

## **O CORPUS C-ORAL-ESQ E A ESTRUTURA INFORMACIONAL DA FALA DE PACIENTES COM ESQUIZOFRENIA**

THE C-ORAL-ESQ CORPUS AND THE INFORMATIONAL STRUCTURE  
OF PATIENTS WITH SCHIZOPHRENIA

Bruno Neves Rati de Melo Rocha | [Lattes](#) | [bbruno791@gmail.com](mailto:bbruno791@gmail.com)

Universidade Federal de Minas Gerais

**Resumo:** Esse trabalho apresenta o C-ORAL-ESQ, corpus oral de pacientes brasileiros com esquizofrenia. O corpus documentará 40 sessões de atendimento psicoterapêutico a pacientes de ambos os sexos, com faixa etária e nível de escolaridade variados. O C-ORAL-ESQ será segmentado em enunciados e unidades tonais a partir de um critério prosódico e pragmático, o qual concebe os enunciados como atos de fala (AUSTIN, 1962), ou seja, sequências com autonomia prosódica e pragmática (CRESTI, 2000). Além disso, o artigo reporta estudos de Cresti, Dovetto e Rocha sobre a estrutura informacional da fala de pacientes italianos com esquizofrenia (CRESTI; DOVETTO; ROCHA, 2015). Os autores notam em alguns pacientes uma tendência para produzir enunciados com menor variação de tipos ilocucionários e de unidades tonais não ilocucionárias.

**Palavras-chave:** Corpus oral; Esquizofrenia; Atos de fala; Prosódia; Language into Act Theory.

**Abstract:** This paper presents the C-ORAL-ESQ, a speech corpus of Brazilian patients with schizophrenia. The corpus will comprise 40 psychotherapy sessions with male and female patients, different age and schooling levels. The C-ORAL-ESQ will be segmented into utterances and intonation units based on a prosodic/pragmatic criterion that regards the utterance as a speech act (AUSTIN, 1962), i.e., a stretch of speech that is autonomous from the prosodic and pragmatic standpoint (CRESTI, 2000). Furthermore, this paper reports studies carried out by Cresti, Dovetto, and Rocha that have looked at informational structure in the speech of Italian patients with schizophrenia (CRESTI; DOVETTO; ROCHA, 2015). The authors have pointed out a tendency for some patients to produce utterances with less variability in terms of both illocutionary types and non-illocutionary tone units.

**Keywords:** Speech corpus; Schizophrenia; Speech acts; Prosody; Language into Act Theory.

## 1 INTRODUÇÃO

Esse artigo (a) apresenta o C-ORAL-ESQ, corpus oral de pacientes brasileiros com esquizofrenia e (b) reporta um conjunto de estudos sobre a estrutura informacional da fala de pacientes italianos com esquizofrenia.

A esquizofrenia é uma doença mental grave que acomete cerca de 1% da população mundial e afeta as relações sociais, o trabalho, o autocuidado e a vida familiar dos pacientes. Tipicamente, tem início entre 15 e 24 anos no sexo masculino e entre 25 e 34 anos no sexo feminino. Caracteriza-se por sintomas positivos (delírios, alucinações, discurso desorganizado e comportamento motor anormal ou desorganizado), sintomas negativos (embotamento afetivo, avolição, alogia, anedonia e associalidade) e déficits cognitivos (prejuízos em capacidades como a abstração, a memória, o raciocínio e a resolução de problemas). Dentre eles, o déficit cognitivo é o principal responsável pelos fortes prejuízos promovidos pela doença. A quantidade e a intensidade dos sintomas em cada paciente podem variar expressivamente. De fato, atualmente, entende-se que a esquizofrenia pode ser melhor compreendida não como uma única doença, mas como um conjunto de síndromes que resultam em um quadro clínico característico (SALGADO, 2008).

Não se tem um consenso sobre a patogênese da esquizofrenia, mas pesquisas recentes indicam que se deve a uma interação entre fatores ambientais e genéticos durante o desenvolvimento neurológico (LEWIS; LEVITI, 2002). No plano neuroquímico, pacientes apresentam alterações na concentração de neurotransmissores como dopamina, serotonina e glutamato. Estudos de neuroimagem mostram que a esquizofrenia correlaciona-se a uma atrofia de regiões cerebrais como o córtex pré-frontal, o giro cíngulo e partes do lobo temporal (SALGADO, 2008). Imagens de fMRI realizadas durante a oitiva de sentenças com e sem emoções mostram que pacientes – especialmente, aqueles com sintomas negativos mais acentuados – possuem maior lateralização no hemisfério direito (BACH et al., 2009).

