalidade. Portanto, é

⁸ Observe que estruturas como (9c) são possíveis apenas com pronomes de 1 & 2Pessoa. Isso indica que pronomes de 1 & 2Pessoa são estruturalmente diferentes de pronomes de 3Pessoa. Consideramos aqui, com base na literatura sobre pronomes (ver Johnson (2013), entre outros), a possibilidade de pronomes de 1 & 2Pessoa serem estruturas mínimas contendo apenas D. Pronomes de 3Pessoas são estruturas maiores, contendo outros conjuntos de traços.

obrigatório o uso de pronomes para fazer referência a uma entidade saliente no discurso sempre que substituir um pronome por um DP pleno não adicionar algo ao significado da sentença (Schlenker, 2005). Essa restrição é denominada *Penalidade do Nome Repetido* (*Repeated Name Penalty*) e foi proposta por Gordon et al. (1993).

Pronomes, sob essa concepção, são considerados itens de substituição e estão sujeitos à necessidade de um antecedente, introduzido na sentença ou discurso. Em uma sentença como (10) abaixo, por exemplo, o pronome em negrito não consegue recuperar um referente sem alguma informação prévia que introduza um possível antecedente. Entretanto, se o mesmo pronome, na mesma situação, for acompanhado de um ato pragmático (como o de apontar, por exemplo), poderá estabelecer referência com uma entidade apresentada por meios extra-linguísticos, tornando seu uso normal (Fetzer, 2012).

(10) **Ele** gosta muito de futebol.

Ademais, as formas pronominais variam de acordo com a posição de antecedentes no nível da sentença. Quando o antecedente se encontra no mesmo domínio oracional que o pronome, a forma pronominal é anafórica (Chomsky 1981; 1986), como em (11a), onde o antecedente é o sujeito da oração subordinada. Quando o antecedente se encontra no domínio do discurso ou dentro da sentença, mas em outro domínio oracional que contém o pronome, então a forma pronominal é não-anafórica (11b). *Her* em (11b), objeto da oração subordinada, pode se referir ao sujeito da oração matriz ou a um antecedente citado no nível do discurso, mas não pode se referir ao sujeito da oração subordinada. Em (11a), dizemos que *herself* é correferente com *Carol*, e em (11b) *her* é correferente com *Ana*.

- (11) a. Ana₁ said that Carol₂ hit **herself**_{*1/2}
 b. Ana₁ said that Carol₂ hit **her**_{1/*2/3}

O contraste em (11) indica que pronomes anafóricos possuem distribuição sintática e restrições semânticas diferentes de pronomes não-anafóricos. Tais

diferenças se devem a restrições estruturais mediadas pela gramática. Sabe-se desde investigações conduzidas durante os anos 80, sobretudo desde a publicação de Chomsky (1981), que a interpretação dos pronomes pode ser sintaticamente condicionada. A Teoria da Ligação (Chomsky, 1981; 1986) evidencia o condicionamento sintático dos dois tipos de pronomes em discussão, anáforas, regidas pelo Princípio A da teoria de Regência e Ligação, e pronomes, regidos pelo Princípio B. Portanto, a Teoria da Ligação procurou descrever e explicar a forma como as anáforas e os pronomes adquirem referência, explicitando os contextos sintáticos que os legitimam:

Binding Theory

(A) An anaphor is bound in its governing category

(B) A pronominal is free in its governing category

(Chomsky, 1981, p. 188)

O Princípio A estabelece que uma anáfora tem necessariamente de estar ligada (compartilhar índice/referência) a um DP dentro do menor domínio sentencial que a contém. O Princípio B, por outro lado, estabelece que um pronome não pode estar ligado a um DP dentro do menor domínio sentencial que o contém. Portanto, os Princípios A e B regulam as estruturas sintáticas específicas que legitimam (11a) e (11b) como anáforas e pronomes, respectivamente. Observe que esses dois Princípios atuam de formas opostas, ou seja, onde A é aplicado, não se aplica B. Observe ainda que as anáforas não são elementos dêiticos, que podem fazer referência a uma entidade apresentada no nível do discurso, mas pronomes podem, embora nem todos os pronomes atuem da mesma maneira, como veremos na próxima seção.

A presente pesquisa investiga apenas pronomes não-anafóricos.

3.2.2.

Tipos de pronomes, estrutura sintática e interpretabilidade

Cardinaletti & Starke (1994) centram-se nas propriedades sintáticas dos pronomes (não-anafóricos), considerando as seguintes classes: pronomes fortes, clíticos e pronomes fracos, segundo exemplificados em (12):

- | | | | |
|------|----|----------------------------|-------------------|
| (12) | a. | A Ana falou com ele | (Pronome forte) |
| | b. | A Ana o viu | (Pronome clítico) |
| | c. | Ø vi a Ana | (Pronome fraco) |

Os autores caracterizam sintaticamente essas três classes da seguinte forma:

i) clíticos e pronomes fracos, em contraposição a pronomes fortes, são estruturalmente “deficientes”, ou seja, possuem menos traços formais, semânticos e fonológicos, e projetam menos categorias funcionais do que pronomes fortes. Pronomes fortes, em oposição, são os únicos a possuir a categoria funcional C, que contém traços referenciais (ver figura 2 abaixo). Além disso, também possuem uma categoria funcional, intermediária (Σ) entre C0 e I0, que contém traços de polaridade e foco. Σ é uma categoria projetada por pronomes fortes e fracos. Clíticos, por sua vez, possuem apenas a categoria mais baixa, I;

ii) pronomes fortes e fracos são projeções máximas (XP), enquanto clíticos são projeções mínimas (X^0);

iii) Pronomes fracos e clíticos são afetados por três tipos de restrições resultantes de sua estrutura empobrecida: (a) não podem introduzir novos referentes no discurso, (b) não podem ser coordenados e (c) não podem ser modificados.

As representações abaixo ilustram as estruturas atribuídas aos três tipos de pronomes:

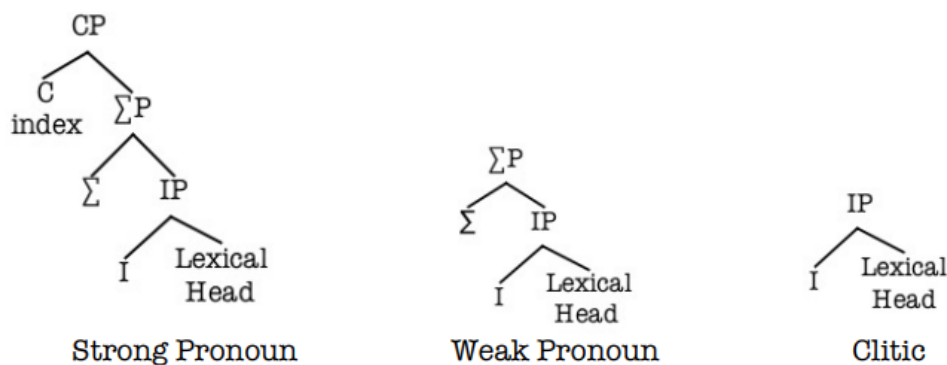


Figura 2. Diferença estrutural - pronomes fortes (*Strong Pronoun*), pronomes fracos (*Weak Pronoun*) e clíticos (*Clitic*).

Em decorrência das características apresentadas, formas pronominais fracas e clíticos apresentam preferência por ter leitura correferencial, isto é, uma leitura que não é dêitica, em que o pronome está referencialmente vinculado a um antecedente presente no nível da sentença. Em oposição, formas fortes têm maior liberdade de interpretação, apresentando preferência por interpretações dêiticas, ou seja, referência a entidades citadas no nível do discurso (Cardinaletti & Starke, 1994; Burzio, 1999).

Essas diferenças entre tipos de pronomes têm reflexos também no processo de aquisição da linguagem. Primeiro, crianças em aquisição da linguagem apresentam um comportamento semelhante ao de adultos no que tange a anáforas já por volta dos 4 anos de idade. Porém, observa-se um atraso no domínio dos pronomes (Chien & Wexler, 1990; Hartman et al., 2012; Brunetto, 2018). Segundo, o atraso em questão não se aplica a todas as formas pronominais. Crianças apresentam mais erros no uso e interpretação de pronomes fortes do que de pronomes fracos e clíticos, que estabelecem referência dentro do nível da sentença. Essa observação foi verificada em diferentes línguas (Solan, 1987; Avrutin & Wexler, 1992; McKee, 1992; Hestvik & Philip, 2000).

McKee (1992) realizou dois estudos experimentais (tarefa de julgamento de valor de verdade), um sobre anáforas e outro sobre pronomes, e, em cada um deles, realizou dois experimentos, um com crianças falantes de inglês e outro com crianças falantes de italiano. O estudo envolveu formas pronominais fortes e fracas. O grupo de crianças falantes de italiano foi composto por 30 crianças entre 3,7 e 5,5 anos, e o grupo de crianças falantes de inglês foi 60 crianças entre 2,6 e 5,3 anos. Os dois grupos participaram de todas as tarefas experimentais. Os

resultados indicam que, em relação a anáforas, não há diferença intergrupar. Em relação a pronomes, houve diferença intergrupar acerca de pronomes fortes: as crianças italianas tiveram uma taxa menor de acertos na correferência de pronomes fortes. No entanto, com clíticos não houve queda na taxa de acertos.

Também contrapondo pronomes (fortes e fracos (nulos)), Costa & Ambulate (2010) apresentam um estudo experimental em português europeu com o intuito de testar se diferenças entre pronomes fortes e fracos estão associadas a problemas de correferência. O experimento teve como foco pronomes na posição de sujeito como em (13). Os resultados indicam que as crianças em processo de aquisição do português europeu apresentam dificuldades no estabelecimento de correferência com pronomes fortes/plenos, mas não com pronomes fracos/nulos. Em estímulos como (13a), as crianças se comportam como adultos, interpretando o pronome nulo como correferente ao sujeito da matriz. Entretanto, em estímulos como (13b), as crianças aceitam, em oposição aos adultos, correferência entre o pronome forte e o sujeito da matriz, em violação, portanto, ao Princípio de Evite Pronome (*Avoid Pronoun Principle* - Chomsky, 1981).⁹

- (13) a. O Noddy disse que ____ tinha fome.
b. O Noddy disse que ele tinha fome.

(Costa & Ambulate, 2010, p. 7-8)

Esses resultados vão ao encontro da teoria de Reinhart (2006) de que a referência livre de pronomes plenos é mais difícil por exigir maior carga memória de trabalho. Para a autora, na interface sintaxe-pragmática, a presença de um pronome livre/forte (i.e., que não está vinculado a outro DP) ativa um processo de comparação derivacional, forçando a construção de uma outra representação sintática em que o pronome está vinculado. Esse processo de construção de representação para fins comparativos e o processo de comparação em si sobrecarregam a memória de trabalho, que em crianças ainda é reduzida.

⁹ Ver Grolla (2005) para dados experimentais do português brasileiro.

3.2.3.

Pronomes em inglês

O inglês, língua objeto de nossa investigação, não apresenta com clareza a distinção tripartite proposta por Cardinaletti & Starke (1994), já que os pronomes nulos nesta língua, na gramática do adulto, estão restritos aos casos chamados de *pro-drop* de diário (Haegeman & Stark, 2021). Não há também no inglês pronomes clíticos como aqueles observados nas línguas românicas. No entanto, como observado em Fiengo & Higginbotham (1981), o inglês contém formas pronominais fracas, formas fonologicamente reduzidas como em (14a), em oposição a formas fortes (14b).

- (14) a. John₁ said that Peter₂ admires ‘m₁
b. John₁ said that Peter₂ admires him₃

Em semelhança ao que é observado nas línguas românicas, em inglês, pronomes fracos apresentam preferência por leitura correferencial: ‘m em (14a), por exemplo, é preferencialmente interpretado como referindo a *John*, sujeito da sentença matriz. Em oposição, pronomes fortes têm preferência por leituras dêiticas. Assim, *him* em (14b) é preferencialmente interpretado como referindo-se a uma entidade mencionada no contexto discursivo e não na sentença.

Em relação à aquisição, as observações feitas para as línguas românicas se aplicam ao inglês. Primeiro, por volta de 3 anos de idade crianças nativas do inglês são capazes de interpretar corretamente anáforas (15a), mas apresentam dificuldades com pronomes (15b), atribuindo correferência com antecedentes locais, em violação ao Princípio B. (Solan, 1987; Chien & Wexler, 1990; Grodzinsky & Reinhart, 1993; Brunetto, 2018).

- (15) a. Peter₁ hurt himself_{1/*2}
b. Peter₁ hurt him_{1/2}

Estudos de Hartman et al. (2012) indicam, no entanto, que a dificuldade de aquisição é relativa somente a pronomes fortes. Os autores realizaram um experimento de julgamento de valor de verdade com 13 crianças (média de idade

de 4 anos e 11 meses) em língua inglesa, envolvendo sentenças com pronomes plenos reduzidos na posição de objeto, como em (16):

- (16) a. I think... Cow washed him
b. I think... Cow washed'm

Os resultados indicam que formas reduzidas em inglês se comportam de uma forma parecida com a aquisição de clíticos em línguas românicas. Crianças não aceitam a correferencialidade do pronome em (16b) com o DP *Cow*, sujeito da sentença encaixada, embora aceitem a mesma correferência em (16a).

Em resumo, a linguística teórica aponta que pronomes possuem subclassificações em relação à distribuição sintática e à interpretação semântica. Essas subclassificações têm sido estudadas especialmente durante o processo de aquisição da linguagem, e as investigações indicam atraso de aquisição de pronomes fortes, possivelmente devido ao não amadurecimento dos processos interpretativos de formas fortes na interface entre sintaxe e pragmática (Reinhart, 2006). Essas considerações tornam possível um novo entendimento do papel desempenhado por pronomes na gramática humana e nos permite ter um melhor entendimento das dificuldades linguísticas observadas em distúrbios mentais, como a esquizofrenia, contribuindo concomitantemente para um melhor entendimento do distúrbio.

Vale ressaltar que o ocorre em (14a) e (16b) acima não pode ser facilmente observado na língua escrita, já que, em inglês, o enfraquecimento fonológico dos pronomes ocorre com frequência na língua oral, mas não tem correspondência clara na escrita. No entanto, embora o inglês não seja uma língua *pro-drop*, sujeitos pronominais nulos, também chamados de elipses ou omissão de sujeito (Haegeman, 2019, Haegeman & Stark, 2021), ocorrem tanto no registro oral como no registro escrito, estando restritos a alguns poucos contextos sintáticos, particularmente na posição de sujeito de orações matrizes, como ilustrado pelos exemplos em (17), retirados de Haegeman (2019, p. 1). Não é claro, no entanto, que em sentenças como (17c), coordenação de orações, temos sujeito nulo na segunda oração coordenada, já que é possível analisar essas coordenações como

coordenação de VPs (McNally, 1992), com apenas um sujeito sendo projetado na estrutura sintática.

- (17) a. (I) wish you were here.
 b. (He) conducts his last symphony with the Zurich Tonhalle Orchestra. Shortly after, (he) dies in his sleep, aged 84, in France. (*Observer* 9.9.12, p.17 c. 2)
 c. Up at 6:45 – (It) rained all day & (it) rained hard at times. (*Sunday*, October 4, 1964)

A omissão de sujeito ocorre com frequência durante a aquisição do inglês (Hyams, 1986, 1991; Rizzi, 1994). Os exemplos abaixo, retirados de Hyams (1986, p. 65-66), ilustram a ocorrência de sujeitos nulos por crianças em aquisição do inglês.

- (18) a. Throw away
 b. want more apples.

Em nossa investigação (ver capítulo 4), analisamos a ocorrência de pronomes fortes e pronomes fracos (nulos) no inglês em posição de sujeito (pronomes fortes e nulos) e em posição de objeto direto (apenas pronomes fortes). Como veremos abaixo, investigações sobre o uso de pronomes por falantes esquizofrênicos de línguas românicas com sujeito nulo (línguas *pro-drop*) evidenciam maior uso de pronomes nulos por esquizofrênicos. Assim, cabe à presente pesquisa investigar se, em línguas não *pro-drop*, como o inglês, ocorre uso significativo de sujeitos nulos na esquizofrenia.

3.3.

Pronomes em gramáticas atípicas

3.3.1.

Envelhecimento atípico

Dificuldades com pronomes foram observadas também em casos de envelhecimento atípico (Kempler, 1995; Almor et al., 1999; Drummond et al., 2015). Almor et al. (1999) realizaram 3 estudos experimentais com o objetivo de investigar se pacientes com Alzheimer apresentam problemas na produção e compreensão de pronomes. O primeiro, que envolvia uma tarefa de produção de discurso espontâneo, apontou que participantes com Alzheimer (11 participantes) produziram uma razão significativamente maior de pronomes no lugar de expressões nominais plenas, em comparação ao grupo controle (9 adultos saudáveis pareados em idade). No segundo experimento, 10 participantes com Alzheimer e 10 controles (pareados em idade) realizaram uma tarefa, consistindo em escolher a melhor alternativa possível (entre duas opções de pronomes diferentes) para completar de forma apropriada uma história contada através de estímulo auditivo. Os resultados obtidos indicam que pacientes com Alzheimer apresentam menor sensibilidade a informações gramaticais necessárias para o processamento de pronomes, fazendo maior uso de pronomes inapropriados para complementar os estímulos auditivos. Em (19), por exemplo, observou-se maior número de escolhas da forma em (19c), forma inapropriada, por participantes do grupo alvo.

- (19) a. The children loved the silly clown at the party. The show was very funny. During the performance the clown threw candies to _____.
- b. them.
- c. him.

O terceiro experimento demonstrou que pacientes apresentaram melhor performance em manter uma informação referencial latente, presente em parágrafos curtos, quando a referência era representada por uma expressão nominal plena. Em oposição a isso, o grupo controle teve uma performance

oposta, apresentando maior facilidade em manter latentes informações referenciais representadas por pronomes durante os mesmos fragmentos de parágrafos curtos.

Os resultados dos testes linguísticos foram correlacionados aos resultados de testes de mensuração de memória semântica e de trabalho (Folstein et al., 1975), observando-se correlações fortes com memória de trabalho, mas não com memória semântica. Estes resultados levaram à conclusão de que prejuízos de memória de trabalho, apresentados por pacientes com Alzheimer, ocasionam problemas em manter latentes informações necessárias para interpretar adequadamente a referência de formas pronominais.

Drummond et al. (2015) analisaram a construção de narrativas, baseada em uma sequência de ações visualmente apresentadas, com 77 participantes nativos do português brasileiro: 41 controles, 22 pacientes com comprometimento cognitivo leve amnésico (a-MCI) e 14 pacientes com Alzheimer, todos pareados em idade e em língua nativa. Os parâmetros analisados foram os seguintes: tempo de narrativa; número total de palavras; tipo de discurso (predominantemente narrativo vs. predominantemente discursivo); coerência geral; coesão referencial (de acordo com os Princípios A, B e C da Teoria da Ligação); índice de eficácia narrativa (obtido da razão do número de palavras pelo número de macroproposições: menos palavras com mais macroproposições caracterizam um melhor índice); e estrutura narrativa (habilidade de formar uma história completa e coesa). O grupo a-MCI demonstrou dificuldades discursivas estabelecidas como um estágio médio entre a performance de controles e do grupo com Alzheimer, divergindo de controles com respeito a coerência global, tipo de discurso e coesão referencial. Os grupos a-MCI e Alzheimer tiveram performances semelhantes, mas apresentaram diferenças em relação a controles nas categorias de palavras recordadas, repetição de palavras na mesma sentença, estrutura narrativa e inclusão de proposições irrelevantes. Os autores discutem que a tarefa empreendida revelou diferenças entre controles saudáveis e grupos com prejuízos cognitivos, inclusive apontando que pacientes com a-MCI, um prejuízo cognitivo de menor grau, já apresentam déficits narrativos, apesar de menos severos do que aqueles apresentados por pacientes com Alzheimer.