A esquizofrenia afeta de maneira expressiva a forma e/ou o conteúdo da linguagem dos pacientes, com alta variabilidade interindividual. Algumas das alterações comumente citadas pela literatura são: aceleração ou redução do fluxo da fala, descarrilamento sintático-semântico, tangencialidade, pobreza de ideias, falta de lógica, ecolalia, palíllia, forte presença de neologismos, concretismo e falta de um projeto discursivo público (DOVETTO, 2014; FROMKIN, 1975; LANGUUPES, 1974; LIDDLE et al., 2002; SCHULTZ; ANDREASEN, 1999)

Com relação à prosódia, são reportadas alterações na quantidade e duração das pausas (maiores em pacientes com alogia) e nos valores médios de f0 e intensidade (menores em pacientes com embotamento afetivo) (COHEN; KIM; NAJOLIA, 2013). Inúmeros estudos documentam que pacientes têm piores desempenhos na percepção e na produção da prosódia emocional (HOEKERT et al., 2007), bem como em tarefas de percepção da diferença de frequência entre tons (LEITMAN et al., 2005). Além disso, pacientes têm maiores dificuldades em sinalizar prosodicamente o final de turnos discursivos (ALPERT et al., 2002).

Atualmente, observa-se um crescente uso de métodos objetivos de análise de dados, ao invés de medições subjetivas (BEDWELL et al., 2014). Ainda assim, há poucos estudos conduzidos a partir da análise da fala espontânea de pacientes, reforçando a necessidade de um corpus como o C-ORAL-ESQ.

## **2 O CORPUS C-ORAL-ESQ**

O C-ORAL-ESQ é um corpus em fase inicial de compilação que pretende documentar e disponibilizar à comunidade a fala de pacientes com esquizofrenia, possibilitando a realização de estudos linguísticos e médicos. O corpus resulta do projeto “A compilação do corpus C-ORAL-ESQ e o estudo da fala de pacientes com esquizofrenia”, de Tommaso Raso (coordenador) e João Salgado (subcoordenador), da Universidade Federal de Minas Gerais.

### **2.1 Características gerais do corpus**

O C-ORAL-ESQ registra sessões regulares de atendimento psicoterapêutico no ambulatório do Instituto Raul Soares (IRS, Fundação Hospitalar do Estado de Minas Gerais – FHEMIG), em Belo Horizonte, e será composto por, no mínimo, 40 gravações com cerca de 1.500 palavras do paciente (às quais se somam às palavras do médico e de outros participantes), em um total de aproximadamente 30.000 palavras de pacientes. Posteriormente, o corpus será disponibilizado para consultas gratuitas no site do projeto C-ORAL-BRASIL<sup>1</sup>.

Para garantir validade ecológica aos dados, são registrados os atendimentos regulares oferecidos pela equipe do IRS nas próprias salas do ambulatório do hospital, às quais os pacientes estão habituados, fazendo uso de equipamentos pouco invasivos (cf. 2.4). São gravadas as vozes de todos os participantes da sessão comunicativa (médicos, enfermeiros, pacientes e familiares de pacientes). Após instalarem adequadamente os apare-

<sup>1</sup> <<http://www.c-oral-brasil.org>>.

lhos de gravação, os pesquisadores se posicionam em uma sala diferente daquela em que ocorre o atendimento, para minimizar o impacto sobre a situação comunicativa.

As gravações do C-ORAL-ESQ apresentam grandes quantidades de enunciados de cada paciente, produzidos espontaneamente, em uma situação que faz parte do seu cotidiano. Assim como estão sendo registradas sessões de acompanhamento médico, espera-se documentar enunciados com diversos temas pertencentes ao dia a dia dos pacientes e com um espectro variado de emoções.

Cada uma das 40 gravações contará com:

1. Arquivo de áudio .wav, com taxa de amostragem de 44100Hz;
2. Arquivo de texto .rtf com a transcrição semiortográfica e a segmentação prosódica dos enunciados;
3. Arquivo de alinhamento texto-som .xml, utilizável no programa *WinPitch* (MARTIN, 2004);
4. Um arquivo de texto .txt com os metadados da sessão.

Por fim, será constituído um subcorpus de 20 textos e 15.000 palavras que passará pelo processo de etiquetagem informacional segundo os preceitos da Language into Act Theory (CRESTI, 2000) (cf. 3.1). Para cada texto do subcorpus, será produzido um arquivo .txt de transcrição e um arquivo .xml de alinhamento texto-som adicional.

Em dezembro de 2017, o projeto C-ORAL-ESQ foi aprovado pelo Comitê de Ética em Pesquisa da UFMG, e até agosto de 2019 foram feitas 22 gravações. Anteriormente ao início das gravações, foram realizadas oficinas de formação para a equipe técnica responsável pela compilação do corpus no LEEL (Laboratório de Estudos Linguísticos e Experimentais da Linguagem), da Faculdade de Letras da UFMG. O treinamento compreendeu oficinas de (a) gravação, (b) transcrição e segmentação prosódica da fala e (c) alinhamento texto-som com programa *Winpitch* (MARTIN, 2004).

## **2.2 A comparabilidade com outros corpora**

No âmbito do projeto C-ORAL-BRASIL, sediado na Universidade Federal de Minas Gerais e coordenado por Tommaso Raso e Heliana Mello, está sendo programada a compilação de um corpus de controle que registrará pelo menos 40 interações médico-paciente com pessoas com doenças crônicas e sem transtornos mentais, pareado por sexo e idade com relação aos pacientes do C-ORAL-ESQ. A escolha de pacientes com doenças crônicas tem o objetivo de reproduzir algumas características encontradas no C-ORAL-ESQ: 1. paciente e médico estão familiarizados um com o outro; 2. paciente

e médico já estão engajados em um tratamento em curso; 3. nesse tipo de consulta, é comum que se façam perguntas sobre o autocuidado do paciente e sobre a sua vida em família; 4. nesse tipo de consulta, é comum que o paciente descreva eventos com alta carga emocional. O corpus de controle está em fase de planejamento, mas é certo que será comparável ao C-ORAL-ESQ do ponto de vista situacional e receberá o mesmo tratamento metodológico. Recentemente, foi aprovada a emenda ao Comitê de Ética em Pesquisa da UFMG para que os dados possam ser coletados.

Além disso, o C-ORAL-ESQ se insere no projeto *Speech Pathology and Cognitive Studies*, coordenado por Francesca Dovetto, que visa construir corpora de pacientes com patologias variadas, falantes de diferentes línguas. Atualmente, conta com o corpus CIPPS (DOVETTO; GEMELLI, 2012), de pacientes italianos com esquizofrenia. Nesse contexto, o C-ORAL-ESQ poderá ser usado como fonte de dados para pesquisas interlinguísticas sobre a fala patológica, ampliando o seu potencial de exploração.

Adicionalmente, O C-ORAL-ESQ será comparável aos corpora da família C-ORAL, que documentam a fala espontânea de pessoas sem transtornos mentais. São eles o C-ORAL-BRASIL (RASO; MELLO, em preparação, 2012), de Português Brasileiro, e o C-ORAL-ROM (CRESTI; MONEGLIA, 2005), de Italiano, Português Europeu, Francês e Espanhol.

O C-ORAL-ROM é um corpus de Italiano, Português Europeu, Francês e Espanhol, fruto de um consórcio financiado pela União Europeia e sediado na Universidade de Florença (Itália). Cada língua é representada por cerca de 300.000 palavras divididas em duas seções. Uma delas é dedicada ao registro da fala informal, e a outra compreende a fala formal (aulas, conferências, discursos políticos etc.), a fala em mídia (telejornais, *talk shows* etc.) e a fala em interações telefônicas. O C-ORAL-ROM tem como principal objetivo registrar a fala espontânea com a maior variação diafásica (variação de situações comunicativas) possível.

O C-ORAL-BRASIL é um corpus de português brasileiro produzido pelo projeto homônimo. O corpus foi elaborado a partir dos critérios estabelecidos para o C-ORAL-ROM e possui a mesma arquitetura, mas nesse corpus foi dada atenção especial à variação diafásica e à qualidade acústica das gravações. Com equipamentos de gravação portáteis de alta definição, foi possível registrar interações tão variadas como uma aula de direção de trânsito e um jogo de futebol. A seção informal do corpus contém 208.130 palavras, divididas em 139 textos de aproximadamente 1.500 palavras (RASO; MELLO, 2012). A seção formal (74 textos e 123.000 palavras), de mídia (101 textos e 144.000 palavras)

e de interações telefônicas (78 textos e 32.000 palavras) está sendo finalizada (RASO; MELLO, em preparação) e conta com duas gravações de atendimentos médicos, análogas às gravações do C-ORAL-ESQ.