3.3.2.

Desenvolvimento atípico

Dificuldade na compreensão e uso de pronomes também são relatados em transtornos do espectro autista (TEA). A literatura aponta troca do traço de pessoa pronominal (*Pronoun shift*) (Tager-Flusberg et al., 2005; Wetherby & Woods, 2006; Luyster & Lord, 2009; Naigles et al., 2017), como em (20), em que a segunda pessoa é utilizada com referência ao locutor (primeira pessoa):

- (20) You want some more milk.
(Intended meaning: I want some more milk)

Hinzen (2017) discute que as diversas questões neurocognitivas do TEA se relacionam com a linguagem e seu uso referencial, observando que dificuldades com o uso anafórico de sintagmas nominais definidos e pronomes são destacados no transtorno. Além disso, também aponta que há relatos de troca de pessoa pronominal, estabelecimento de referência por meio de gestos e dificuldade de entendimento de termos dêiticos referenciais.

Banney et al. (2015) compararam, em tarefa de construção narrativa, um grupo autista entre 9 e 15 anos com um grupo neurotípico, ambos combinados em habilidades linguísticas e inteligência não-verbal. Os resultados encontrados indicam que autistas apresentam menor complexidade sintática (calculada a partir do número de orações por unidades terminais mínimas), menor diversidade sintática (quantificada pelo número de sentenças complexas (sentenças com dois ou mais verbos principais e categorizadas em relação ao tipo)), e maior número de pronomes ambíguos.

Em relação a pronomes de 3ª Pessoa referencial, a literatura aponta que estes pronomes são evitados por autistas, com preferência por repetição de DPs plenos, apresentando, portanto, violação da Penalidade do Nome Repetido. Novogrodsky (2013) relata resultados de tarefas de contar e recontar histórias, realizadas com 23 crianças diagnosticadas com autismo (de 6 a 14 anos) e 17 crianças controle, em desenvolvimento típico (de 5 a 14 anos). Esses experimentos demonstraram que crianças com autismo apresentam dificuldades em estabelecer antecedentes para pronomes, apresentando uso ambíguo de pronomes, especialmente em relação à

terceira pessoa pronominal. Foi apresentado, por crianças autistas, um número significativamente maior de pronomes sem antecedentes claros ou intencionais, caracterizando anormalidades referenciais, como pode ser visto em (21):

- (21) a. No antecedent: “Once upon a time there was a frog” and he said “frog where are you?” (In the picture, a boy is looking for the frog.)
- b. A different referent: “The bees were chasing the dog. He had climbed up on rock and went into a tree”. (In the picture, the boy climbed up the tree, not the dog.)
- c. Agreement error: “and they said mom”, “can I go outside to play?”

(Novogrodsky, 2013 p. 88)

Novogrodsky & Edelson (2015) exploraram a produção pronominal de terceira pessoa em posição de sujeito, objeto e possessivo, além de habilidades sintáticas (complexidade sentencial e erros morfossintáticos) gerais de 24 crianças autistas (entre 6 e 14 anos) em tarefas de contar e recontar histórias, comparadas com 17 crianças em desenvolvimento típico (de 5 a 14 anos). Diferenças de grupo em habilidades sintáticas não foram significativas, mas as medidas de uso de pronomes mostraram que crianças autistas produzem mais pronomes ambíguos do que crianças do grupo controle.

Hobson & Lee (2009) realizaram estudo para avaliar se crianças autistas produzem e compreendem expressões e gestos atípicos com função dêitica. O primeiro estudo foi realizado com 20 crianças e adolescentes com autismo, comparados a 20 crianças e adolescentes controle, sem diagnóstico específico, mas com retardamento mental, consistindo em tarefa de apontamento de posicionamento de objetos (animais de plástico). A maior parte dos participantes produziu termos dêiticos, mas o grupo autista apresentou incongruências no uso de pronomes demonstrativos como, por exemplo, *this*, utilizando-os para fazer referência a locais distantes. O segundo estudo envolveu 15 jovens com autismo e 15 sem, porém, também com retardamento mental. O teste também envolveu o posicionamento de objetos e teve uma parte verbal e uma não-verbal, além de uma triagem. Na parte verbal, o teste envolvia o uso de termos dêiticos (8 termos)

semelhantes em formas e instruções como em (22). O resultado indica pior performance em autistas em responder a instruções envolvendo termos dêiticos contrastivos para perspectiva (*this* vs. *that*)

(22) Put a duck in THAT field and put a horse in THIS field.

De acordo com Hinzen (2017), a dificuldade principal presente no TEA parece residir na referenciação dêitica e no traço de Pessoa gramatical.

3.3.3.

Pronomes em esquizofrenia

O comportamento verbal em esquizofrenia é marcado por formas pronominais com diversas peculiaridades: (a) ambiguidade referencial: diversos possíveis referentes para um mesmo pronome (Barch & Berenbaum, 1996; Docherty et al., 2003; Ditman & Kuperberg, 2010); (b) ausência de referência: nenhuma menção prévia a algum possível referente para um determinado pronome (Rochester & Martin, 1979; Frith, 1992; Ditman & Kuperberg, 2010; Chaves et al., 2020); (c) evasão pronominal: preferência pelo uso de DPs plenos em detrimento de pronomes (Rochester & Martin, 1979; Frith, 1992); e (d) troca de traços de pessoa: segunda pessoa ou terceira pessoa sendo utilizadas no lugar da primeira pessoa do discurso (Watson et al., 2012; Hinzen, 2017; Çokal et al., 2018).

Hoffman et al. (1985) aplicou, em 24 pacientes esquizofrênicos e 24 pacientes psiquiátricos não-esquizofrênicos, o teste de Hunt (Hunt, 1970), que consiste na reestruturação de uma sequência de 32 frases curtas que exigem que o participante utilize estratégias sintáticas para sintetizar o texto. As frases utilizadas no experimento tinham como tema a composição, aspecto e outras características do metal *alumínio*. Os participantes (controle e esquizofrênicos) foram expostos a frases curtas e desconectadas, e a tarefa consistiu em reescrever as frases dadas da melhor forma possível, sendo possível combinar frases, trocar a ordem das palavras e omitir palavras que ocorressem muitas vezes, mas não era permitido deixar de incluir alguma das informações presentes nas frases dadas. Os resultados apontam para dificuldades relacionadas a relações de correferência de

ordem pronominal em esquizofrênicos. As correferências tecidas pelos pacientes foram ambíguas e muitas vezes percebidas pelos próprios participantes e corrigidas através da inserção da expressão nominal plena, evitando, assim, o uso de pronomes. Em (23) apresentamos exemplos relatados em Hoffman et al. (1985, p. 192), onde a segunda ocorrência do pronome relativo é ambígua em (23a) e o pronome *it* (23b) não tem um referente claro.

(23) a. Aluminum is a metal, which is abundant. It has many uses, **which** comes from bauxite.

b. Aluminum is an abundant metal with many uses. Coming from bauxite, an ore, **it** contains other substances.

(Hoffman et al., 1985, p. 192)

Chaves (2022), considerando dados do português brasileiro, analisou narrativas de sonhos e de atividades de vigília em esquizofrênicos e controles. Os resultados sugerem maior uso de pronomes nulos e maior volume de pronomes nulos com anomalia referencial em esquizofrênicos. Esses resultados são semelhantes aos de Tovar et al. (2019b), que também apresentam resultados experimentais indicando que esquizofrênicos com FTD, nativos de espanhol, produzem mais pronomes nulos.

O estudo recente de Tovar et al. (2019a) investigou narrativas de alucinação de 19 pacientes esquizofrênicos que escutavam vozes frequentemente. Essas narrativas foram selecionadas para as variáveis: expressões nominais (primeira, segunda e terceira pessoa diferenciadas), subordinação, coordenação, adjuntos, parataxe, expressão nominal anafórica, conteúdo pessoal e impessoal (vago e não vago), vagueza de conteúdo, erro gramatical, parafasia, violações semânticas, neologismos e ecolalia. As variáveis foram individualizadas e contadas para comparações estatísticas. Os resultados indicam que (i) de forma consistente, pronomes de 1ª Pessoa gramatical foram menos usados do que pronomes de 2ª Pessoa e, em alguns casos, ausentes; (ii) não houve ocorrência de expressões nominais anafóricas, ilustrando falta de conectividade entre enunciados; (iii) houve um grande número de parataxes; (iv) o conteúdo no nível sentencial foi

amplamente pessoal em contraposição a impessoal; (v) não há taxa significativa de erros sintáticos e/ou semânticos.

Docherty et al. (1997) observam que pronomes com referências vagas, confusas ou ambíguas são comuns em esquizofrenia. O dado em (24), apresentado pelos pesquisadores (Docherty et al., 1997, p. 502) evidencia referência ambígua. O pronome *he*, sujeito da segunda sentença, pode estar se referindo tanto a George como a Lester, mencionados na sentença anterior.

(24) I saw George and Lester at the store. He looked very sad.

Chaika & Lambe (1989) e Frith (1992) também apontam para ausência de referentes para pronomes. Foram analisadas narrativas, criadas a partir de uma história em vídeo, de pacientes psicóticos (14 com esquizofrenia e 8 com mania) com desordens de fala. Esquizofrênicos, apresentaram, em contraste com pacientes com mania e grupo controle, maior uso de pronomes sem referência, introduzindo uma nova entidade no discurso, como em (25), onde a anomalia pronominal é marcada em negrito.

(25) and I didn't think **that** was fair the way the way **they** did **that** either, so that's why I'm kinda like asking could **we** just get together for one big party or something ezz it hey if it we'd all in which is in not **they**'ve been here, so why **you** jis now discovering it?

(Chaika & Lambe, 1989, p. 418)

Em relação a trocas de traços de pessoa, estudo experimental longitudinal realizado por Watson et al. (2012) indicam que indivíduos com risco de esquizofrenia (parentes de primeiro grau de esquizofrênicos) durante o período prodromal apresentam significativamente maior uso de pronomes de segunda pessoa no lugar de 1Pessoa (i.e., 1Pessoa por 2Pessoa). O experimento comparou os resultados de 4 grupos de participantes: (a) indivíduos com risco e que desenvolveram esquizofrenia (i.e., indivíduos em período prodromal); (b) indivíduos que tiveram sintomas psicóticos, mas que não desenvolveram esquizofrenia; (c) indivíduos com risco de esquizofrenia, mas que não

apresentaram experiência psicótica; (d) controles. Foram coletadas e comparadas avaliações em 3 testes (ressonância magnética, testes psicométricos e *Present State Examination* - PSE) ao longo de um período de 18 meses. A amostra linguística foi fornecida pelo teste PSE, que consistia em uma transcrição de narrativas de 200 palavras sobre saúde mental e física de modo geral. A partir da transcrição, a porcentagem do número total de palavras para cada variável (substantivos singulares e plurais, verbos no infinitivo, no tempo progressivo e passado, pronomes, artigos, conjunções e erros) foi obtida. Os resultados evidenciam maior troca pronominal da primeira pessoa por segunda pessoa no grupo (a), dando suporte à ideia de que pessoas que desenvolvem esquizofrenia apresentam precocemente disfunção linguística relativa a pronomes.

Outro experimento recente foi realizado por Çokal et al. (2018), que compararam a produção linguística de 30 participantes esquizofrênicos (15 marcadamente com +FTD e 15 -FTD) com 15 parentes de primeiro grau de esquizofrênicos e 15 controles saudáveis, com foco em expressões nominais definidas, incluindo pronomes, e complexidade sintática. Pacientes com esquizofrenia +FTD produziram o maior número de expressões nominais anômalas. Dentre essas anomalias, são relatadas trocas de traços de pessoa gramatical em pronomes, como ilustrado em (26):

(26) Resk saves the day. They come home.

(Çokal et al., 2018, p. 6)

De modo geral, pode-se dizer que problemas na construção da referência pronominal são mais salientes em casos de esquizofrenia com sintomas de FTD. Assim como Çokal et al. (2018), Rochester & Martin (1979) encontraram diferenças significativas entre pacientes esquizofrênicos com FTD, em comparação a pacientes sem FTD e ao grupo controle. Nas análises, foram considerados pronomes pessoais, pronomes demonstrativos e termos comparativos (e.g. *smaller than*, *bigger than*). Observou-se que esquizofrênicos com FTD têm maior dificuldade na construção de referência pronominal, em comparação com o grupo controle e com esquizofrênicos sem FTD. Resultados semelhantes são relatados em Harvey (1983), que comparou 20 maníacos e 40

esquizofrênicos (20 com FTD e 20 sem FTD) e 10 controles, todos pareados em idades. Observou-se também maior taxa de referências vagas e uma menor taxa de elementos de coesão no subgrupo +FTD.

3.4.

Pronomes na escrita

Estudos sobre linguagem em esquizofrenia baseados em textos escritos são bem mais escassos do que estudos baseados em narrativas orais ou em experimentos controlados. Ainda pouco se sabe sobre o impacto da esquizofrenia na modalidade escrita.

Fineberg et al. (2015) fizeram uso de *software* de contagem de palavras aplicado a relatos (em inglês) em primeira pessoa em esquizofrenia em comparação com desordem do humor e ansiedade. Os textos do grupo esquizofrenia foram compostos por 77 amostras descrevendo a experiência da esquizofrenia publicados no jornal *Schizophrenia Bulletin*. O grupo de desordem do humor e ansiedade foi considerado controle não-psiquiátrico e foi composto por 22 textos retirados da página *Anxiety and Depression Association of America*, também sobre relatos em primeira pessoa. Os autores compararam estatisticamente as amostras obtidas através de software de contagem considerando itens funcionais (e.g. artigos, pronomes e preposições) e categorias substantivas (i.e. nomes, verbos regulares e advérbios e adjetivos), totalizando 70 categorias. Acerca de categorias funcionais, observou-se que esquizofrênicos usam mais referências externas, fazendo maior uso de pronomes como *they* até mesmo em relatos de primeira pessoa. Em relação a categorias lexicais, esquizofrênicos fizeram mais uso de palavras de atribuição externa (e.g. “humano” e “religião”).

Birnbaum et al. (2017) realizaram um estudo com o objetivo de refinar a identificação de diagnósticos de esquizofrenia em postagens de mídias sociais através da exploração de técnicas de aprendizado de máquina, empregando análises da linguística computacional concatenadas com avaliações clínicas. Para isso, foram coletadas postagens da rede social Twitter, extraídas de 671 usuários com autodeclarações de diagnóstico de esquizofrenia, que foram então avaliadas em autenticidade por especialistas clínicos. Diagnósticos validados foram usados

para construir um classificador com o objetivo de distinguir usuários esquizofrênicos de controles saudáveis. A amostra final contou com 146 usuários do grupo alvo e 146 usuários do grupo controle, composto por uma amostra aleatória de usuários do *Twitter* sem menções de esquizofrenia ou psicose em suas postagens. O classificador, então, teve sua performance comparada a de avaliações de especialistas a partir da avaliação de dados novos usuários do *Twitter* não-avaliados. Cada postagem na página de cada participante foi analisada através de um modelo linguístico de n-Gramas e através da ferramenta LIWC, um programa de análise textual que processa contagens de palavras para classes semânticas e traços estruturais a partir de um dicionário interno que mapeia palavras, utilizado para identificar medidas linguísticas para as categorias: atributos afetivos (e.g. emoção positiva e negativa, raiva, ansiedade, dentre outros); atributos cognitivos, que incluem categorias de cognição (e.g. discrepâncias, inibição, negação e etc.) e categorias de percepção (e.g. ver, ouvir, sentir e etc.); e atributos de estilo linguístico, que compreendem densidade lexical (e.g. verbos, auxiliares, preposições, conjunções e etc.), referências temporais (passado, presente e futuro), preocupação social/pessoal (e.g. familiar, amizade, social, trabalho e etc.), foco e percepção interpessoal (pronomes de 1Pessoa, 2Pessoa, 3Pessoa do singular e do plural). Os resultados da análise apontaram diferenças linguísticas significativas no grupo esquizofrenia no uso pronominal, caracterizado por um maior uso de pronomes de 1Pessoa. Os resultados do classificador apresentaram uma taxa de acerto de 88% utilizando apenas dados linguísticos. Os autores, por fim, argumentam que o emprego de técnicas de avaliação de material escrito e colaborações de diferentes áreas de conhecimento no entendimento da esquizofrenia se mostram relevantes para um maior entendimento de distúrbios mentais como um todo.

Zomick et al. (2019) exploram indicadores linguísticos de esquizofrenia em fóruns de discussões na plataforma social *Reddit* através do empreendimento de técnicas estatísticas e de aprendizado de máquina. Para isso, coletaram e analisaram uma grande base de dados de postagens na plataforma social de usuários que declararam diagnóstico de esquizofrenia, comparados a grupo controle, totalizando um número de 159 usuários para cada grupo (318 ao todo) com um mínimo de 10 postagens, totalizando 66,454 comentários do grupo alvo e 113,570 do grupo controle. Marcadores linguísticos foram analisados através da

ferramenta *Linguistic Inquiry and Word Count* (LIWC). Ao analisar um texto, o programa busca por palavras-alvo no dicionário e computa frequências para cada uma das dimensões: i) porcentagem de itens funcionais (e.g. pronomes e artigos); ii) marcadores de processos psicológicos (e.g. afetação, palavras cognitivas e sociais); iii) pontuação; iv) e medidas de formalidade (e.g. palavras). Para realizar comparações, os autores empregaram t-testes independentes para determinar se a diferença média da frequência de cada categoria do programa LIWC era significativa para distinguir o grupo alvo do controle. Os resultados foram consistentes com estudos anteriores acerca da linguagem em esquizofrenia: esquizofrênicos utilizaram mais palavras associadas a problemas de saúde, ansiedade, emoções negativas, pronomes de 1Pessoa e de 3Pessoa no plural, além de palavras com semântica de incerteza (e.g. “talvez”).

Bae et al. (2021) realizaram um estudo semelhante com o objetivo de determinar se técnicas de aprendizado de máquina podem ser usadas de forma efetiva na determinação de marcadores linguísticos para esquizofrenia, em textos de usuários de mídias sociais. Para isso, os autores coletaram postagens do subreddit */r/schizophrenia* na plataforma *Reddit* utilizando a API *pushshift* para compor o grupo alvo, assim como postagens relacionadas a temas fora do escopo de saúde mental (*/r/jokes*, */r/fitness*, */r/meditation*, */r/parenting*, */r/relationships*, e */r/teaching*) para compor o grupo controle. O banco de dados final foi composto por 13,156 postagens do grupo esquizofrenia e 247,569 do grupo controle. Foram extraídos traços linguísticos e tópicos de conteúdo de todas as postagens através da ferramenta LIWC, que foram, então, classificados como pertencentes a cada grupo (alvo e controle) com técnicas de aprendizado de máquina não-supervisionado. Para esse fim, os textos foram processados e foram extraídas diversas categorias da ferramenta LIWC, como: contagem de palavras, contagem de palavras de mais de 6 letras, palavras funcionais (pronomes), orientações de tempo (passado, presente e futuro) e itens de referência a processos psicológicos (emoção positiva, emoção negativa, surpresa, raiva e outros). Os autores identificaram diferenças significativas entre os grupos, com a esquizofrenia estando associada a um menor número de contagem de palavras e uso de pronomes de primeira pessoa (plural), pronomes de terceira pessoa (singular), tempo verbal passado e palavras de emoção positiva. Além disso, pessoas que participaram do subreddit */r/schizophrenia* apresentaram maior uso da segunda

pessoa pronominal, terceira pessoa plural, pronomes impessoais, tempo verbal presente e palavras de emoção negativa.