A comparabilidade com os corpora C-ORAL se dá, em primeiro lugar, pelo fato de representarem a fala espontânea, ou seja, a fala não programada, que é planejada durante a sua execução (NENCIONI, 1983). Tão importante quanto isso, todos os corpora C-ORAL são segmentados a partir do critério prosódico-pragmático inspirado pela Language into Act Theory (CRESTI, 2000) (cf. seção 3.1).

A comparabilidade com corpora de pacientes com esquizofrenia e de pessoas sem transtornos mentais é um dos pontos fortes do C-ORAL-ESQ, pois permitirá investigar quais das manifestações linguísticas observadas nos pacientes são específicas da doença, quais manifestações são próprias de uma língua e como se dá a interação entre essas manifestações.

### 2.3 Os pacientes

As gravações são feitas com pacientes em atendimento ambulatorial, de ambos os sexos, com idade e nível de escolaridade variados. A equipe médica do Instituto Raul Soares (ISR, FHEMIG) seleciona, exclusivamente, pacientes estabilizados, com *score* de até 19 pontos em sintomas positivos na escala PANSS (KEY; FISZBEIN; OPLER, 1987)<sup>2</sup>.

Antes de cada sessão de gravação, os pesquisadores apresentam, resumidamente, o projeto e pedem consenso para que a consulta seja registrada, especificando que as gravações poderão ser usadas exclusivamente para fins de pesquisa e didáticos. O consenso é formalizado pela assinatura de um Termo de Consentimento Livre e Esclarecido pelo participante ou por seu representante legal, bem como pela equipe de gravação.

### 2.4 Os equipamentos e a qualidade acústica das gravações

Uma das dificuldades relativas à compilação de um corpus de fala espontânea reside no fato de que suas gravações não são feitas em laboratório, dificultando a obtenção de dados que possam ser analisados no nível segmental. Nesse sentido, é preciso tomar uma série de precauções para garantir **qualidade acústica** e **validade ecológica** para os dados. No corpus C-ORAL-BRASIL, que registra situações comunicativas em que os

---

<sup>2</sup> A Escala das Síndromes Positiva e Negativa (*Positive and Negative Syndrome Scale – PANSS*) é uma escala psicométrica usada para avaliar a intensidade dos sintomas positivos e negativos em pacientes com Esquizofrenia. A PANSS é composta por uma escala de 7 perguntas relativas a sintomas positivos, 7 perguntas relativas a sintomas negativos e 16 perguntas sobre a psicopatologia geral.

participantes frequentemente se encontram em movimento, foram usados microfones de lapela conectados a um gravador de alta resolução (intermediados, quando necessário, por uma mesa de som). Os equipamentos usados garantiram liberdade de movimentação aos participantes e, ao mesmo tempo, permitiram a obtenção de gravações com uma qualidade acústica adequadas não somente para o estudo de aspectos lexicais e morfossintáticos, mas também para análises fonéticas e prosódicas.

No C-ORAL-ESQ, considerando que todas as gravações ocorrerão em consultórios, seria possível fazer uso de microfones condensadores de mesa conectados ao gravador. Todavia, tendo em vista que o corpus registrará a fala de pacientes com esquizofrenia, é fundamental que sejam usados equipamentos não invasivos. Além disso, os microfones de lapela, por serem muito discretos, interferem significativamente menos na situação comunicativa, favorecendo a validade ecológica das gravações. Por esses motivos, optou-se por usar, no C-ORAL-ESQ, dois microfones omnidirecionais sem fio Sennheiser EK100/SK100 ligados a um gravador Tascam DR-100 ou Marantz PMD660. Os microfones são afixados na parte central da gola da camisa dos participantes, para registrar as suas vozes com uma presença reduzida de ruídos de fundo e diminuir a quantidade e a intensidade de sobreposições entre vozes de falantes diferentes. Excepcionalmente, poderão ser usados microfones adicionais conectados ao gravador por meio de uma mesa de som Behringer Xenyx 1222 FX. Todas as gravações são feitas no formato estéreo e salvas em um arquivo *.wav*, com taxa de amostragem de 44.100 Hz. Na etapa de alinhamento, é criada uma cópia do arquivo de som, com taxa de amostragem de 22.050 Hz, para uso no software *Wipitch* (MARTIN, 2004), mas é mantido o arquivo original.

A qualidade acústica de cada gravação será classificada por meio de um procedimento semiautomático baseado em um *script* do *Praat* (BOERSMA; WEENINK, 2011) desenvolvido para a avaliação das gravações do C-ORAL-BRASIL II (RASO; MELLO, em preparação). O *script* (VIEIRA; RASO; OLIVEIRA, em preparação) se baseia nos seguintes parâmetros:

- Acurácia no cálculo da curva de  $f_0$  dos segmentos vozeados;
- Acurácia no cálculo dos formantes das vogais;
- Relação entre sinal e ruído de fundo;
- Percentual de sobreposição entre enunciados.

À exemplo do C-ORAL-BRASIL, serão selecionadas para compor o corpus somente as gravações que se encaixarem nas faixas descritas pela tabela a seguir, com preferência pela faixa A.

**Quadro 1** – Faixas de qualidade acústica, adaptado a partir de Raso (2012a, p. 74)

Faixa	Descrição
A	Ótima resposta dos microfones. Quase toda a gravação possui qualidade apropriada para quase todo tipo de análise fonética. Quase nenhuma sobreposição. Quase nenhum ruído de fundo. Computação da f0 possível em (quase) todo o arquivo.
AB	Ótima resposta dos microfones. Grande parte da gravação possui qualidade apropriada para quase todo tipo de análise fonética. Poucas sobreposições. Pouco ruído de fundo. Computação da f0 possível em (quase) todo o arquivo.
B	Boa ou média resposta dos microfones. Muitas partes da gravação possuem qualidade apropriada para análise fonética. Computação da f0 possível em grande parte do arquivo. Não há excesso de sobreposições e de ruído de fundo.
BC	Média resposta dos microfones. Computação da f0 possível em pelo menos 60% do arquivo. Mesmo quando não é possível a computação de f0, a gravação continua clara para a escuta.
C	Qualidade baixa. Média ou baixa resposta dos microfones. Computação da f0 possível em pelo menos 60% do arquivo. Algumas partes da gravação podem não ser claras para a escuta.

## 2.5 Transcrição e segmentação

Será feita uma transcrição semiortográfica de acordo com os critérios estabelecidos para o C-ORAL-BRASIL (MELLO et al., 2012), que garantem a legibilidade dos textos ao mesmo tempo em que registram fenômenos em processo de gramaticalização no português brasileiro. Alguns dos fenômenos preservados são *aférese* (estava > tava), *rotacismo* (bicicleta > bicireta), *redução pronominal* (ela > ea, ele > e', etc.) e *perda da marcação de plural* (as meninas > as menina). A adoção desses critérios possibilita que os textos sejam posteriormente submetidos à anotação léxico-morfossintática com o *parser* Palavras (BICK, 2012), adaptado para o português brasileiro para o C-ORAL-BRASIL.

A escolha do critério a ser usado para segmentar a fala em enunciados<sup>3</sup> é uma das principais decisões metodológicas para a compilação de um corpus oral, pois (a) exprime uma concepção sobre como o texto oral é organizado e, conseqüentemente, (b) condiciona em alguma medida a maneira pela qual os dados serão interpretados. No C-ORAL-ESQ, será feita uma segmentação de base prosódica e pragmática inspirada pela *Language into Act Theory* (CRESTI, 2000; MONEGLIA; RASO, 2014; RASO, 2012b), seguindo os critérios adotados pelo C-ORAL-BRASIL.

O quadro sinóptico a seguir reporta os símbolos que serão usados na segmentação dos enunciados, com exemplos de uso no C-ORAL-BRASIL e a indicação de seus arquivos de áudio. Ao longo de todo esse trabalho, os arquivos de áudio serão referenciados

<sup>3</sup> A presente seção não adota a distinção terminológica entre *enunciados* e *stanzas*, própria da *Language into Act Theory*, introduzida na seção 3.1. Assim, *enunciado* deve ser entendido como *unidade de referência da fala*, em oposição a *frase*, *unidade de referência da escrita*.



pela sigla [ex-*n*], em que *n* indica o número do exemplo.