Em resumo, podemos considerar, portanto, que anormalidades de referência pronominal, já bem estudadas no registro oral de pacientes esquizofrênicos, também ocorrem no registro escrito. Portanto, apesar de a escrita minimizar problemas linguísticos associados à redução de memória de trabalho (ver capítulo 2, seção 2.5), os problemas com pronomes persistem.

No próximo capítulo, apresentaremos o experimento realizado por nós em textos escritos no inglês, com foco em pronomes e complexidade sintática no nível sentencial.

4

Pronomes em esquizofrenia: análise de textos escritos no contexto de mídia social

Conforme já estabelecido, a presente pesquisa visa analisar o uso de pronomes pessoais na esquizofrenia, com observação adicional da complexidade sintática no nível sentencial. A amostra é composta por postagens de usuários autodeclarados esquizofrênicos na plataforma social *Reddit* e de controles sem declarações de transtornos mentais. Baseando-se no que foi apresentado e discutido nos capítulos anteriores, a análise foi conduzida considerando as seguintes hipóteses de trabalho (ver seção 1.3.):

(a') A esquizofrenia leva a perdas de complexidade estrutural tanto no nível da sentença como no nível nominal.

(b') Os pronomes são afetados tanto na forma como no significado.

(c') Essas dificuldades são observadas também na escrita, já que a escrita é uma tarefa cognitiva que induz aumento de memória de trabalho.

Dadas as hipóteses estabelecidas e nos parâmetros linguísticos analisados (ver seção 4.3.1), previmos que, em comparação ao grupo controle, o grupo alvo (esquizofrenia) apresentará:

(P1) Menor complexidade no nível sintático, com maior ocorrência de sentenças matrizes e maior ocorrência de sentenças truncadas anômalas

(P2) Maior ocorrência de sentenças truncadas

(P3) Maior ocorrência de truncadas anômalas

(P4) Maior ocorrência de pronomes nulos

(P5) Maior ocorrência de pronomes +Locutor (1Pessoa e 2Pessoa)

(P6) Maior ocorrência de pronomes não-referenciais

(P7) Maior ocorrência de pronomes anômalos

4.1.

Método de coleta de dados

A presente pesquisa foi aprovada em sua totalidade pela Câmara de Ética em Pesquisa da PUC-Rio (protocolo 111/2020).

Os dados foram coletados na interface de programação de aplicações (API) *pushshift* (<https://pushshift.io/>), um grande diretório de armazenamento de dados que copia todo o conteúdo da plataforma *Reddit*, ou seja, possui os mesmos dados de sua API. O uso da API *pushshift* se justifica por possuir um número reduzido de limitações para captura de dados (e.g. limite de dados disponíveis por tempo e de quantidade de dados coletados por vez), sendo considerada pela nossa equipe como a mais eficiente para os objetivos da investigação conduzida.

Visto que dados dispostos na API em questão são cópias dos dados dispostos na plataforma *Reddit*, considera-se que são, também, de domínio público, e todos os dados foram acessados e armazenados através de um código em linguagem Python elaborado pela nossa própria equipe.

A partir da API *pushshift*, selecionamos os usuários no subreddit */r/schizophrenia*, para composição do grupo alvo por meio de *flairs de inclusão* (figura 3)¹⁰ autoatribuição de esquizofrenia (ver Anexo 1). Observe, portanto, que as autodeclarações de esquizofrenia consideradas foram feitas em um contexto específico (*subreddit /r/schizophrenia*), onde o termo *schizophrenia* é tomado como um termo técnico, referindo-se à desordem mental correspondente. Consideramos, portanto, que é improvável que um usuário que realize autodefinição com termos/*flairs* associados à esquizofrenia, nesse contexto específico, esteja usando o termo escolhido em um sentido não carregado, mais informal (e.g. como sinônimo de “doido”).

A imagem mostra uma interface de usuário de uma postagem no Reddit. No topo, há uma barra cinza com o texto "Enviado por" seguido de um nome de usuário parcialmente visível. À direita, há um retângulo verde com o texto "Schizophrenia" em branco, seguido do texto "há 9 horas".

Figura 3. Exemplo de *flair* (*Schizophrenia* em verde).

Quando um usuário faz uma caracterização de si mesmo via *flair*, esta informação aparece em todas as suas postagens, sendo representada ao lado de seu

¹⁰ *Flairs* são semelhantes a rótulos que usuários podem atribuir a si mesmos dentro de um *subreddit* específico.

nome (omitido na Figura 3). Dessa forma, a autodeclaração por *flair* é de caráter identitário, diferentemente de uma pista lexical ou declarações realizadas pontualmente em um texto (postagem ou comentário). Por isso, priorizamos a autodeclaração por *flairs*.

Logo após, filtramos os usuários encontrados a partir de uma lista de *flairs de exclusão* (ver Anexo 2), de autodeclarações de outros distúrbios mentais com efeitos sobre a linguagem, de acordo com o DSM-V (APA, 2013). Os usuários que usaram quaisquer *flairs* presentes no Anexo 2 foram excluídos. Em seguida, capturamos todo o conteúdo postado pelos usuários selecionados (todas as postagens de tópico ou comentário presentes no perfil de cada um) e utilizamos códigos para a procura de informações de *background* (e.g. sexo, idade, nacionalidade) em suas postagens, utilizando listas de itens lexicais elaboradas pela nossa equipe (e.g. “man”, “woman”, “husband”, “autistic”, “twenty”, “birthday”, “America” e etc.). Com isso, fizemos uma nova procura por distúrbios conflitantes com a esquizofrenia (Anexo 3). Usuários com pistas lexicais confirmadas nessa etapa também foram excluídos. Após, procuramos por novas pistas lexicais acerca da idade (Anexo 4), sexo (Anexo 5) e língua nativa (Anexo 6) de cada usuário restante. Consideramos importante filtrar usuários não-nativos do inglês com o intuito de controlar efeitos de segunda língua ou falta de domínio da língua inglesa. As buscas foram primeiramente realizadas de forma automática com um código escrito em Python, que localizou as pistas lexicais e armazenou as postagens encontradas.

Decidimos não excluir da nossa amostra usuários que se declararam afetados por depressão e/ou ansiedade, visto que são condições relacionadas aos sintomas da esquizofrenia, como apresenta o DSM-V. Além disso, essa decisão se mostrou necessária para a seleção do grupo alvo e controle, pois a exclusão de usuários com autodeclaração para tais condições levaria a um corte de 44% (ver seção 4.2.) da amostra final.

O grupo controle foi capturado de forma a fazer o melhor pareamento possível com o grupo alvo em variáveis de *background*. Todos os usuários passaram pelas mesmas filtragens do grupo alvo. Usuários do grupo controle possuem participação nos mesmos *subreddits* ou *subreddits* de temas semelhantes aos usuários do grupo alvo a fim de evitar viés temático nos discursos coletados.

Observamos que foi feito pareamento também para autodeclaração de ansiedade e depressão.

Foram selecionados textos com um mínimo de 100 palavras, contadas pela ferramenta Word. Na contagem de palavras, contrações (e.g. *I'm*, *he's*, *don't*) foram expandidas através do uso da biblioteca *contractions*¹¹. A etiquetagem morfosintática foi realizada nas primeiras 500 palavras de cada texto (ver figura 4). Dentro desse limite, 11 dos 78 textos da amostra total não foram analisados inteiramente, pois ultrapassavam o limite de 500 palavras.

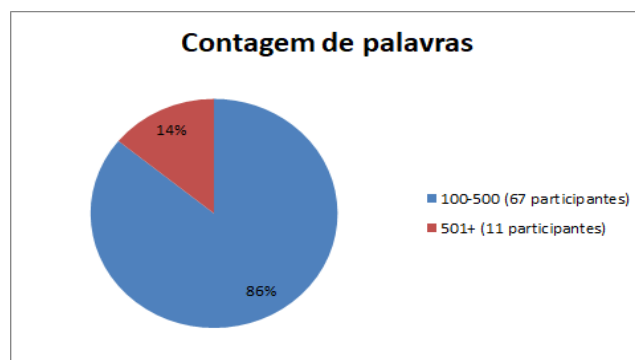


Figura 4. Gráfico com distribuição (em quantidade numérica e percentual) dos textos que pertencem ao grupo de 100 a 500 palavras e textos que ultrapassaram o limite de 500 palavras.

Para todos os usuários (alvos e controles), foram selecionadas postagens fora do subreddit */r/schizophrenia* com o intuito de controlar a interferência de viés temático/discursivo. Para cada postagem do grupo alvo em um determinado *subreddit*, selecionamos uma postagem de usuário do grupo controle que tenha sido realizada no mesmo *subreddit*, salvo um caso em que foi escolhida uma postagem realizada em *subreddit* de tema semelhante devido à impossibilidade de pareamento total (arbitrariedade dos dados), procurando controlar os temas/assuntos tratados nos dois grupos e evitar que escolhas de tópicos discursivos pudessem afetar o resultado da análise.

Dessa forma, em nossa amostra final tivemos, nos dois grupos (alvo e controle), a seguinte distribuição de postagens por *subreddit*:

¹¹ A biblioteca *contractions* (<https://pypi.org/project/contractions/>) expande todas as ocorrências com apóstrofe (e.g. *it's*), porém não expande as contrações sem apóstrofe (e.g. *its* quando se quer dizer *it's*, caso comum na escrita coloquial da internet). Nesses casos, corrigimos manualmente durante o processo de *tagging*.

Número de postagens coletadas por subreddit (grupo alvo)

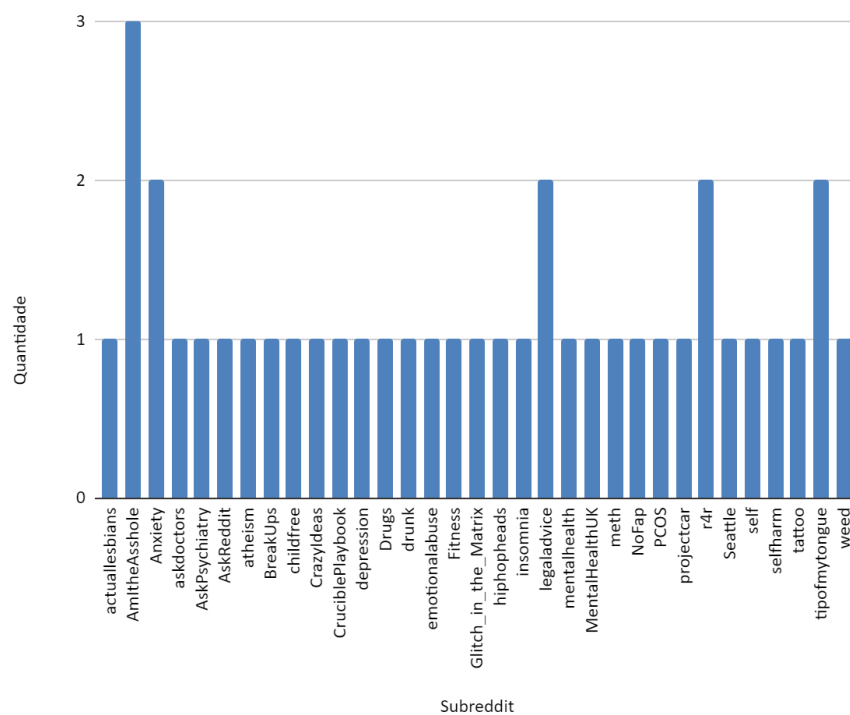


Figura 5. Número de postagens coletadas por cada *subreddit* selecionado (grupo alvo).

Número de postagens coletadas por subreddit (grupo controle)

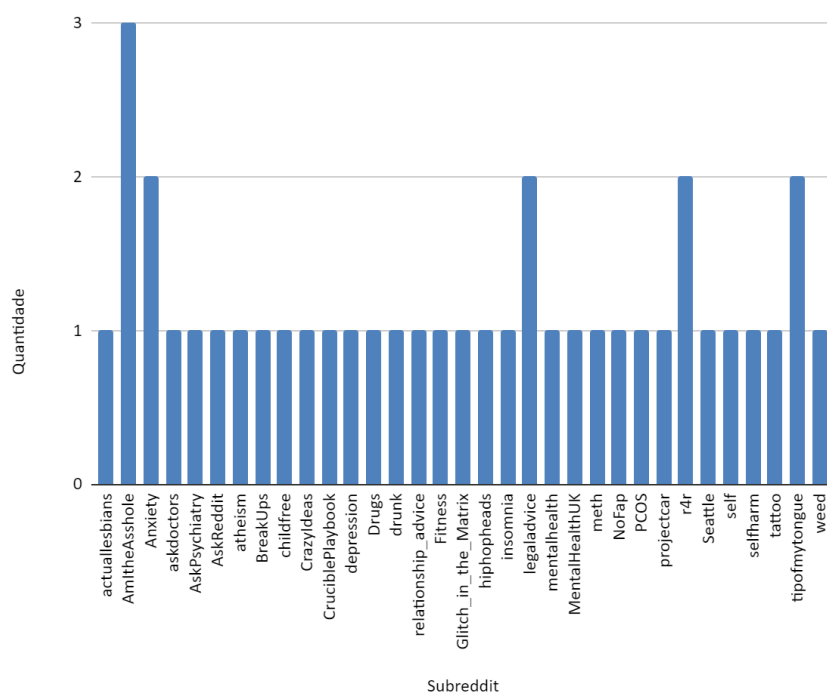


Figura 6. Número de postagens coletadas por cada *subreddit* selecionado (grupo controle).

Concluída a fase de coleta dos dados, preparamos os textos selecionados para o esquema de anotações (etiquetagem) manual adotado na pesquisa. Portanto, foram retirados dos textos todos os ícones para-linguísticos, como *emojis*.

4.2.

Participantes

A amostra final contém textos de 78 usuários (39 do grupo esquizofrenia e 39 do grupo controle), sendo a maioria do sexo masculino; isto é, com apresentação de pistas lexicais para sexo masculino. Isso se justifica pelo fato de a plataforma *Reddit* possuir uma maior quantidade documentada de usuários do sexo masculino (Finlay, 2014). Ambos os grupos possuem a mesma incidência de ansiedade e depressão e, também, agrupamentos de idade semelhantes, conforme indicado nas tabelas abaixo:

Distribuição por sexo (grupo alvo e controle)						
Sexo	Usuários (alvo)	%	Usuários (controle)	%	Total	%
N (não informado)	9	23%	8	20%	17	22%
M (Masculino)	20	51%	21	54%	41	52%
F (Feminino)	10	26%	10	26%	20	26%

Tabela 1. Distribuição por sexo (grupo alvo e controle).

Distribuição por idade (grupo alvo e controle)						
Grupo de idade	Usuários (alvo)	%	Usuários (controle)	%	Total	%
N (não informado)	4	10%	6	15%	10	13%
-18	1	3%	1	3%	2	3%
18-24	19	49%	17	44%	36	46%
25-34	13	33%	13	33%	26	33%

35+	2	5%	2	2%	4	5%
-----	---	----	---	----	---	----

Tabela 2. Distribuição por idade (grupo alvo e controle).

Média e desvio padrão por idades (grupo alvo e controle)		
Grupos	Média	Desvio padrão
Alvo	24,22	6,3
Controle	24,39	6,4
Ambos	24,30	6,4

Tabela 3. Média e desvio padrão por idades (grupo alvo e controle).

Distribuição por declarações de ansiedade e depressão (grupo alvo e controle)						
Declaração	Usuários (alvo)	%	Usuários (controle)	%	Total	%
N (sem declaração)	22	56%	22	56%	44	56%
Ansiedade	3	8%	3	8%	6	8%
Depressão	5	13%	4	10%	9	12%
Ambos	9	23%	10	26%	19	24%

Tabela 4. Distribuição por declarações de ansiedade e depressão.

4.3.

Análise dos dados

4.3.1. Parâmetros de análise

Os seguintes parâmetros de análise foram estabelecidos:

No nível sentencial:

- (a) Tipo de sentença em função da complexidade sintática (Matriz vs. Subordinada)
- (b) Tipo de sentença em função de truncamento (elisão) (+Truncada, vs. -Truncada)

(c) Tipo de truncamento sentencial (gramatical (-Anômalo) vs. agramatical (+Anômalo))

No nível nominal:

(d) Tipo de pronome em função da forma fonológica (Pleno vs. Nulo)

(e) Tipo de pronome em função de referencialidade (+Referencial vs. -Referencial)

(f) Tipo de pronome em função dos traços de pessoa (+Locutor (1Pessoa, 2Pessoa), vs. -Locutor (3Pessoa))

(g) Interpretabilidade de pronome de 3Pessoa +Referencial (+Anômalo vs. -Anômalo)

4.3.2.

Esquema de anotação morfossintática (etiquetagem)

O processo de etiquetagem morfossintática dos textos foi realizado manualmente, seguindo os parâmetros de análise estabelecidos. Apenas sentenças e expressões nominais pronominais na posição de sujeito ou objeto direto não-preposicionado foram etiquetadas.

Todas as sentenças foram classificadas como matriz (SM) ou encaixada (SE). Matrizes e encaixadas truncadas foram etiquetadas como tal (TRUNC), e classificadas como não-anômalas (TRUNC-A) ou anômalas (TRUNC+A). Os truncamentos não-anômalos são casos canônicos de elipses. Os casos anômalos são casos agramaticais de aplicação de operações de elisão. Foram consideradas sentenças apenas unidades de texto contendo um predicado verbal. Etiquetamos também sentenças coordenadas (SC), porém estas não foram contabilizadas para análises estatísticas. Por motivo de tempo de execução e apresentação da pesquisa, deixamos comparações envolvendo sentenças coordenadas para desenvolvimentos futuros.

Pronomes na posição de sujeito sentencial e de complemento verbal direto, não-preposicionado, foram etiquetados levando em consideração a forma (plenos (O(vert)) ou nulos (N)), a referencialidade (+Referencial (+R), -Referenciais (-R)) e o traço de Pessoa (1Pessoa (1P), 2Pessoa (2P) e 3Pessoa (3P)). A informação semântica de referencialidade (+Referencial (+R) e -Referencial (-R)) levou em

consideração pronomes com referência genérica e expletivos classificados como -Referenciais, comparados a pronomes com Referência específica, considerados +Referenciais. Ainda, pronomes 3P+R foram classificados como anômalos (+A) e não-anômalos (-A), onde -A são casos em que o referente/antecedente do pronome pôde ser facilmente encontrado no nível da sentença ou inferido do contexto verbal da postagem. Casos de pronomes +A são aqueles em que não foi possível definir o referente/antecedente do pronome. A etiquetagem de pronomes levou em consideração também a posição sintática do pronome (sujeito [SUJ] ou objeto [OBJ]), porém, devido ao tempo de conclusão da pesquisa e aos parâmetros estabelecidos anteriormente, não consideramos na análise estatística o fator posição sintática. Do mesmo modo, etiquetamos, mas não consideramos na presente análise, o contexto sintático sentencial de ocorrência de pronomes de 3Pessoa anômalos, se dentro do domínio de matrizes ou de encaixadas.