**Quadro 2** – Símbolos usados na segmentação prosódica do C-ORAL-ESQ

Símbolo	Valor	Descrição	Exemplos
//	Fronteira prosódica terminal	Marca fronteiras de <i>enunciados</i> , entendidos como unidades com autonomia prosódica que veiculam um ato de fala (AUSTIN, 1962; CRESTI, 2000)	Exemplo 1 [ex-1] *ECR: aí tá parado // de vez em quando mexe aí mas pouco //
/	Fronteira prosódica não terminal	Marca a fronteira entre unidades tonais do mesmo enunciado	Exemplo 2 [ex-2] *SIL: porque enquanto descansa / carrega pedra //
[/n]	Retração ( <i>retracting</i> )	Indica que o falante desconsiderou uma quantidade de palavras produzidas por ele e retomou o enunciado anteriormente a essas palavras	Exemplo 3 [ex-3] *BRU: não / eu vou beber &dire [/4] eu vou beber pouco / né //
+	Enunciado interrompido	Sinaliza que o falante abandonou o enunciado que estava produzindo	Exemplo 4 [ex-4] *JHP: eu vou botar + aqui tem um [/1] um default dele / que tá / a quinhentos quilobits por segundo //

Para compreender a opção por um critério prosódico-pragmático, é pertinente que se faça uma breve análise de um conjunto de sequências extraídas do corpus C-ORAL-BRASIL.

Primeiramente, consideremos o exemplo (5), para o qual é apresentado uma transcrição sem nenhuma segmentação. No plano sintático, (5) poderia ser analisado como um período composto formado por três orações coordenadas, sendo que a primeira delas apresenta um adjunto adverbial. De acordo com um critério de segmentação sintático, o exemplo poderia ser representado como um único enunciado, como em (5').

Exemplo 5:

na hora do banho ela ia pra casa da outra avó e eu tomava meu banho e ia pra casa da outra avó também

Exemplo 5':

Na hora do banho, ela ia pra casa da outra avó e eu tomava meu banho e ia pra casa da outra avó também.

Ao ouvir o áudio associado ao exemplo [ex-5], no entanto, não restam dúvidas de

que “na hora do banho ela ia pra casa da outra avó” [ex-5a] é prosodicamente autônomo com relação ao conteúdo subsequente [ex-5b]. Por esse motivo, parece mais adequado representar essa sequência como um enunciado independente. Dentro dele, há também uma fronteira prosódica entre “na hora do banho” e “ela ia pra casa da outra avó” [ex-5c, ex-5d]. Todavia, essa fronteira, ao invés de conferir autonomia às partes, sinaliza que elas pertencem à mesma estrutura prosódica. Assim, “na hora do banho” e “ela ia pra casa da outra avó” são unidades tonais diferentes do mesmo enunciado. Entre a segunda e a terceira orações do exemplo (“e eu tomava meu banho” e “e ia pra casa da outra avó também”), também há uma fronteira prosódica que sinaliza que essas orações compõem juntas um enunciado [ex-5e, ex-5f].

Na transcrição a seguir, já no formato do C-ORAL-ESQ, as fronteiras terminais (que conferem autonomia a uma sequência linguística indicando fim de enunciado) são marcadas por ‘//’. Para as fronteiras não terminais (que indicam os limites das unidades tonais de um enunciado), é usado o símbolo ‘/’. A sigla de três letras seguida de asterisco indica o falante.

Exemplo 5:

\*DFL: na hora do banho / ela ia pra casa da outra avó // e eu tomava meu banho / e ia pra casa da outra avó também //

Façamos agora uma análise de (6) e (7).

Exemplo 6:

\*BAL: tá saindo de uma garrafinha que tem um bico muito pequeno

Exemplo 7:

\*BAL: cê tá com um jarro d’água que tem uma espessura assim

Do ponto de vista sintático, seria possível dizer que, em ambos os casos, uma oração relativa modifica um dos argumentos do verbo da oração principal. Todavia, mais uma vez, a prosódia mostra-se fundamental para interpretar a estrutura da fala: enquanto (6) parece formado por um único enunciado [ex-6], o exemplo (7) apresenta duas sequências autônomas prosodicamente, ou seja, dois enunciados distintos [ex-7, ex-7a, ex-7b].