Em todos os casos de coordenação que poderiam ser analisados como coordenação de sintagmas verbais (SVs), optamos por não marcar o sujeito da segunda coordenada, considerando a possibilidade de uma análise de coordenação de SVs. Isso foi feito para evitar um falso positivo em relação ao número de pronomes nulos na posição de sujeito.

O processo de etiquetagem foi realizado por 2 dois pesquisadores envolvidos na pesquisa: João Victor de Oliveira Miranda e Silva, autor desta dissertação (JVS), e Cilene Rodrigues, orientadora da pesquisa (CR). O processo foi realizado em 3 fases:

Fase 1: JVS, orientado por CR, fez a primeira etiquetagem;

Fase 2: CR fez a revisão da etiquetagem realizada na fase 1;

Fase 3: JVS e CR discutiram e reanalisaram todos os casos de discordância de etiqueta. Primeiramente, etiquetamos os textos do grupo controle e, em seguida, os textos do grupo alvo.

4.3.3.

Exemplos da amostra final

Exemplo 1 - Grupo alvo

[SM 24 [M4F] Anywhere - Looking for a cute girl [^{SE} to shower with attention]] <https://imgur.com/a/HDbIO3n> About Me very monogamous currently a senior in university [SM music is my passion], [SM I^{O1PSUJ} listen to a wide range of genres, mostly older stuff] [SM pro^{N3P+R+ASUJ} loves discovering new music more than anything] interested in audio equipment like headphones [^{SC} [SM pro^{N1PSUJ} recently started [^{SE} collecting vinyl bikes]] and [SM pro^{N3P+R+ASUJSM} is looking [^{SE} to do some touring over the summer]] [SM pro^{N3P+R+ASUJSM} plays video games [SM pro^{N3P+R+ASUJSE} has PC, PS4, and Xbox One)]] About You:] [^{SC} [SMTRUNC-A cute girl [^{SE} who is interested in a monogamous romantic relationship] would start online] but [SM I^{O1PSUJ} expect us^{O1POBJ} [^{SE} to meet up eventually]]]] shorter and lighter than me [SM I^{O1PSUJ} am 6' and a healthy weight]) [SM pro^{N3P+R+ASUJSM} has an interest in music, [^{SE} pro^{N3P-R-AExpSUJSE} does not matter what genres] [^{SC} [SMTRUNC+A kind of needy/clingy] and [SM pro^{N3P+R+ASUJSM} loves attention]] talkative/chatty, [SM pro^{N3P+R+ASUJSM} likes leading conversations] [SM pro^{N3P+R+ASUJMS} wants [^{SC} [^{SE} to tell me^{O1POBJ} all about themselves, their day,] and [^{SE} whatever is on their minds]]] understanding done with high school [SM Send me^{O1POBJ} a pic of yourself [^{SC} [^{SE} if this sounds like you] and [^{SE} you^{O2PSUJ} are interested]]]. [SMTRUNC+A let us^{O1POBJ} get [^{SE} to know each other]]!

Exemplo 2 - Grupo alvo

[SM I^{O1PSUJ} have had the conversation with past girlfriends before but in the opposite direction] - [SM they^{O3P+R-ASUJ} eventually wanted children, [^{SE} where I^{O1PSUJ} knew [^{SE} since I^{O1PSUJ} was a teenager [^{SE} that I^{O1PSUJ} did not]]]]. [SM My advice, [^{SE} echoing most of this thread], would be [^{SE} to do [^{SE} what I^{O1PSUJ} did and let [^{SE} the relationship go]]]. [^{SC} [SM Give him^{O3P+R+AOBJSM} the freedom [^{SE} to find a partner [^{SE} who wants children [^{SE} as he^{O3P+R+ASUJSE} does]]]], and [SM give yourself the same freedom [^{SE} to have a happy child free life with a good man [^{SE} who wants that as well]]]]. [SM The longer you^{O2PSUJ} wait, [^{SC} [^{SE} the harder it^{O3P-R-AExpSUJ} will be]], and

[^{SE} the resentment will be tougher [^{SE} to let [^{SE} go]]]]. [SM You^{O2PSUJ} do not deserve that].

Exemplo 3 - Grupo controle

A hospital for sad people

[SM I^{O1PSUJ} have been sad lately]. [SM Today during my evening walk I^{O1PSUJ} thought of this idea – A place [^{SE} where people can go to [^{SE} when they^{O3P+R-ASUJ} are sad, stressed or lonely]]]. [SM In my mind, [^{SE} walking into this hospital] will make you^{O2P-R-AGenSUJ} [^{SE} feel like [^{SE} you^{O2P-R-AGenSUJ} are closer to nature]]]. [SM You^{O2P-R-AGenSUJMS} are welcomed with a fully loaded ice-cream cone]. [SM You^{O2P-R-AGenSUJ} can find people [^{SC} [^{SE} to talk to], engage in creating art and music with other people and find good people [^{SE} to hug and sleep next to]]]. [SM I^{O1PSUJ} know so many people [^{SE} who need it^{O3P+R-AOBJ}]]. [SM Just take a look at r/depressed or r/lonely]. [^{SC} [SM I^{O1PSUJ} do not have the money [^{SE} to build something like this]. [^{SC} [SM I^{O1PSUJ} wish [^{SETRUNC-A} I^{O1PSUJ} could]] But [^{SMTRUNC-A} [^{SE} if you^{O2PSUJ} have the money and like the idea], please do]]]!

Exemplo 4 - Grupo controle

[SM Can your girlfriend be friends with an ex?]

[SM My misses still occasionally has contact with her ex]. [^{SC} [SM He^{O3P+R-ASUJ} likes all of her photos on Instagram] and [SM she^{O3P+R-ASUJ} likes his stuff on insta and Facebook]]. [SM He^{O3P+R-ASUJ} sometimes messages her^{O3P+R-AOBJ} [^{SE} asking [^{SE} how her daughter is (not his)]]]. [^{SC} [SM She^{O3P+R-ASUJ} insisted [^{SC} [^{SE} they^{O3P+R-ASUJ} was not serious] and [^{SE} there^{O3P-R-AExpSUJ} is nothing there]]], but [SM something in my gut feels odd about this situation]]. [^{SC} [SM She^{O3P+R-ASUJ} hates [^{SE} the thought of me doing it^{O3P+R-AOBJ} with an ex of mine]], but [^{SC} [SM it^{O3P-R-AExpSUJ} seems [^{SE} to be okay for her]] and [SM pro^{N3P+R+ASUJSM} does not see the potential issues]]]. [^{SC} [SM Am I^{O1PSUJ} just being "Emotionally immature"] or [SM do I^{O1PSUJ} have a right [^{SE} to kind want [^{SE} this shit to stop]]]]...

4.3.4.

Análise estatística

Dada a heterogeneidade da nossa amostra quanto à presença de declaração de outros sofrimentos mentais (ansiedade e/ou depressão), a análise estatística foi conduzida em duas etapas, visando compreender possíveis efeitos de ansiedade e depressão nos resultados obtidos.

Etapa 1: Amostra total (i.e., amostra heterogênea): comparação estatística dos fatores linguísticos entre todos os participantes de cada grupo (39 alvos e 39 controles, totalizando 78 participantes); ou seja, incluindo a incidência de participantes com autodeclaração de ansiedade e/ou depressão;

Etapa 2: Amostra parcial (i.e., amostra homogênea): comparação estatística dos fatores linguísticos removendo da amostra participantes com autodeclaração de ansiedade e/ou depressão (17 participantes do grupo alvo e 17 do grupo controle), totalizando um corte de 34 participantes.

Dado que os textos/postagens analisadas variam em tamanho (número de palavras, de 100 a 500 palavras), todos os fatores manipulados foram quantificados como percentuais de ocorrência do item alvo pelo número total de palavras da postagem. Por exemplo, o número total de matrizes por postagem foi obtido considerando a razão percentual do total de sentenças matrizes (truncadas e não-truncadas) pelo número total de palavras. Do mesmo modo, o número total de pronomes plenos em uma postagem é representado na análise estatística pela razão percentual do número total de pronomes plenos pelo número total de palavras da mesma postagem.

A análise estatística foi conduzida no software SPSS (IBM). Dado que os dados não apresentaram uma distribuição normal, foi realizada análise não-paramétrica, testes de comparação Mann-Whitney U , com valor de P fixo no nível .05.

4.4.

Resultados

Etapa 1: Análise da amostra total (heterogênea)

No nível da sentença, comparações de Mann-Whitney U foram significativas para número total de sentenças matrizes truncadas ($p = .031$) e para número total de sentenças matrizes truncadas não-anômalas ($p = .004$) (tabela 5). O grupo alvo produziu um número significativamente menor de sentenças matrizes truncadas e de sentenças matrizes truncadas não-anômalas (figura 7):

	U de Mann-Whitney	Wilcoxon W	Z	Significância Sig. (bilateral)
MSTotal	758,000	1538,000	-0,025	0,980
MS-TRUNC	678,000	1458,000	-0,824	0,410
MS+TRUNCTotal	546,000	1326,000	-2,162	0,031
MS+TRUNC-A	489,000	1269,000	-2,845	0,004
MS+TRUNC+A	746,000	1526,000	-0,157	0,875
ESTotal	689,000	1469,000	-0,715	0,475
ES-TRUNC	715,000	1495,000	-0,455	0,649
ES+TRUNCTotal	679,500	1459,500	-0,954	0,340
ES+TRUNC-A	757,000	1537,000	-0,060	0,952
ES+TRUNC+A	622,500	1402,500	-1,766	0,077

Tabela 5. Comparações Mann-Whitney U entre grupos (amostra total) para categorias do nível da sentença. Variáveis com efeitos significativos destacadas em azul.

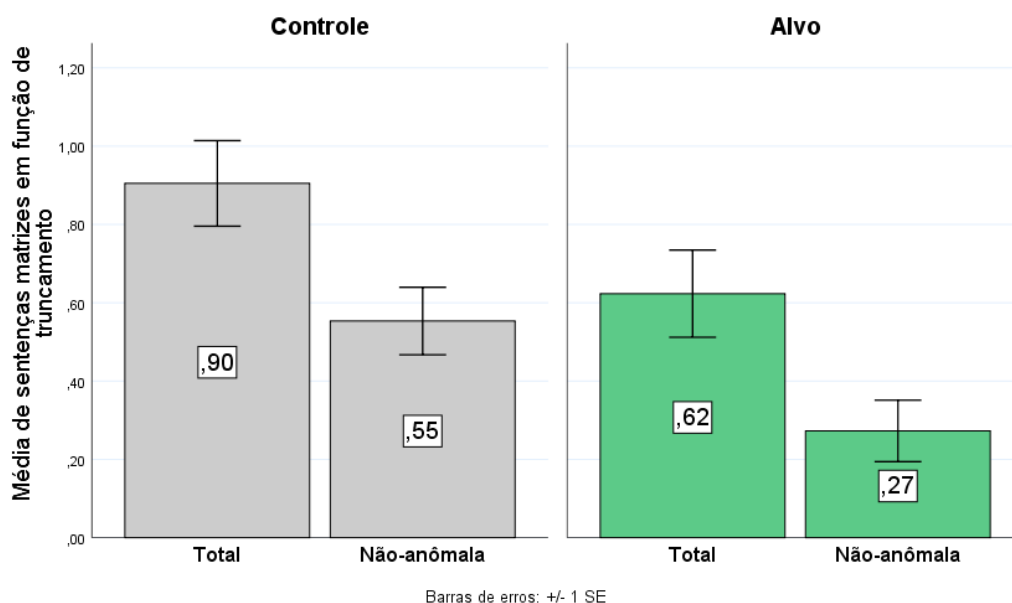


Figura 7. Comparações entre grupos da média de sentenças matrizes truncadas (MS+TRUNCTotal) e sentenças matrizes truncadas não-anômalas (MS+TRUNC-A).

No nível pronominal, comparações de Mann-Whitney U foram significativas apenas para a categoria de Pronome Pleno de 3Pessoa Referencial Anômalo (O+R-L3P+A) ($p = .038$) (tabela 6). O grupo alvo apresentou um número significativamente maior de pronomes plenos referenciais de terceira pessoa anômalos (figura 8)¹²:

	U de Mann-Whitney	Wilcoxon W	Z	Significância Sig. (bilateral)
PronTotal	628,000	1408,000	-1,324	0,185
O (Pronome pleno)	636,000	1416,000	-1,244	0,213
N (Pronome nulo)	737,500	1517,500	-0,244	0,807
O+R (Pleno +Referencial)	629,000	1409,000	-1,314	0,189
O-R (Pleno -Referencial)	720,500	1500,500	-0,400	0,689
N+R(Nulo +referencial)	708,000	1488,000	-0,561	0,575
N-R (Nulo -referencial)	702,000	1482,000	-1,755	0,079
O+R+L (Pleno +Referencial +Locutor)	733,000	1513,000	-0,275	0,783
N+R+L (Nulo +Referencial +Locutor)	664,000	1444,000	-1,082	0,279
O+R+L1P (1Pessoa Pleno +Referencial +Locutor)	713,000	1493,000	-0,475	0,635
O+R+L2P (2Pessoa Pleno +Referencial +Locutor)	723,000	1503,000	-0,437	0,662
O+R-L3P (3Pessoa Pleno +Referencial -Locutor)	639,500	1419,500	-1,210	0,226
O+R-L3P-A (3Pessoa Pleno +Referencial -Locutor -Anômalo)	579,000	1359,000	-1,815	0,070
O+R-L3P+A (3Pessoa Pleno +Referencial -Locutor +Anômalo)	572,000	1352,000	-2,080	0,038
N+R+L1P (1Pessoa Nulo +Referencial +Locutor)	664,000	1444,000	-1,082	0,279
N+R+L2P (2Pessoa Nulo +Referencial +Locutor)	760,500	1540,500	0,000	1,000
N+R-L3P (3Pessoa Nulo +Referencial -Locutor)	715,500	1495,500	-0,672	0,501

¹² Importante ressaltar que o gráfico de comparação entre grupos da média de pronomes plenos referenciais de terceira pessoa anômalos (figura 8) apresenta efeito de perda de significância apesar da análise estatística (Mann-Whitney U) ter apontado o parâmetro como significativo, resultado que atribuímos à presença de *outliers* no grupo controle. Esse efeito de presença de *outliers* foi reduzido ao retirarmos da amostra, na etapa 2 da análise estatística, usuários com autodeclaração de ansiedade e/ou depressão (figura 10).

N+R-L3P-A (3Pessoa Nulo +Referencial -Locutor -Anômalo)	737,000	1517,000	-0,388	0,698
N+R-L3P+A (3Pessoa Nulo +Referencial -Locutor +Anômalo)	741,000	1521,000	-0,585	0,559

a. Variável de Agrupamento: Grupo
b. *P* significativo no nível .05

Tabela 6. Comparações de **Mann-Whitney *U*** entre grupos (amostra total) para categorias do nível pronominal. Variável com efeito significativo destacada.

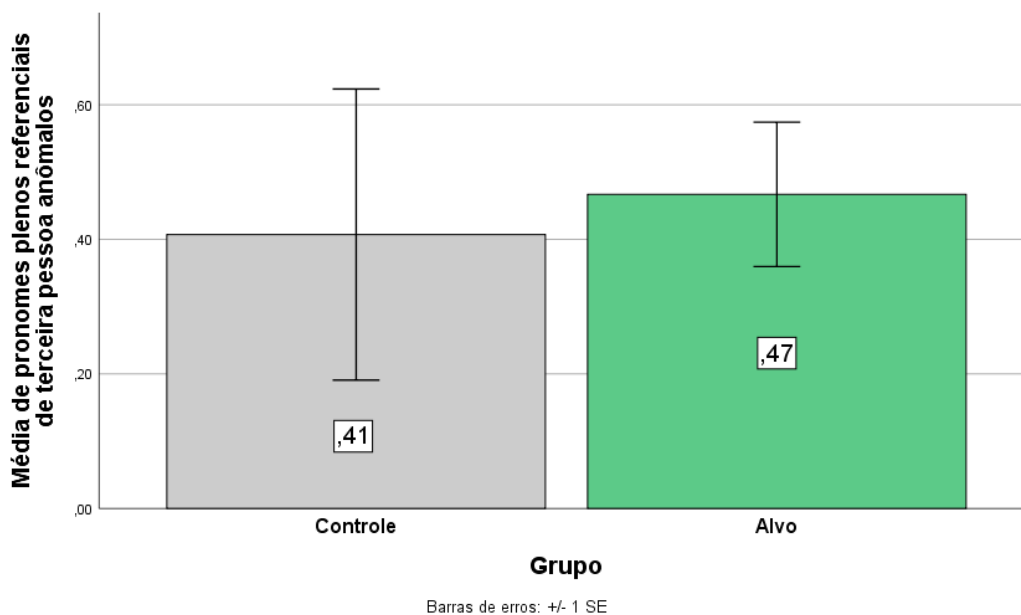


Figura 8. Comparações entre grupos da média de pronomes plenos referenciais de terceira pessoa anômalos (O+R-L3P+A).

Etapa 2: Análise da amostra parcial (homogênea)

Nesta etapa, no nível da sentença, comparações de Mann-Whitney *U* foram também significativas para número total de sentenças truncadas não-anômalas no nível de matriz ($p = .028$) (tabela 7), com produção significativamente menor de sentenças matrizes truncadas não-anômalas (figura 9) pelo grupo alvo. Esses resultados reiteram, portanto, os resultados obtidos na etapa 1.

	U de Mann-Whitney	Wilcoxon W	Z	Significância Sig. (bilateral)
MS-TRUNC	236,000	489,000	-0,141	0,888
MS-TRUNC	208,000	461,000	-0,798	0,425
MS+TRUNCTotal	161,000	414,000	-1,916	0,055
MS+TRUNC-A	152,000	405,000	-2,203	0,028

MS+TRUNC+A	220,500	473,500	-0,535	0,593
ESTotal	233,000	486,000	-0,211	0,833
ES-TRUNC	227,000	480,000	-0,352	0,725
ES+TRUNCTotal	203,500	456,500	-1,070	0,285
ES+TRUNC-A	235,000	488,000	-0,258	0,796
ES+TRUNC+A	183,500	436,500	-1,806	0,071

a. Variável de Agrupamento: Grupo
b. *P* significativo no nível .05

Tabela 7. Comparações de **Mann-Whitney *U* entre grupos (amostra parcial)** para **categorias do nível da sentença**. Variável com efeito significativo destacada.

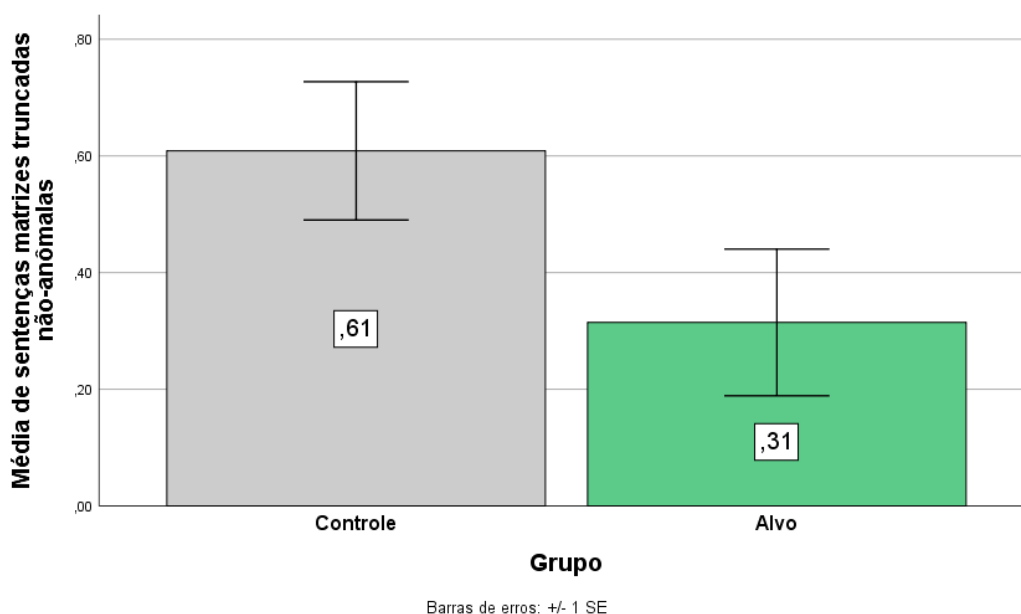


Figura 9. Comparações entre grupos da média de sentenças matrizes truncadas não-anômalas (MS+TRUNC-A).

No nível pronominal, comparações de Mann-Whitney *U* também foram significativas para a categoria de Pronome Pleno de 3Pessoa Referencial Anômalo (O+R-L3P+A) ($p = .001$) (tabela 8). O grupo alvo apresentou um número significativamente maior de pronomes plenos referenciais de terceira pessoa anômalos (figura 10), reforçando, portanto, o achado da etapa 1.

	U de Mann-Whitney	Wilcoxon W	Z	Significância Sig. (bilateral)
PronTotal	195,000	448,000	-1,103	0,270
O (Pronome pleno)	199,000	452,000	-1,009	0,313

N (Pronome nulo)	228,000	481,000	-0,343	0,732
O+R (Pleno +Referencial)	199,000	452,000	-1,009	0,313
O-R (Pleno -Referencial)	234,000	487,000	-0,188	0,851
N+R(Nulo +referencial)	216,000	469,000	-0,641	0,521
N-R (Nulo -referencial)	220,000	473,000	-1,431	0,153
O+R+L (Pleno +Referencial +Locutor)	235,000	488,000	-0,164	0,869
N+R+L (Nulo +Referencial +Locutor)	220,000	473,000	-0,571	0,568
O+R+L1P (1Pessoa Pleno +Referencial +Locutor)	224,000	477,000	-0,423	0,673
O+R+L2P (2Pessoa Pleno +Referencial +Locutor)	211,000	464,000	-0,861	0,389
O+R-L3P (3Pessoa Pleno +Referencial -Locutor)	218,000	471,000	-0,564	0,572
O+R-L3P-A (3Pessoa Pleno +Referencial -Locutor -Anômalo)	175,500	428,500	-1,566	0,117
O+R-L3P+A (3Pessoa Pleno +Referencial -Locutor +Anômalo)	116,000	369,000	-3,194	0,001
N+R+L1P (1Pessoa Nulo +Referencial +Locutor)	220,000	473,000	-0,571	0,568
N+R+L2P (2Pessoa Nulo +Referencial +Locutor)	242,000	495,000	0,000	1,000
N+R-L3P (3Pessoa Nulo +Referencial -Locutor)	233,000	486,000	-0,332	0,740
N+R-L3P-A (3Pessoa Nulo +Referencial -Locutor -Anômalo)	231,000	484,000	-0,468	0,639
N+R-L3P+A (3Pessoa Nulo +Referencial -Locutor +Anômalo)	241,500	494,500	-0,033	0,974
a. Variável de Agrupamento: Grupo				
b. <i>P</i> significativo no nível .05				

Tabela 8. Comparações de **Mann-Whitney *U*** entre grupos (amostra parcial) para categorias do nível pronominal. Variável com efeito significativo destacada.

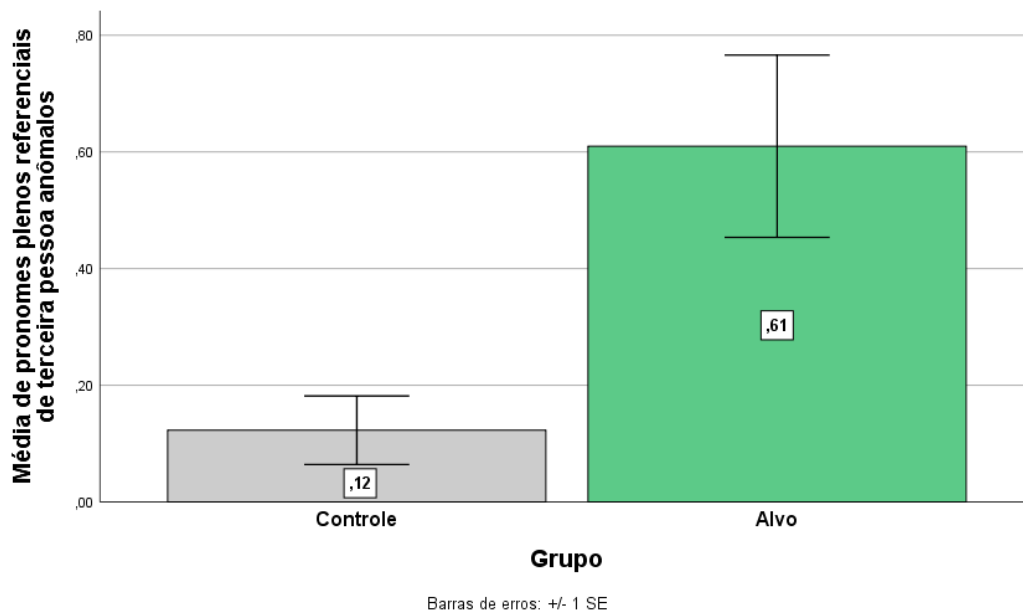


Figura 10. Comparações entre grupos da média de pronomes plenos referenciais de terceira pessoa anômalos (O+R-L3P+A).

4.5.

Discussão dos resultados

Conforme estabelecido no capítulo 1, o objetivo da presente pesquisa é analisar o uso pronominal, e adicionalmente a complexidade estrutural no nível da sentença, em textos escritos em inglês produzidos por usuários do *Reddit* com autodeclaração de esquizofrenia (grupo alvo), comparando-os a usuários sem a referida autodeclaração (grupo controle). A pesquisa partiu das seguintes hipóteses de trabalho:

- (a) A esquizofrenia leva a perdas de complexidade estrutural tanto no nível da sentença como no nível nominal.
- (b) Os pronomes são afetados tanto na forma como no significado.
- (c) Essas dificuldades podem ser observadas na escrita, já que a escrita demanda maior carga de memória de trabalho (ver capítulo 2).

Considerando estas hipóteses e os parâmetros linguísticos analisados na investigação, estabelecemos as seguintes previsões para o grupo alvo, em comparação ao controle:

(P1) Menor complexidade no nível sintático, com maior ocorrência de sentenças matrizes e maior ocorrência de sentenças truncadas anômalas

(P2) Maior ocorrência de sentenças truncadas

(P3) Maior ocorrência de truncadas anômalas

(P4) Maior ocorrência de pronomes nulos

(P5) Maior ocorrência de pronomes +Locutor (1Pessoa e 2Pessoa)

(P6) Maior ocorrência de pronomes não-referenciais

(P7) Maior ocorrência de pronomes anômalos

A previsão 1 (P1) não teve suporte nos resultados obtidos. Não houve diferença intergrupar significativa para tipo de sentença (matriz vs. subordinada). Este resultado não está, portanto, na mesma direção de achados de pesquisas anteriores (Özcan et al., 2017; Chaves, 2022) que relatam maior volume de sentenças matrizes em pacientes com esquizofrenia. Esta diferença pode refletir diferenças em tarefas cognitivas referentes à produção de texto escrito e de texto oral, já que as amostras de Özcan et al. (2017) e Chaves (2022) foram compostas por narrativas orais. Embora a produção de texto escrito exija uma carga de memória de trabalho expressiva (ver seção 2.5.), a publicação de textos escritos em plataformas sociais possibilita ao usuário/escritor maior tempo de elaboração e revisões. Assim, é possível que nossos participantes, de modo geral, podem ter tido vantagem neste processo, podendo apresentar textos mais complexos do ponto de vista sintático.

As previsões 2 e 3 (P2 e P3) foram parcialmente conferidas. Em oposição à P2, as nossas observações indicam maior produção de sentenças truncadas no grupo controle, mas menor volume de sentenças truncadas não-anômalas no grupo alvo, resultado consistente com P3. Os dados em (27a) e (27b), trechos de postagens do grupo controle e grupo alvo, respectivamente, ilustram o uso de sentenças truncadas não-anômalas (27a) e anômalas (27b), destacadas em negrito:

(27) a. what is the best type of savings account I can get to save for a car? More info inside. I am a 21y/o living with my parents in Maine, no longer a college student, with student loans I need to start paying soon. **Making \$1200/month currently at a full-time job. Able to save ~\$800/month.**

b. I need some advice for stomach tattoo. Hey I need some advice on stomach tattoo pain. **I did my stomach a few days** and well obviously it hurt like a bitch.

Acreditamos que, de modo geral, esses achados indicam falta de habilidade dos esquizofrênicos em lidar com o processo de elisão. Em comparação com o grupo controle, o grupo alvo realizou menos elisões e as elisões realizadas apresentaram significativamente mais anomalias. A elisão é um processo gramatical que envolve informações sintáticas (tipo de subestrutura a ser elidida), informações semânticas e pragmáticas (recuperação do antecedente do constituinte elidido) e informações fonológicas (apagamento de constituinte) (Merchant, 2001). Portanto, nossos resultados sugerem dificuldades no nível gramatical e vão ao encontro de pesquisas anteriores como DeLisi (2001), que relata maior número de erros sintáticos em esquizofrenia, Moro et al. (2015), que também relatam maior dificuldade em operações sintáticas, como formação de perguntas -qu e inversão sintática, e Chaves (2022), que apresenta resultados semelhantes aos nossos, indicando maior habilidade do grupo controle em processos de truncamento. De modo geral, portanto, acreditamos que essas dificuldades indicam menor complexidade sintática no grupo alvo.

Em relação ao uso pronominal, nossas previsões também obtiveram sustentação parcial. Primeiramente, não se observa maior número de pronomes nulos (P4). Isto é, não é possível afirmar que a produção de textos escritos por esquizofrênicos nativos do inglês apresenta maior ocorrência de pronomes nulos. Esses resultados diferem, portanto, de observações registradas na literatura sobre o uso de pronomes nulos na esquizofrenia (ver seção 3.3.3.). Tanto Tovar et al. (2019b) como Chaves (2022) registram maior uso de pronomes nulos em narrativas produzidas por esquizofrênicos nativos de espanhol e português brasileiro (línguas de sujeito nulo). Essa discrepância é provavelmente devido ao fato de o inglês não ser uma língua de sujeito nulo, usando pronomes nulos na posição de sujeito apenas em contextos sintáticos específicos, conforme discutido no capítulo 3. Ainda, pode ser que o fato de textos escritos envolverem um registro formal leve à queda na produção de sujeitos nulos no inglês. De qualquer modo, nossos achados também não nos possibilitam concluir que não há diferenças entre esquizofrênicos e controles, no inglês, em relação à produção de

pronomes fortes e fracos. Conforme salientamos no capítulo 3, os pronomes fracos do inglês ocorrem especialmente em posição de objeto e estão restritos ao registro oral. Portanto, os nossos dados não nos possibilitam verificar a ocorrência desses pronomes.

Não obtivemos suporte para P5, já que não houve, no grupo alvo, em comparação com o grupo controle, maior ocorrência de pronomes de primeira e segunda pessoa em oposição a pronomes de 3Pessoa. Salientamos que esses resultados não estão de acordo com observações obtidas por outros pesquisadores que, com base em textos retirados de plataformas sociais, incluindo o *Reddit*, indicam maior uso de pronomes de 1Pessoa (Birnbaum et al., 2017; Bae et al., 2021). No momento, não temos uma boa hipótese para essa discrepância com resultados de pesquisas anteriores, mas acreditamos que a frequência e a distribuição dos traços de pessoa dependem do tópico/tema das postagens (ver Chaves (2022) para uma discussão sobre esta questão).

Também não obtivemos suporte para P6. Não houve diferença intergrupar significativa para o fator referencialidade pronominal. O grupo alvo não produziu maior número de pronomes não referenciais. Entretanto, observamos que, de modo geral, o pronome *it* apresentou, no grupo alvo, um grande volume de anomalias. Considere, como exemplo, o trecho abaixo, retirado de uma das postagens do grupo alvo. Essa questão fica em aberto, para futuras investigações.

(28) I would definitely want one of those 10 speed transmissions from the newer truck, this build would be mainly for highway hauling and maybe some camping, and I do not think the old transmission would hold up well with a coyote. An issue I do see is **it** possibly not working with the older transfer case, that makes me wonder how many parts the donor truck is going to be giving.

A previsão 7 (P7) está plenamente de acordo com as observações obtidas. O grupo alvo produziu um número significativamente maior de anomalias referenciais de pronomes plenos de 3Pessoa. A diferença intergrupar se tornou ainda mais forte na segunda etapa da análise estatística, quando uma amostra homogênea foi considerada. Esse resultado vai ao encontro de resultados anteriores. Docherty & Gottesman (2000), Docherty et al. (1997), Rochester &

Martin (1979) e Harvey (1983) apontam que referências vagas, confusas ou ambíguas são comuns em esquizofrenia. Chaves et al. (2020) e Chaves (2022) também indicam maior ocorrência de pronomes nulos anômalos na esquizofrenia. Igualmente, Sevilla et al. (2018) relatam anomalias na referência pronominal (ver ainda Fineberg et al., 2015; Zomick et al., 2019; Bae et al., 2021; e Tovar et al., 2019b para o uso de pronomes de 3Pessoa na esquizofrenia). Como exemplo de anomalia de pronome referencial de 3Pessoa, encontrada em nossa amostra, considere o trecho a seguir (grupo alvo):

(29) I am not sure if I am going insane or not (kind of long)

So first things first (f20) I have schizophrenia which means I have got more things to work on in the cognitive and communication side of things than most I get that. It also means I might be completely in the wrong about what I am seeing or feeling and I would love to have a couple extra opinions on whether this is abuse towards me, **him**, or if were just bad to each other.

Em (29), o pronome em negrito *him* não possui nenhum antecedente prévio para estabelecer referência, caracterizando-se como um uso anômalo do pronome.

Nossos resultados indicam, de uma forma geral, que a esquizofrenia apresenta problemas na aplicação de operações gramaticais (truncamento) e problemas na construção de referência de pronomes de 3Pessoa. Acreditamos que essas observações alinham-se à hipótese de Hinzen & Rosselló (2015), Hinzen (2017) e Çokal et. al. (2018) de que a esquizofrenia leva a uma redução da estrutura sintática, e especialmente a dificuldades na construção da estrutura de sintagmas nominais (DPs) definidos.

Observamos ainda que nossos achados indicam também que a ansiedade e a depressão não possuem efeitos linguísticos significativos na esquizofrenia.

De modo geral, como vimos no capítulo 2, a literatura aponta que a esquizofrenia afeta a cognição de maneira ampla, prejudicando recursos cognitivos como memória de trabalho e controle executivo, além de habilidades sociais como teoria da mente. Portanto, pesquisas como a que realizamos podem, em um primeiro momento, indicar tanto um problema de competência (conhecimento gramatical) como de desempenho (acesso e uso de conhecimento

gramatical), caracterizando-se como uma redução nos recursos cognitivos necessários para uso da gramática. Pesquisas como a de Moro et al (2015) indicam tratar-se, no entanto, de um déficit de competência, já que as dificuldades com complexidade sintática observadas não se correlacionam com redução de memória de trabalho, geralmente associada à execução de tarefas complexas. Em nossa pesquisa, consideramos a questão como sendo de natureza gramatical, sem nos aprofundar na distinção competência vs. desempenho. Embora esta distinção seja importante para a delimitação de questões específicas da linguagem humana, na presente discussão, ela não se mostra tão produtiva já que, independente da sua origem, os déficits observados são de natureza gramatical, cabendo, portanto, uma caracterização com base no funcionamento interno da gramática.

5

Conclusão

O objetivo primário desta dissertação foi verificar a complexidade sintática na produção escrita de esquizofrênicos falantes da língua inglesa, centrando-se no uso de pronomes, considerando também a complexidade no nível sentencial. Os dados analisados foram coletados na plataforma social *Reddit*.

Contribuímos para avanços nos estudos sobre linguagem na esquizofrenia com um estudo experimental acerca do uso de pronomes e sobre a complexidade estrutural no nível sentencial, comparando postagens escritas realizadas na plataforma social *Reddit* realizadas por usuários com autodeclaração de diagnóstico de esquizofrenia (grupo alvo) com postagens realizadas na mesma plataforma por usuários sem a referida autodeclaração.

A amostra final, coletada por meio de técnicas da Ciência da Computação, contém 78 textos (39 do grupo alvo e 39 do grupo controle). Essa amostra é heterogênea em relação à presença de sofrimentos mentais adicionais (ansiedade e depressão). Os textos selecionados foram etiquetados manualmente considerando tipos de sentença e de pronomes. Os parâmetros analisados no nível sentencial foram: (a) tipo de sentença em função da complexidade sintática (Matriz vs. Subordinada); (b) tipo de sentença em função de truncamento (elisão) (+Truncada, vs. -Truncada); e (c) tipo de truncamento sentencial (gramatical (-Anômalo) vs. agramatical (+Anômalo)). No nível pronominal, os parâmetros analisados foram: (d) tipo de pronome em função da forma fonológica (Pleno vs. Nulo); (e) referencialidade pronominal (+Referencial vs. -Referencial); (f) tipo de pronome em função dos traços de pessoa (+Locutor (1Pessoa, 2Pessoa), vs. -Locutor (3Pessoa)); (g) interpretabilidade de pronome de 3Pessoa +Referencial (+Anômalo vs. -Anômalo).

Os resultados das análises estatísticas realizadas indicam diferenças significativas entre os grupos para os parâmetros referentes à anomalia: parâmetros (c) - tipo de truncamento sentencial (gramatical (-Anômalo) vs. agramatical (+Anômalo)), e (g) - interpretabilidade de pronome de 3Pessoa +Referencial (+Anômalo vs. -Anômalo), com o grupo alvo apresentando menor número total de sentenças matrizes truncadas não-anômalas e maior número de pronomes referenciais de 3Pessoa anômalos. Estes resultados estão de acordo com achados

já relatados na literatura. Acreditamos que os mesmos reforçam a hipótese de que esquizofrenia leva à redução da complexidade estrutural, particularmente em processos derivacionais de elisão e construção da referência pronominal de 3Pessoa, alinhando-se, portanto, com as hipóteses de Hinzen & Rosselló (2015), Hinzen (2017) e Çokal et. al. (2018) de que a esquizofrenia leva a uma redução da estrutura sintática, e especialmente a dificuldades na construção da estrutura de sintagmas nominais (DPs) definidos. Apesar da literatura apontar que esquizofrênicos falantes de línguas *pro-drop* têm preferência por pronomes nulos (formas pronominais fracas) (Tovar et al., 2019b; Chaves, 2022), nossos resultados não evidenciam essa preferência, observação que interpretamos como possivelmente associada à nossa língua de análise, o inglês, por não ser uma língua *pro-drop*. Nossos achados indicam ainda que ansiedade e depressão não afetam (i.e., reduzem) a complexidade sintática.

Acreditamos que nossa pesquisa contribui para o entendimento da esquizofrenia (e seus sintomas associados) e da linguagem como um mecanismo cognitivo por contribuir para uma caracterização mais refinada de condições específicas associadas à linguagem, em particular aquelas relativas aos processos gramaticais de elisão no nível sentencial e da construção referencial de pronomes fortes de 3Pessoa. Esses achados também contribuem para um maior entendimento da Gramática humana como um todo, apontando para a possibilidade de alguns processos cognitivos derivacionais (elisão e referência de pronomes fortes) demandarem uma carga mais alta de recursos cognitivos envolvidos no uso adequado da linguagem, como a memória de trabalho. Além disso, nossa pesquisa contribuiu para a criação de um banco de dados de textos escritos, retirados da plataforma *Reddit*, que podem ser futuramente utilizados em novas pesquisas acerca da linguagem no contexto de transtornos mentais. Contribuiu também para o refinamento de métodos de etiquetagem manual, que podem ser utilizados para técnicas futuras de aprendizagem de máquina, e a comparação entre uso de formas plenas e nulas no inglês.

Enfatizamos ainda que pesquisas interdisciplinares são relevantes no contexto de estudos da linguagem na esquizofrenia. A nossa investigação obteve vantagens no emprego de métodos da Ciência da Computação para coleta de dados, que nos possibilitaram trabalhar com um número maior de dados, selecionando os dados da maneira mais cuidadosa possível.

Entretanto, algumas considerações são relevantes em relação à pesquisa conduzida. Primeiro, a amostra foi composta com base na presença/ausência de autodeclaração de diagnóstico para esquizofrenia. Assim, não é possível, a partir do tipo de dado que trabalhamos, garantir a veracidade do diagnóstico nem considerar questões relevantes para o estudo da linguagem na esquizofrenia, como desordem do pensamento formal (FTD). Salientamos, no entanto, que nossos achados foram na mesma direção de achados descritos na literatura, resultantes de pesquisas com controle de variáveis externas, incluindo diagnóstico mais preciso.

Outra limitação da nossa pesquisa é a ausência de controle do período de publicação das postagens analisadas. A amostra foi composta por postagens realizadas antes e depois da pandemia (COVID-19). Estudos recentes apontam, no entanto, para possíveis efeitos da pandemia na saúde mental da população no geral (Mota et al., 2020). Assim, observamos que uma análise por período (pré-pandemia e pós-pandemia) pode ser relevante para investigações sobre o uso da linguagem.

É importante apontar, também, que nossa amostra é pequena, compreendendo uma quantidade reduzida de postagens após as filtragens realizadas durante a coleta. Como consequência, nossos dados podem estar suscetíveis ao risco beta (erro tipo II).

Para melhor entendimento da gramática na esquizofrenia, investigações futuras devem ser realizadas levando em consideração diferenças translinguísticas e diferenças intralinguísticas. Por exemplo, estudos focados em produção oral no inglês poderão considerar a diferença entre formas pronominais fortes e fracas, possibilitando assim uma análise mais ampla e mais detalhada sobre os efeitos da variação linguística na esquizofrenia. Mostram-se igualmente importantes comparações entre coordenação e subordinação sentencial na esquizofrenia, já que esses dois processos apresentam diferenças entre si em relação à complexidade derivacional.

Referências bibliográficas

ABNEY, S. P. **The English noun phrase in its sentential aspect**. Tese de Doutorado. Massachusetts Institute of Technology, 1987.

ADGER, D. **Core syntax: A minimalist approach**. Oxford: Oxford University Press, 2003.

ALMOR, A. Why do Alzheimer patients have difficulty with pronouns? Working memory, semantics, and reference in comprehension and production in Alzheimer's disease. **Brain and language**, v. 67, n. 3, p. 202-227, 1999.

American Psychiatric Association (APA). **Diagnostic and Statistical Manual of Mental Disorders**. 5th ed. Washington, DC: The Association, 2013.

ANDREASEN, N. C. Thought, language, and communication disorders: I. Clinical assessment, definition of terms, and evaluation of their reliability. **Archives of general Psychiatry**, v. 36, n. 12, p. 1315-1321, 1979a.

ANDREASEN, N. C. Thought, language, and communication disorders: II. Diagnostic significance. **Archives of general Psychiatry**, v. 36, n. 12, p. 1325-1330, 1979b.

ANDREASEN, N. C. Scale for the assessment of thought, language, and communication (TLC). **Schizophrenia bulletin**, v. 12, n. 3, p. 473, 1986.

AOUN, J. Resumption and last resort. **DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada**, v. 16, n. spe, p. 13-43, 2000.

ARIEL, M. Accessing noun-phrase antecedents. Routledge, Londres. **Linguistics, COLING**, v. 96, p. 113-118, 1990.

ASUDEH, A. **Resumption as resource management**. Stanford University dissertation. Stanford, 2004.

AVRUTIN, S.; WEXLER, K. Development of Principle B in Russian: Coindexation at LF and coreference. **Language acquisition**, v. 2, n. 4, p. 259-306, 1992.

BAAUW, S. **Grammatical features and the acquisition of reference: A comparative study of Dutch and Spanish**. Routledge, 2013.

BAE, Y. J.; SHIM, M.; LEE, W. H. Schizophrenia Detection Using Machine Learning Approach from Social Media Content. **Sensors**, v. 21, n. 17, p. 5924, 2021.

BAGNER, D. M. Language comprehension and working memory language comprehension and working memory deficits in patients with schizophrenia. **Schizophrenia Research**, v. 60, n. 2-3, p. 299-309, 2003.

BANNEY, R. M.; HARPER-HILL, K.; ARNOTT, W. L. The Autism Diagnostic Observation Schedule and narrative assessment: Evidence for specific narrative impairments in autism spectrum disorders. **International Journal of Speech-Language Pathology**, v. 17, n. 2, p. 159-171, 2015.

BARCH, D. M.; BERENBAUM, H. Language production and thought disorder in schizophrenia. **Journal of Abnormal Psychology**, v. 105, n. 1, p. 81, 1996.

BECKER, T. M. et al. Cognitive control components and speech symptoms in people with schizophrenia. **Psychiatry Research**, v. 196, n. 1, p. 20-26, 2012.

BELLETTI, A.; BENNATI, E.; SORACE, A.. Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. **Natural language & linguistic theory**, v. 25, n. 4, p. 657, 2007.

BERNSTEIN, J. B. The DP Hypothesis: Identifying Clausal Properties in the Nominal Domain. **The handbook of contemporary syntactic theory**, p. 536, 1991.

BERNSTEIN, J. B. Reformulating the determiner phrase analysis. **Language and Linguistics Compass**, v. 2, n. 6, p. 1246-1270, 2008.

BERWICK, R. C. et al. Evolution, brain, and the nature of language. **Trends in cognitive sciences**, v. 17, n. 2, p. 89-98, 2013.

BERWICK, R. C.; CHOMSKY, N. **Why only us: Language and evolution**. MIT press, 2016.

BILDER, R. M. et al. Neuropsychology of first-episode schizophrenia: initial characterization and clinical correlates. **American Journal of Psychiatry**, v. 157, n. 4, p. 549-559, 2000.

BIRNBAUM, M. L. et al. A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals. **Journal of medical Internet research**, v. 19, n. 8, p. e7956, 2017.

BLEULER, E. **Dementia praecox or the group of schizophrenias**. 1911.

BOLHUIS, J. J. et al. How could language have evolved?. **PLoS Biol**, v. 12, n. 8, p. e1001934, 2014.

BOWIE, C. R.; HARVEY, P. D. Cognitive deficits and functional outcome in schizophrenia. **Neuropsychiatric disease and treatment**, v. 2, n. 4, p. 531, 2006.

BRUENING, B. Selectional asymmetries between CP and DP suggest that the DP hypothesis is wrong. **University of Pennsylvania working papers in linguistics**, v. 15, n. 1, p. 5, 2009.

BRUENING, B. Binding, coreference, and presuppositions: A presuppositional Precede-and-Command binding theory. **Unpublished draft. University of Delaware**, 2018.

BRUNETTO, V. The Pronoun Interpretation Problem in romance complex predicates. **Lingua**, v. 161, p. 82-100, 2015.

BURZIO, L. Weak anaphora. In: CINQUE, G. et al. **Paths towards Universal Grammar. Studies in honor of Richard S. Kayne**. Washington, DC: Georgetown University Press. p. 59-84, 1994.

BYRNE, M. E.; CROWE, T. A.; GRIFFIN, P. S. Pragmatic language behaviors of adults diagnosed with chronic schizophrenia. **Psychological reports**, v. 83, n. 3, p. 835-846, 1998.

CARDINALETTI, A.; STARKE, M. The typology of structural deficiency: On the three grammatical classes. **Working Papers in Linguistics**, 4.2, p. 41-109, 1994.

CARLSON, G. N. **Reference to kinds in English**. University of Massachusetts Amherst, 1977.

CENSITS, D. M. et al. Neuropsychological evidence supporting a neurodevelopmental model of schizophrenia: a longitudinal study. **Schizophrenia research**, v. 24, n. 3, p. 289-298, 1997.

CHAIKA, E.; LAMBE, R. A. Cohesion in schizophrenic narratives, revisited. **Journal of communication disorders**, v. 22, n. 6, p. 407-421, 1989.

CHAIKA, E. O. **Understanding psychotic speech: Beyond Freud and Chomsky**. Charles C Thomas, Publisher, 1990.

CHAMPAGNE-LAVAU, M.; STIP, E. Pragmatic and executive dysfunction in schizophrenia. **Journal of Neurolinguistics**, v. 23, n. 3, p. 285-296, 2010.

CHAVES, M. F. F.; RODRIGUES, C. The impact of schizotypy on pragmatics: analyzing if context-information is integrated into the semantic meaning of definite singular DPs. **Cadernos de Estudos Linguísticos**, v. 62, p. e020014-e020014, 2020.

CHAVES, M. et al. M190. USE OF NULL PRONOUNS IN SCHIZOPHRENIA. **Schizophrenia Bulletin**, v. 46, n. Suppl 1, p. S208, 2020.

CHAVES, M. F. **Failures in reference-set computations: pronouns in schizophrenia**. Tese de doutorado. PUC-Rio, 2022.

CHIEN, Y.; WEXLER, K. Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. **Language acquisition**, v. 1, n. 3, p. 225-295, 1990.

CHIERCHIA, G.; MCCONNELL-GINET, S. **Meaning and grammar: an introduction to semantics**. 1990.

CHOMSKY, N. **Aspects of the Theory of Syntax**. Cambridge, MA: MIT Press, 1965.

CHOMSKY, N. Remarks on nominalization. In: JACOBS, R.; ROSENBAUM, P. **Readings in English Transformational Grammar**. Ginn, Waltham. Mass. 1970.

CHOMSKY, N. **Lectures on government and binding**. Dordrecht: Foris Publications, 1981.

CHOMSKY, N. **Knowledge of language: Its nature, origin, and use**. Greenwood Publishing Group, 1986.

CHOMSKY, N. **The Minimalist Program**. Cambridge, MA: MIT, 1995.

CHOMSKY, N. **Novos horizontes no estudo da linguagem e da mente**. Unesp, 2005.

CHOMSKY, N. The language capacity: architecture and evolution. **Psychonomic bulletin & review**, v. 24, n. 1, p. 200-203, 2017.

CINQUE, G. On the evidence for partial N-movement in the Romance DP. In: CINQUE, G. et al. **Paths towards universal grammar. Studies in honor of Richard S. Kayne**. Georgetown University Press, 1994.

CLEMMER, E. J. Psycholinguistic aspects of pauses and temporal patterns in schizophrenic speech. **Journal of Psycholinguistic Research**, v. 9, n. 2, p. 161-185, 1980.

COHEN, A. S.; MITCHELL, K. R.; ELVEVÅG, B. What do we really know about blunted vocal affect and alogia? A meta-analysis of objective assessments. **Schizophrenia research**, v. 159, n. 2-3, p. 533-538, 2014.

COSTA, J.; AMBULATE, J. The acquisition of embedded subject pronouns in European Portuguese. In: **Proceedings of the 2009 mind/context divide workshop**. Somerville, MA: Cascadilla Press, p. 1-12, 2010.

ÇOKAL, D. et al. The language profile of formal thought disorder. **npj Schizophr** 4, 18, 2018.

CONDRAV, R.; STEINHAEUER, S. R.; GOLDSTEIN, G. Language comprehension in schizophrenics and their brothers. **Biological Psychiatry**, v. 32, n. 9, p. 790-802, 1992.

CONDRAV, R. et al. Working memory capacity predicts language comprehension in schizophrenic patients. **Schizophrenia Research**, v. 20, n. 1-2, p. 1-13, 1996.

CONDRAV, R. et al. The language system in schizophrenia: effects of capacity and linguistic structure. **Schizophrenia Bulletin**, v. 28, n. 3, p. 475, 2002.

COVINGTON, M. A. et al. Schizophrenia and the structure of language: the linguist's view. **Schizophrenia research**, v. 77, n. 1, p. 85-98, 2005.

DANON, G. **Syntactic definiteness in the grammar of Modern Hebrew**. 2001.

DANON, G. The definiteness feature at the syntax-semantics interface. **Features Perspectives on a Key Notion in Linguistics**, v. 8, n. 40, p. 144-165, 2010.

DELISI, L. E. et al. Anomalous cerebral asymmetry and language processing in schizophrenia. **Schizophrenia bulletin**, v. 23, n. 2, p. 255-271, 1997.

DELISI, L. E. Speech disorder in schizophrenia: review of the literature and exploration of its relation to the uniquely human capacity for language. **Schizophrenia bulletin**, v. 27, n. 3, p. 481-496, 2001.

DE VITA, F. et al. The Role of Working Memory on Writing Processes. **Frontiers in Psychology**, v. 12, 2021.

DICKINSON, D. et al. General and specific cognitive deficits in schizophrenia. **Biological psychiatry**, v. 55, n. 8, p. 826-833, 2004.

DITMAN, T.; KUPERBERG, G. R. Building coherence: A framework for exploring the breakdown of links across clause boundaries in schizophrenia. **Journal of neurolinguistics**, v. 23, n. 3, p. 254-269, 2010.

DOCHERTY, N. M. et al. Working memory, attention, and communication disturbances in schizophrenia. **Journal of Abnormal Psychology**, v. 105, n. 2, p. 212, 1996.

DOCHERTY, N. M.; MILLER, T. N.; LEWIS, M. A. Communication disturbances in the natural speech of schizophrenic patients and

non-schizophrenic parents of Patients. **Acta Psychiatrica Scandinavica**, v. 95, n. 6, p. 500-507, 1997.

DOCHERTY, N. M.; GORDINIER, S. W. Immediate memory, attention and communication disturbances in schizophrenia patients and their relatives. **Psychological Medicine**, v. 29, n. 1, p. 189-197, 1999.

DOCHERTY, N. M.; GOTTESMAN, I. I. A twin study of communication disturbances in schizophrenia. **The Journal of nervous and mental disease**, v. 188, n. 7, p. 395-401, 2000.

DOCHERTY, N. M. et al. Stability of formal thought disorder and referential communication disturbances in schizophrenia. **Journal of Abnormal Psychology**, v. 112, n. 3, p. 469, 2003.

DOCHERTY, N. M. On identifying the processes underlying schizophrenic speech disorder. **Schizophrenia Bulletin**, v. 38, n. 6, p. 1327-1335, 2012.

DRUMMOND, C. et al. Deficits in narrative discourse elicited by visual stimuli are already present in patients with mild cognitive impairment. **Frontiers in Aging Neuroscience**, v. 7, p. 96, 2015.

FETZER, A. **Textual coherence as a pragmatic phenomenon**. 2012.

FIENGO, R.; HIGGINBOTHAM, J. Opacity in NP. **Linguistic Analysis**, 7, 395-421, 1981.

FINEBERG, S. K. et al. Word use in first-person accounts of schizophrenia. **The British Journal of Psychiatry**, v. 206, n. 1, p. 32-38, 2015.

FINLAY, S. C. **Age and gender in Reddit commenting and success**. 2014.

FIORAVANTI, M. et al. A meta-analysis of cognitive deficits in adults with a diagnosis of schizophrenia. **Neuropsychology review**, v. 15, n. 2, p. 73-95, 2005.

FOLSTEIN, M. F.; FOLSTEIN, S. E.; MCHUGH, P. R. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. **Journal of psychiatric research**, v. 12, n. 3, p. 189-198, 1975.

FORBES, N. F. et al. Working memory in schizophrenia: a meta-analysis. **Psychological medicine**, v. 39, n. 6, p. 889-905, 2009.

FRASER, W. I. et al. The diagnosis of schizophrenia by language analysis. **The British Journal of Psychiatry**, v. 148, n. 3, p. 275-278, 1986.

FRITH, C. D. **The cognitive neuropsychology of schizophrenia**. Psychology press, 1992.

GERNSBACHER, M. A.; TALLENT, K. A.; BOLLIGER, C. M. Disordered discourse in schizophrenia described by the structure building framework. **Discourse Studies**, v. 1, n. 3, p. 355-372, 1999.

GOLD, S. et al. Longitudinal study of cognitive function in first-episode and recent-onset schizophrenia. **American Journal of Psychiatry**, v. 156, n. 9, p. 1342-1348, 1999.

GOLDBERG, T. E. et al. Neuropsychological assessment of monozygotic twins discordant for schizophrenia. **Archives of general psychiatry**, v. 47, n. 11, p. 1066-1072, 1990.

GOLDBERG, T. E. et al. Genetic risk of neuropsychological impairment in schizophrenia: a study of monozygotic twins discordant and concordant for the disorder. **Schizophrenia research**, v. 17, n. 1, p. 77-84, 1995.

GOLDBERG, T. E.; GREEN, M. Neurocognitive functioning in patients with schizophrenia: an overview. **Psychopharmacology: The fourth generation of progress**, p. 1245-1257, 1995.

GORDON, P. C.; GROSZ, B. J.; GILLIOM, L. A. Pronouns, names, and the centering of attention in discourse. **Cognitive science**, v. 17, n. 3, p. 311-347, 1993.

GORDON, P. C.; HENDRICK, R.; JOHNSON, M. Effects of noun phrase type on sentence complexity. **Journal of memory and Language**, v. 51, n. 1, p. 97-114, 2004.

GRODZINSKY, Y.; REINHART, T. The innateness of binding and coreference. **Linguistic inquiry**, v. 24, n. 1, p. 69-101, 1993.

GROLLA, E. Resumptive pronouns as last resort: Implications for language acquisition. **University of Pennsylvania Working Papers in Linguistics**, v. 11, n. 1, p. 7, 2005.

HABTEWOLD, T. D. et al. Schizophrenia symptoms are inherently heterogeneous: a systematic review of cluster and group-based studies. **bioRxiv**, p. 599498, 2019.

HAEGEMAN, L. Register-based subject omission in English and its implication for the syntax of adjuncts. **Anglophonia. French Journal of English Linguistics**, n. 28, 2019.

HAEGEMAN, L.; STARK, E. Register-specific subject omission in English and French and the syntax of coordination. In: **Continuity and Variation in Germanic and Romance**. Oxford University Press, p. 15-43, 2021.

HARVEY, P. D. Speech competence in manic and schizophrenic psychoses: The association between clinically rated thought disorder and

cohesion and reference performance. **Journal of Abnormal Psychology**, v. 92, n. 3, p. 368, 1983.

HARTMAN, J.; SUDO, Y.; WEXLER, K. Principle B and phonologically reduced pronouns in child English. **5th Generative Approaches to Language Acquisition, North America [GALANA 5]**, October, p. 11-13, 2012.

HAUSER, M. D.; CHOMSKY, N.; FITCH, W. Tecumseh. The faculty of language: what is it, who has it, and how did it evolve?. **science**, v. 298, n. 5598, p. 1569-1579, 2002.

HYAMS, N. **Language acquisition and the theory of parameters**. Dordrecht: Reidel, 1986.

HYAMS, N. A Reanalysis of Null Subjects in Child Language. **Theoretical issues in language acquisition: Continuity and change in development**, p. 249, 1991.

HEATON, R. K. et al. Stability and course of neuropsychological deficits in schizophrenia. **Archives of general psychiatry**, v. 58, n. 1, p. 24-32, 2001.

HEIM, I. R. **The semantics of definite and indefinite noun phrases**. University of Massachusetts Amherst, 1982.

HEIM, I. R.; KRATZER, A. **Semantics in generative grammar**. Oxford: Blackwell, 1998.

HESTVIK, A.; PHILIP, W. Binding and coreference in Norwegian child language. **Language Acquisition**, v. 8, n. 3, p. 171-235, 2000.

HINZEN, W.; ROSSELLÓ, J. The linguistics of schizophrenia: Thought disturbance as language pathology across symptoms. **Frontiers in Psychology**, 2015, vol. 6, num. 971, 2015.

HINZEN, W. Reference across pathologies: a new linguistic lens on disorders of thought. **Theoretical Linguistics**, v. 43, n. 3-4, p. 169-232, 2017.

HOBSON, R. P.; GARCÍA-PÉREZ, R. M.; LEE, A. Person-centred (deictic) expressions and autism. **Journal of autism and developmental disorders**, v. 40, n. 4, p. 403-415, 2010.

HOFFMAN, R. E. et al. Message disruptions during syntactic processing in schizophrenia. **Journal of Communication Disorders**, v. 18, n. 3, p. 183-202, 1985.

HORNSTEIN, N. **Move**. Oxford: Blackwell, 2001.

HORNSTEIN, N. Pronouns in a minimalist setting. **The Copy Theory of Movement**, John Benjamins, p. 351-385, 2007.

HUNT, K. W. Syntactic maturity in schoolchildren and adults. **Monographs of the society for research in child development**, v. 35, n. 1, p. iii-67, 1970.

JOHNSON, Kyle. Pronouns vs. definite descriptions. **Generative linguistics and acquisition: Studies in honor of Nina M. Hyams**, v. 54, p. 157, 2013.

KELLOGG, R. T.; OLIVE, T.; PIOLAT, A. Verbal, visual, and spatial working memory in written language production. **Acta Psychologica**, v. 124, n. 3, p. 382-397, 2007.

KEMPLER, D. Language changes in dementia of the Alzheimer type. **Dementia and communication**, p. 98-114, 1995.

KENEDY, E. Gerativismo. In: MARTELOTTA, M. E. **Manual de lingüística**. São Paulo: Contexto, v. 1, p. 127-140, 2008.

KNUD, L. Information structure and sentence form. **Cambridge (Mas): Cambridge UP**, 1994.

KRAEPELIN, E. **Dementia praecox and paraphrenia**. Livingstone, 1919.

KUPERBERG, G. R. et al. Behavioral and electrophysiological approaches to understanding language dysfunction in neuropsychiatric disorders: Insights from the study of schizophrenia. **The Neuropsychology of Mental Illness. Cambridge University Press, New York**, p. 67-95, 2009.

KUPERBERG, G. R. Language in schizophrenia part 1: an introduction. **Language and linguistics compass**, v. 4, n. 8, p. 576-589, 2010.

LEZAK, M. D. et al. **Neuropsychological assessment**. Oxford University Press, USA, 2004.

LEES, R. B.; KLIMA, E. S. Rules for English pronominalization. **Language**, v. 39, n. 1, p. 17-28, 1963.

LONGOBARDI, G. Reference and proper names: A theory of N-movement in syntax and logical form. **Linguistic inquiry**, p. 609-665, 1994.

LONGOBARDI, G. **The structure of DPs: Some principles, parameters, and problems**. na, 2001.

LONGOBARDI, G. Toward a unified grammar of reference. **Zeitschrift für Sprachwissenschaft**, v. 24, n. 1, p. 5-44, 2005.

LUYSTER, R.; LORD, C. Word learning in children with autism spectrum disorders. **Developmental psychology**, v. 45, n. 6, p. 1774, 2009.

LYONS, C. **Definiteness**. Cambridge University Press, 1999.

MCKEE, C. A comparison of pronouns and anaphors in Italian and English acquisition. **Language acquisition**, v. 2, n. 1, p. 21-54, 1992.

MCKENNA, P. J.; OH, T. M. **Schizophrenic speech: Making sense of bathroofs and ponds that fall in doorways**. Cambridge University Press, 2005.

MCNALLY, L. VP coordination and the VP-internal subject hypothesis. **Linguistic Inquiry**, p. 336-341, 1992.

MERCHANT, J. **The syntax of silence: Sluicing, islands, and the theory of ellipsis**. Oxford University Press on Demand, 2001.

MITCHELL, R. L.; CROW, T. J. Right hemisphere language functions and schizophrenia: the forgotten hemisphere?. **Brain**, v. 128, n. 5, p. 963-978, 2005.

MOHAMED, S. et al. Generalized cognitive deficits in schizophrenia: a study of first-episode patients. **Archives of general psychiatry**, v. 56, n. 8, p. 749-754, 1999.

MONTEIRO, L. C.; LOUZÃ, M. R. Alterações cognitivas na esquizofrenia: consequências funcionais e abordagens terapêuticas. **Archives of Clinical Psychiatry (São Paulo)**, v. 34, p. 179-183, 2007.

MORICE, R.; MCNICOL, D. The comprehension and production of complex syntax in schizophrenia. **Cortex**, v. 21, n. 4, p. 567-580, 1985.

MORO, A. et al. Detecting syntactic and semantic anomalies in schizophrenia. **Neuropsychologia**, v. 79, p. 147-157, 2015.

MOTA, N. B. et al. Dreaming during the Covid-19 pandemic: Computational assessment of dream reports reveals mental suffering related to fear of contagion. **PloS one**, v. 15, n. 11, p. e0242903, 2020.

MÜLLER, S. **Head-Driven Phrase Structure Grammar: Eine Einführung**. 2013.

NAIGLES, L. R. et al. Neural correlates of language variability in preschool-aged boys with autism spectrum disorder. **Autism Research**, v. 10, n. 6, p. 1107-1119, 2017.

NESTOR, P. G. et al. A neuropsychological analysis of schizophrenic thought disorder. **Schizophrenia Research**, v. 29, n. 3, p. 217-225, 1998.

NIENDAM, T. A. et al. Neurocognitive performance and functional disability in the psychosis prodrome. **Schizophrenia research**, v. 84, n. 1, p. 100-111, 2006.

NOVOGRODSKY, R. Subject pronoun use by children with autism spectrum disorders (ASD). **Clinical linguistics & phonetics**, v. 27, n. 2, p. 85-93, 2013.

NOVOGRODSKY, R.; EDELSON, L. R. Ambiguous pronoun use in narratives of children with Autism Spectrum Disorders. **Child Language Teaching and Therapy**, v. 32, n. 2, p. 241-252, 2016.

ÖZCAN, A. et al. The production of simple sentence structures in schizophrenia. **International Journal of Arts & Sciences**, v. 9, n. 4, p. 159, 2017.

PARK, S.; GOODING, D. C. Working memory impairment as an endophenotypic marker of a schizophrenia diathesis. **Schizophrenia Research: Cognition**, v. 1, n. 3, p. 127-136, 2014.

PAROLA, A. et al. Pragmatics, Theory of Mind and executive functions in schizophrenia: Disentangling the puzzle using machine learning. **Plos one**, v. 15, n. 3, p. e0229603, 2020.

PELLETIER, F. J. Did Frege believe Frege's principle?. **Journal of Logic, Language and information**, v. 10, n. 1, p. 87-114, 2001.

PENN, D. L.; SANNA, L. J.; ROBERTS, D. L. Social cognition in schizophrenia: an overview. **Schizophrenia bulletin**, v. 34, n. 3, p. 408-411, 2008.

PERRY, W. et al. Working memory in schizophrenia: transient “online” storage versus executive functioning. **Schizophrenia bulletin**, v. 27, n. 1, p. 157-176, 2001.

PEUSKENS, J.; DEMILY, C.; THIBAUT, F. Treatment of cognitive dysfunction in schizophrenia. **Clinical Therapeutics**, v. 27, p. S25-S37, 2005.

PICALLO, M. C. **Nominals and nominalizations in Catalan**. 1991.

PICALLO, M. C. **On Gender and Number**. 2007.

RODRIGUEZ-FERRERA, S.; MCCARTHY, R. A.; MCKENNA, P. J. Language in schizophrenia and its relationship to formal thought disorder. **Psychological Medicine**, v. 31, n. 2, p. 197, 2001.

ROSSETTI, I.; BRAMBILLA, P.; PAPAGNO, C. Metaphor comprehension in schizophrenic patients. **Frontiers in psychology**, v. 9, p. 670, 2018.

REINHART, T. **Anaphora and Semantic Interpretation**. Chicago: University of Chicago Press. 1983a.

REINHART, T. Coreference and bound anaphora: A restatement of the anaphora questions. **Linguistics and Philosophy**, v. 6, n. 1, p. 47-88, 1983b.

REINHART, T. **Interface strategies: Optimal and costly computations**. The MIT Press, 2006.

REINHART, T. Processing or pragmatics? Explaining the coreference delay. **The processing and acquisition of reference**, p. 157-94, 2011.

REINHART, T. **Anaphora and semantic interpretation**. Routledge, 2016.

RITTER, E. Evidence for number as a nominal head. **GLOW Newsletter**, 1991.

RITTER, E. Where's gender?. **Linguistic inquiry**, v. 24, n. 4, p. 795-803, 1993.

RITTER, E. On the syntactic category of pronouns and agreement. **Natural Language & Linguistic Theory**, v. 13, n. 3, p. 405-443, 1995.

RIZZI, L. Early null subjects and root null subjects. **Language acquisition studies in generative grammar**, p. 151-176, 1994.

ROBERTS, I. **Parameter hierarchies and universal grammar**. Oxford University Press, 2019.

ROCHESTER, S.; MARTIN, J. R. **Crazy Talk: A study of the discourse of schizophrenic speakers**. Nova York, Plenum. 1979.

ROSSETTI, I.; BRAMBILLA, P.; PAPAGNO, C. Metaphor comprehension in schizophrenic patients. **Frontiers in psychology**, v. 9, p. 670, 2018.

SAAB, A. Syntax or nothing: Some theoretical and empirical remarks on implicit arguments. **Borealis—An International Journal of Hispanic Linguistics**, v. 3, n. 2, p. 125-183, 2014.

SALZMANN, M. The NP vs. DP debate. Why previous arguments are inconclusive and what a good argument could look like. Evidence from

agreement with hybrid nouns. **Glossa: a journal of general linguistics**, v. 5, n. 1, 2020.

SCHETTINO, A. et al. The comprehension of idiomatic expressions in schizophrenic patients. **Neuropsychologia**, v. 48, n. 4, p. 1032-1040, 2010.

SCHLENKER, P. Minimize restrictors!(Notes on definite descriptions, Condition C and epithets). In: **Proceedings of Sinn und Bedeutung**. p. 385-416, 2005.

SEVILLA, G. et al. Deficits in nominal reference identify thought disordered speech in a narrative production task. **Plos one**, v. 13, n. 8, p. e0201545, 2018.

SHLONSKY, U. Resumptive pronouns as a last resort. **Linguistic inquiry**, v. 23, n. 3, p. 443-468, 1992.

SITSKOORN, M. M. et al. Cognitive deficits in relatives of patients with schizophrenia: a meta-analysis. **Schizophrenia research**, v. 71, n. 2-3, p. 285-295, 2004.

SOLAN, L. Parameter setting and the development of pronouns and reflexives. In: **Parameter setting**. Springer, Dordrecht, p. 189-210, 1987.

SWANSON, H. L. Individual and age-related differences in children's working memory. **Memory & Cognition**, v. 24, n. 1, p. 70-82, 1996.

SZABOLCSI, A. **Model theoretic semantics of performatives**. 1982.

SZABOLCSI, A. The noun phrase. **The syntactic structure of Hungarian**, p. 179-274, 1994.

TAGER-FLUSBERG, H.; PAUL, R.; LORD, C. **Language and communication in autism**. 2005.

THOMAS, P. et al. Syntactic processing and written language output in first onset psychosis. **Journal of communication disorders**, v. 26, n. 4, p. 209-230, 1993.

THORNTON, R.; WEXLER, K. **Principle B, VP ellipsis, and interpretation in child grammar**. MIT press, 1999.

TITONE, D.; LEVY, D. L.; HOLZMAN, P. S. Contextual insensitivity in schizophrenic language processing: evidence from lexical ambiguity. **Journal of Abnormal Psychology**, v. 109, n. 4, p. 761, 2000.

TORREY, E. F. Studies of individuals with schizophrenia never treated with antipsychotic medications: a review. **Schizophrenia Research**, v. 58, n. 2-3, p. 101-115, 2002.

TOVAR, A. et al. The linguistic signature of hallucinated voice talk in schizophrenia. **Schizophrenia research**, v. 206, p. 111-117, 2019a.

TOVAR, A. et al. Language disintegration under conditions of severe formal thought disorder. **Glossa: a journal of general linguistics**, v. 4, n. 1, 2019b.

WALENSKI, M. et al. Grammatical processing in schizophrenia: Evidence from morphology. **Neuropsychologia**, v. 48, n. 1, p. 262-269, 2010.

WATSON, A. R. et al. Use of second-person pronouns and schizophrenia. **The British Journal of Psychiatry**, v. 200, n. 4, p. 342-343, 2012.

WETHERBY, A. M.; WOODS, J. J. Early social interaction project for children with autism spectrum disorders beginning in the second year of life: A preliminary study. **Topics in Early Childhood Special Education**, v. 26, n. 2, p. 67-82, 2006.

ZAMPARELLI, R. The structure of (in) definiteness. **Lingua**, v. 115, n. 6, p. 915-936, 2005.

ZIMMERER, V. C. et al. Deictic and propositional meaning—New perspectives on language in schizophrenia. **Frontiers in Psychiatry**, v. 8, p. 17, 2017.

ZOMICK, J.; LEVITAN, S. I.; SERPER, M. Linguistic analysis of schizophrenia in Reddit posts. In: **Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology**. p. 74-83, 2019.

Anexo 1 - *flairs* de autodeclaração (esquizofrenia)

Schizophrenia
 Schizophrenic
 Residual Schizophrenia
 Paranoid Schizophrenia
 Paranoid Schizophrenic
 schizophrenia unspecified type
 paranoid schizophrenia
 schizophrenia
 Schitzophrenic
 Schizophrenia
 schizophrenic
 Schizophrenia Girl
 residual schizophrenia
 undifferentiated schizophrenia
 A schizophrenic
 schizophrenic psycho
 Disorgarinzed Scihozhpernia
 Catatonic Schizophrenia
 Diagnosed with schizohpernia
 Paranoid schizophrenia and tapering off Rexulti
 Undifferentiated schizophrenia
 Schizophrenia, but I will always love myself <3
 Paranoid schizophrenia & taking Rexulti 4 mg
 Paranoid schizophrenia and taking Rexulti
 paranoid schizophrenic
 Schizophrenia (Neg Hell)
 (I was?) Simple Schizophrenic
 schizophrenia, undifferentiated type
 Residual Schizophrenic
 Paranoid schizophrenia
 Hebephrenic Schizophrenic
 Schizophrenia Undifferentiated
 Paranoid Schizophrenia
 Schizophrenia 🌙
 Undifferentiated Schizophrenic
 Schizophrenia | Managing with meds
 simple schizophrenia
 Schizophrenic NOS
 Catatonic Schizophrenia
 Refractory Paranoid Schizophrenia
 paranoid schizophrenic
 Disorganized Schizophrenia
 disorganized schizophrenic
 Paranoid schizophrenic
 Schizophrenia, but happy
 Schizophrenia residual
 Childhood-Onset Schizophrenic
 Schizophrenic

schizophrenia paranoid type
Catatonic schizophrenic
Schizophrenia 🦊
Schizophrenia 🦊
Unspecified Schizophrenia
Schizophrenic Mod
Disorganised Schizophrenia (Delusional/Chuunibyou)
schizophrenic/clairvoyant
Schizophrenia paranoid psychotic
Schizophrenia ^-^
Schizophrenia, but I will always love myself <3
Unspecified schizophrenia
schizophrenia @ SAD
Paranoid schizophrenic

Anexo 2 - *flairs* excludentes

Schizoaffective, BP type
 Schizoaffective-Bipolar Type
 Bipolar : Schizoaffective : PTSD
 Schizoaffective Bipolar
 Schizoaffective Bipolar Subtype
 Schizoaffective - Bipolar
 autism 1 with psychosis
 schizoaffective bipolar
 schizoaffective something
 Psychotic Depression
 Schizoaffective Bipolar Type
 Schizoaffective-bipolar
 SZ and OCD
 mdd w psychosis
 paranoid schizophrenia, major depression, and anxiety
 SZA BP
 Schizoaffective - Bipolar type
 Schizoaffective bipolar type
 schizoaffective, bipolar type
 Schizoaffective Bipolar
 schzodepressive
 schizoaffective depressive type
 Psychotic, GAD, Depression, PTSD
 Schizoaffective bi polar
 Schizoaffective Depressive
 Schizoaffective, BP Type
 SZA-BPD1, OCD
 sza bp etc. etc.
 Schizo/Dep/Anx
 Schizoaffective (BP type)
 Schizoaffective (bipolar type)
 Sza, Bipolar
 Schizoaffective d/o bipolar type
 SZA Depressive
 Loved One | anxiety and other problems
 Schizoaffective: bipolar type, apparently
 psychosis and OCD, hope it's not schizophrenia :(
 Schizodepressive Disorder
 Schizoaffective Bipolar type
 Schizoaffective, Bipolar Type
 Schizotypal + OCD 'n' mild Tourette
 psychosis and depression
 Schizoaffective Depressive type
 depression
 schizo affective depression
 Schizophrenia + ADD
 Schizoaffective Depressive Type
 bipolar flavored

Schizoaffective (Depressive Type)
 Schizoaffective/Anxiety
 Schizoaffective (Bipolar)
 sza bipolar
 Schizophrenia + depression
 schizodepressive
 paranoid schizophrenia bipolar
 SZA bipolar
 Schizoannoying 🦋 • (schizoaffective bp)
 Schizoaffective (Depressive Subtype)
 WIP Diagnosis, Cognitive Impairment
 PTSD w/ psychosis
 28 | Schizoaffective Depressive
 BP1 w/ psychosis
 Mood Disorder NOS + Psychosis
 Visual Snow, SZA-D
 Signal Noise, sza-d
 Schizoaffective Disorder, Depressive Type
 schizoaffective -bipolar type
 schizoaffective - bipolar type
 disorganized? adhd? even my psych can't tell
 SZA/PTSD/OCD
 Bipolar?
 Schizoaffective-bipolar
 Schizoaffective - Depressive
 schizoaffective, bipolar type
 schizoaffective d/o bipolar type
 Bipolar 2 w/psychotic features
 schizoaffective (bipolar type)
 Schizoaffective and BPD
 Schizoaffective (Bipolar Subtype)
 Level 2 ASD w/SZ
 SZA/PTSD
 Schizo/Dep/Anx/OCD
 SZA-Depressive Type
 Schizophrenia & ADHD
 bipolar w delusions
 Schizoaffective by Ticcing
 Bipolar and paranoid
 schizoaffective & OCD
 Schizophrenia + BPD
 The OCD dude
 Schizoaffective—Bipolar Type
 sad/ocd
 Schizoaffective - Bipolar Type
 Sza-bipolar type w/bpd
 Psychosis, OCD, PTSD
 Dx'd STPD/SZA BP
 schizoaffective bipolar type
 Schizoaffective, ADHD-C, PTSD

Schizoaffective & Autistic
 SZA and OCD
 SZ / BPD
 STPD
 SZA Ticcing Borderline
 Cotard's Delusion + Capgras Syndrome
 schizoaffective (bipolar i + schiz)
 depressive w/ psychotic features
 Severe depression w/ psychotic and paranoid effects
 Schizoaffective bipolar
 bipolar I
 Schizotypal w/ PF
 Chaplain w/CPTSD & Schizo family
 Schizophrenia in Remission + ADD
 Schizotypal & Psychotic Depression
 SCZASPD
 schizoaffective & OCD
 Depression & Alcoholism
 Bipolar Schizophrenic
 paranoid schizoaffective / bpd
 SZA Bipolar Type
 Schizoaffective (bipolar subtype)
 Schizotypal/Cyclothymia
 SZA Ebenezer Bad
 Psychotic depression
 Schizoaffective depressive type
 BPD
 Schizophrenia, Anxiety disorder, ASPD
 MDNOS + Acute Psychosis
 SPD + OCD 'n' mild Tourette
 STPD, OCD
 Schizoaffective depressive
 Schizophrenia & ADHD
 Schizoaffective - Depression
 Schizoaffective(Bipolar Type)
 Undxed Sleep disorder(hypnagogic hallucinations)
 Visual Snow, SZA-D, Episodes
 Schizoaffective Bipolar I
 simple schizophrenia and borderline personality
 Schizoaffective—Bipolar Subtype
 psychosisdepression
 depression with psychosis
 stpd, ocd, avpd
 Diagnosed USS
 Sza-bipolar type
 Personality disordered beyond repair
 Schizoaffective (Mixed)
 Dual Diagnosis
 Intense Signal Noise, sza-d
 Psychotic Ultradian Bipolar 1, Paranoid SZ

OCD and maybe some Psychotic Disorder
 Depresson w severe psychosis
 OCD + psychosis
 psychosis/bpd/ocd
 Chaplain w/CPTSD & Schizo loved ones
 Misdiagnosed survivor
 Chronic Childhood Schizophrenia & Complex PTSD
 Schizoaffective, Bipolar type
 Schizophrenic/Bipolar
 Autism Spectrum Disorder
 Depression, Psychotic Features
 Schizophrenic Depressive
 Schizoaffective & Autistic
 osdd/bpd
 Schizoaffective depressive type.
 Bipolar. BF is schizoaffective.
 Schizoaffective Bi-Polar
 SZA, depressive
 OCD w/ Psychotic Features
 Obsessive Compulsive Disorder
 Major Depression with Psychotic Features
 scizophrenia ocd depression
 Schizophrenia, BP1, PAD, OCD
 Schizoaffective BP
 bipolar w/ psychotic symptoms
 SZ + apparently bipolar rapid-cycling
 Schizoaffective DID
 Prodromal Schizo + OCD 'n' Tourette
 Schizophrenia + depression
 On SSDI: schizophrenia, PTSD, clinical depression
 Autistic
 depressive subtype
 schizoaffective & BPD
 bipolar subtype
 19M/Schizo/PTSD/Autism
 18M/Schizo/PTSD/Autism
 big bipolar dyke bitch
 Manic
 BP1 with psychosis | BPD
 Multiple mental and physical illnesses
 GAD | PTSD | Schizophrenia
 Paranoid schizophrenic | PTSD
 I have crippling depression
 Schizoaffective - Depressive
 ASPD
 Depression
 ADHD-C
 Bipolar 1
 ADHD
 ADHD & Schizophrenic

Bipolar 1 w/ psychotic features
Major Depression with Psychosis
Schizophrenia, PTSD, major depression, GAD
Paranoid schizophrenia & taking Rexulti 4 mg
Adult Autistic
On SSDI. Schizophrenia, PTSD, clinical depression, GAD.
On SSDI/SSI. Schizophrenia, PTSD, major depression, GAD.
Paranoid schizophrenia and tapering off Rexulti
Paranoid schizophrenia and taking Rexulti
28F with PTSD, N-mom & E-dad, now 2,000 miles away & debating LC
Schizophrenia, PTSD, major depression, anxiety NOS
On SSDI/SSI. Schizophrenia, PTSD, major depression, anxiety NOS.
BPD|GAD|Dysphoric|Psychotic
Schizoaffective bipolar type
Schizoaffective bipolar
Asperger's
paranoid schizophrenia and ocd
GAD, Schizophrenia

Anexo 3 - lista lexical excludente (transtornos mentais)

SCHIZOAFFECTIVE-RELATED

SZA-B

SZA-D

SZA-BP

SZA-DP

SZABP

Schzodepressive

Schizobipolar

SZAD

BIPOLAR

bipolar

bipolarity

ANXIETY

aniexty

anxiety

DEPRESSION

depression

depressed

depress

depressive

PTSD

post-traumatic

PTSD

ADHD

ADHD-C

ADHD

Attention deficit

attention-deficit

attention deficit

atention-deficit

OCD

Obsessive-Compulsive

Obsessive Compulsive

Compulsive-Obsessive

Compulsive Obsessive

Obsesive-Compulsive

Obsesive Compulsive

OCD

LANGUAGE AND SPEECH DISORDERS

language disorder

speech disorder

language disorder
 speech disorder
 language disorder
 speech disorder
 Speech Sound Disorder
 speech sound disorder
 speech disorder
 speech sound disorder

MDD
 MDD

TOURETTES
 Tourette
 tourete

TIC DISORDER
 Ticking
 Tic Disorder
 tic disorder
 ticc

ALZHEIMERS
 Alzheimer
 Alzeimer

ASPERGERS
 Asperger
 Aspeger
 Aspie

DISABILITIES
 intellectual disability
 intelectual disability
 developmental disability

STUTTERING
 Stuttering
 Stammering

COMMUNICATION DISORDERS
 SCD
 Social Communication Disorder
 Pragmatic Communication disorder
 social communication disorder
 pragmatic communication disorder

AUTISM
 Autism
 Autistic

Autie

SPECIFIC LEARNING DISORDER

SLD

learning disability

leaningdisability

DYSLEXIA

dyslexia

dyslexic

NERUDEVELOPMENTAL DISORDERS

neurodevelopmental

CATATONIA

catatonia

catatonic

MANIA

mania

manic

maniac

maniacal

hypomanic

CYCLOTHYMIA

Cyclothymic

Cyclothymia

MDD

MDD

Mood dysregulation

DMDD

DYSTHYMIA

Dysthymia

Dysthymic

ACUTE STRESS

Acute stress

stress disorder

stress acute

stres acute

stres disorder

DISSOCIATIVE DISORDERS

OSDD

Dissociative disorder

dissociativedisorder

disociative disorder

disociativedisorder
DissociativeIdentity
disociativeidentity
dissociative identity
dissociative identity

AMNESIA
amnesia

NERUCOGNITIVE DISORDER
NeurocognitiveDisorder
neurocognitive disorder

DEMENTIA
Dementia

PARKINSONS
Parkinson

HUNTINGTONS
Huntington

SCLEROSIS
Sclerosis

ENCEPHALOPATIES
Hydrocephalus
encephalopathy
encephalopathies

BORDERLINE
Borderline

SUFFERED STROKE
Stroke

Anexo 4 - lista lexical (idade)

birthday
birthdate
bday
bparty
anniversary
aniversary
ten
eleven
twelve
thirteen
thirten
fourteen
fourten
forteen
forten
fifteen
fiften
sixteen
sixten
seventeen
seventen
eighteen
eighten
nineteen
nineten
twenty
thirty
forty
fifty
sixty
seventy
eighty
ninety
hundred
lifetime
summers
sumers
winters
springs
autumns
life
maturity
youth
adolescence
adult
boy
childhood
childhod

girl
infancy
lifetime
majority
milestone
senility
seniority
elderliness
generation
date
grow
mature
ago
kid
born

Anexo 5 - lista lexical (sexo)

tomboy
ponytail
sex
makeup
nails
earrings
abs
biceps
pp
fap
jack off
jill off
schlick
squirt
masturbat
ejaculat
porn
rape
glans
clitoris
breast
nipple
niple
erogenous
touch
nsfw
baby
daddy
dady
mommy
momy
sugar
feet
nurse
pee pee
fornication
intimacy
love
sexual
feminine
femininity
masculinity
masculine
reproduction
sensual
afab
amab
androgyny

gyno
dysphoria
cis
trans
gender
demi
femme
ftm
mtf
identity
fluid
normative
queer
lgbt
homo
hetero
gyne
gay
lesbian
pan
poly
straight
romantic
relationship
genital
gina
urologist
pussy
penis
vagina
boy
girl
son
daughter
brother
sister
father
mother
boyfriend
girlfriend
woman
man
tie
dress
skirt
panty
husband
wife
male
female

vasectomy
my period
pregnant
beard

Anexo 6 - lista lexical (nativos)

usa
uk
america
asia
oceania
europe
north
south
west
east
southern
northern
western
eastern
native
speaker
language
hometown
country
nationality
city
continent
province
republic
county
township
island
territory
federation
birthplace
motherland
homeland
state
bilingual
sign
afrikaans
amharic
arabic
aymara
basque
bengali
bislama
catalan
chamorro
chichewa
divehi
dzongkha
english

faroese
galician
greenlandic
guarani
hebrew
hindi
icelandic
kazakh
kirghiz
kirundi
kongo
korean
kurdish
latin
lingala
luba katanga
luba
katanga
luxembourgish
malay
manx
maori
nepali
ndebele
nynorsk
occitan
panjabi
punjabi
pashto
persian
peul
quechua
rwandi
sango
shona
sinhalese
slovak
southern sotho
swahili
swati
tamil
tigrinya
tsonga
tswana
turkmen
urdu
uzbek
venda
xhosa
zulu

afghanistan
albania
algeria
angola
anguilla
antarctica
antigua
barbuda
argentina
armenia
aruba
australia
austria
azerbaijan
bahamas
bahrain
bangladesh
barbados
belarus
belgium
belize
benin
bermuda
bhutan
bolivia
bosnia
herzegovina
botswana
bouvet
brazil
brunei
darussalam
bulgaria
burkina faso
burundi
cambodia
cameroon
canada
cape verde
cayman
chad
chile
china
cocos
keeling
colombia
comoros
korea
congo
cook

costa rica
ivory
divoire
cote
d ivoire
croatia
cuba
cyprus
denmark
dominica
east timor
ecuador
egypt
salvador
eritrea
estonia
ethiopia
falkland
faroe
fiji
finland
france
guiana
polynesia
gabon
gambia
georgia
germany
ghana
greece
greenland
grenada
guadeloupe
guam
guatemala
guinea
guyana
haiti
heard
mcdonald
holy see
honduras
hungary
iceland
india
indonesia
iran
iraq
ireland
israel

italy
jamaica
japan
jordan
kazakstan
kenya
kosovo
kyrgyzstan
laos
latvia
lebanon
lesotho
liberia
jamahiriya
liechtenstein
lithuania
macau
macedonia
madagascar
malawi
malaysia
maldives
mali
malta
marshall
martinique
mauritania
mauritius
mayotte
mexico
micronesia
moldova
mongolia
montserrat
montenegro
morocco
mozambique
myanmar
namibia
nauru
nepal
netherlands
antilles
new caledonia
zealand
nicaragua
niger
nigeria
niue
norfolk

northern mariana
norway
oman
pakistan
palau
palestinian
paraguay
peru
philippines
pitcairn
poland
portugal
puerto rico
qatar
romania
russia
rwanda
saint helena
saint kitts
nevis
saint lucia
saint pierre
miquelon
saint vincent
grenadines
samoa
principe
arabia
senegal
serbia
seychelles
sierra leone
slovakia
slovenia
somalia
africa
spain
sri lanka
sudan
suriname
svalbard
jan mayen
swaziland
sweden
switzerland
taiwan
tajikistan
tanzania
thailand
togo

tokelau
tonga
trinidad
tobago
tunisia
turkey
turkmenistan
turks
caicos
tuvalu
uganda
ukraine
arab
emirates
united kingdom
united states
united
great britain
uruguay
uzbekistan
vanuatu
venezuela
vietnam
wallis
futuna
sahara
yemen
zambia
zimbabwe
afghan
albanian
algerian
american
andorran
angolan
antiguans
argentinean
armenian
australian
austrian
azerbaijani
bahamian
bahraini
bangladeshi
barbadian
barbudans
batswana
belarusian
belgian
belizean

beninese
bhutanese
bolivian
bosnian
brazilian
british
bruneian
bulgarian
burkinabe
burmese
burundian
cambodian
cameroonian
canadian
cape verdean
chadian
chilean
chinese
colombian
comoran
congolese
costa rican
croatian
cuban
cypriot
czech
danish
dutch
timorese
ecuadorean
egyptian
emirian
equatorial
eritrean
estonian
ethiopian
fijian
filipino
finnish
french
gabonese
gambian
georgian
german
ghanaian
greek
grenadian
guatemalan
bissauan
guinean

guyanese
haitian
herzegovinian
honduran
hungarian
kiribati
icelander
indian
indonesian
iranian
iraqi
israeli
italian
ivorian
jamaican
japanese
jordanian
kazakhstani
kenyan
kittian
nevisian
kuwaiti
kyrgyz
laotian
latvian
lebanese
liberian
libyan
liechtensteiner
lithuanian
luxembourger
macedonian
malagasy
malawian
malaysian
maldivian
malian
maltese
marshallese
mauritanian
mauritian
mexican
micronesian
moldovan
monacan
mongolian
moroccan
mosotho
motswana
mozambican

namibian
nauruan
nepalese
zealander
ni vanuatu
nivanuatu
nicaraguan
nigerian
nigerien
irish
norwegian
omani
pakistani
palauan
panamanian
papua
paraguayan
peruvian
polish
portuguese
qatari
romanian
russian
rwandan
saint lucian
salvadoran
samoan
san marinese
sao tomean
saudi
scottish
senegalese
serbian
seychellois
sierra leonean
singaporean
slovakian
slovenian
islander
solomon
somali
african
spanish
sri lankan
lankan
sudanese
surinamer
swazi
swedish
swiss

syrian
taiwanese
tajik
tanzanian
thai
togolese
tongan
trinidadian
tobagonian
tunisian
turkish
tuvaluan
ugandan
ukrainian
uruguayan
uzbekistani
venezuelan
vietnamese
welsh
yemenite
zambian
zimbabwean
ulaan baatar
vientian
katmandu
casablanca
petersburg
wien
new york
los angeles
rio de janeiro
karachi
chongqing
kolkata
istanbul
lagos
tianjin
guangzhou
shenzhen
lahore
bangalore
chennai
nagoya
hyderabad
chengdu
nanjing
wuhan
ho chi minh
ahmedabad
xian

xi an
dongguan
hangzhou
foshan
shenyang
surat
suzhou
pune
harbin
belo horizonte
fukuoka
johannesburg
qingdao
dalian
yangon
alexandria
jinan
guadalajara
dubai
barcelona
chicago
san francisco
las vegas
sydney
milan
san diego
vancouver
zurich
houston
seattle
hong kong
frankfurt
sao paulo
tel aviv
calgary
orlando
atlanta
dallas
hamburg
osaka
austin
phoenix
naples
denver
philadelphia
delhi
portland
valencia
brisbane
minneapolis

shanghai
lyon
adelaide
edmonton
stuttgart
baltimore
auckland
cologne
new orleans
kuwait city
kyiv
hanover
perth
nashville
dusseldorf
sacramento
glasgow
salt lake
raleigh
kracow
melbourne
montreal
manchester
mumbai
toronto
birmingham
miami
porto
munich
boston
marseille
capital
abu dhabi
abuja
accra
adamstown
addis ababa
algiers
alofi
amman
amsterdam
andorra
ankara
antananarivo
apia
ashgabat
asmara
astana
asuncion
athens

avarua
baghdad
baku
bamako
bandar seri begawan
bangkok
bangui
banjul
basseterre
beijing
beirut
belgrade
belmopan
berlin
bern
bishkek
bissau
bogota
brasil
bratislava
brazzaville
bridgetown
brussels
bucharest
budapest
buenos aires
bujumbura
cairo
canberra
caracas
castries
cayenne
charlotte amalie
chisinau
cockburn
conakry
copenhagen
dakar
damascus
dhaka
dili
djibouti
dodoma
dar es salaam
doha
douglas
dublin
dushanbe
edinburgh
el aaiun

episkopi cantonment
flying fish cove
freetown
funafuti
gaborone
town
georgetown
gibraltar
grytviken
guatemala city
gustavia
hagatna
hamilton
hanga roa
hanoi
harare
hargeisa
havana
helsinki
honiara
islamabad
jakarta
jamestown
jerusalem
ramallah
gaza
juba
kabul
kampala
kathmandu
khartoum
kiev
kigali
kingston
kingstown
kinshasa
kuala lumpur
putrajaya
kuwait
libreville
lilongwe
lima
lisbon
ljubljana
lome
london
luanda
lusaka
luxembourg
madrid

majuro
malabo
managua
manama
manila
maputo
marigot
maseru
mata utu
matautu
mbabane
lobamba
melekeok
ngerulmud
minsk
mogadishu
monaco
monrovia
montevideo
moroni
moscow
muscat
nairobi
nassau
naypyidaw
djamena
niamey
nicosia
nouakchott
noumea
nuku alofa
nukualofa
nuuk
oranjestad
oslo
ottawa
ouagadougou
pago pago
palikir
panama
papeete
paramaribo
paris
philipsburg
phnom penh
plymouth
brades estate
podgorica
cetinje
port louis

port moresby
port vila
port au prince
port of spain
porto novo
cotonou
prague
praia
pretoria
bloemfontein
cape town
pristina
pyongyang
quito
rabat
reykjavik
riga
riyadh
rome
roseau
saipan
san jose
san juan
san marino
san salvador
sana a
sanaa
santiago
valparaiso
santo domingo
sao tome
sarajevo
seoul
singapore
skopje
sofia
sri jayawardenepura kotte
colombo
st george s
georges
helier
john s
johns
peter port
pierre
stanley
stepanakert
stockholm
sucre
la paz

sukhumi
suva
taipei
tallinn
tarawa
tashkent
tbilisi
kutaisi
tegucigalpa
tehran
thimphu
tirana
tiraspol
tokyo
torshavn
tripoli
tskhinvali
tunis
ulan bator
vaduz
valletta
the valley
vatican city
victoria
vienna
vientiane
vilnius
warsaw
washington
wellington
willemstad
windhoek
yamoussoukro
abidjan
yaounde
yaren
yerevan
zagreb