Exemplo 6’:

\*BAL: tá saindo de uma garrafinha que tem um bico muito pequeno //

Exemplo 7’:

\*BAL: cê tá com um jarro d’água // que tem uma espessura assim //

Há uma ampla bibliografia que discute os parâmetros envolvidos na realização de

cada tipo de fronteira prosódica. Apesar de não haver consenso entre os autores, alguns dos parâmetros frequentemente mencionados são: parâmetros de  $f_0$ , de intensidade, de ritmo, pausas e mudanças na qualidade vocálica (MITTMANN; BARBOSA, 2016). Dentre eles, as pausas são um dos parâmetros mais estudados, visto que elas produzem necessariamente uma quebra prosódica no fluxo discursivo. Todavia, nem a *presença*, nem a *duração* de uma pausa determinam se a fronteira será terminal ou não terminal (RASO; MITTMANN; MENDES, 2015). Isso será ilustrado com os exemplos (8) e (9).

Em (8), há um silêncio de 1,5 segundos entre duas unidades tonais do mesmo enunciado (“por causa do excesso” e “de doce”) [ex-8]. Já em (9), há uma pausa de 0,88 segundos separando o penúltimo enunciado do último (“agora não sei” e “se vai fazer ou se desanimou”) [ex-9]. No mesmo exemplo, os dois primeiros enunciados são separados por uma quebra terminal sem pausa.

Exemplo 8: CAB fala sobre suas preferências alimentícias.

\*CAB: também não gosto muito de doce de leite / por causa do excesso / de doce // Exemplo 9: ECR discute sobre a construção de uma casa.

\*ECR: é / e diz ele que tem vontade de fazer lá embaixo // agora eu não sei // se vai fazer ou se desanimou //

A discussão feita até aqui indica que a adoção de um critério prosódico permite representar adequadamente a maneira pela qual a fala espontânea é segmentada. Mais do que um mecanismo de desambiguação sintática, a prosódia constitui a principal marca formal para estruturar o material linguístico. Os exemplos também mostram que a autonomia prosódica de uma sequência não depende da sintaxe: ainda que sejam frequentes os enunciados com estrutura verbal saturada, muitos são formados apenas por sintagmas isolados, em alguns casos sem verbos. O exemplo (10) ilustra essa discussão.

Exemplo 10: FLA e REN conversam no supermercado.

\*FLA: esses dias / a gente tava procurando toalha / pa cobrir o / carneirinho // **que a gente disseca** // <que ele> [/1] ele +

\*REN: <o quê> //

\*FLA: na aula de anatomia //

\*REN: <ah> //

\*FLA: <porque vai ressecando> / né //

\*REN: ah / <entendi> //

\*FLA: <uhn / pão> **de alho** // Nossa / <eu quero tudo de> comida também / <né> //

\*REN: <delícia / né> // <é> / hoje cê tá faminta / né / filha //

Nesse diálogo, há vários enunciados sem verbo (“o quê //”, “na aula de anatomia

//” e “delícia / né //”) [ex-10a, ex-10b, ex-10c], um dos quais se constitui somente de uma interjeição (“ah //”) [ex-10d]. Também há enunciados em que o único verbo está fora da oração principal (“que a gente disseca //” e “porque vai ressecando / né //”) [ex-10e, ex-10f].

Sequências como essas são amplamente atestadas em corpora de fala espontânea de diferentes línguas<sup>4</sup> e parecem atender aos pressupostos básicos de estruturação da fala, que não necessariamente coincidem com os da escrita. Em face disso, é importante que o critério de segmentação de um corpus seja formulado de modo a atribuir a essas sequências o mesmo valor conferido aos enunciados com verbo.

Visto que a autonomia prosódica não depende da presença de estrutura oracional (e, portanto, não necessariamente realiza predicação), há de se perguntar qual é a função básica dessas unidades na comunicação. Como será mostrado em 3.1, há uma correspondência sistemática entre unidades autônomas prosodicamente e unidades de ação (atos de fala) (AUSTIN, 1962). Nesse sentido, o critério pragmático para que uma sequência linguística possa ser autônoma é o seu valor enquanto ação que o falante realiza sobre o interlocutor. Os enunciados sem verbo do exemplo (10) atendem a esse critério.

Antes de prosseguir, é necessário ressaltar que o critério para identificar quebras prosódicas durante o processo de transcrição e segmentação das gravações é de natureza perceptual. Por isso, a equipe envolvida nas diversas fases de tratamento dos dados do C-ORAL-ESQ passou por longas oficinas de treinamento, para aumentar a confiabilidade de suas decisões. Adicionalmente, serão feitos testes *kappa* de Fleiss (1971) para validar a anotação de quebras terminais e não terminais, assim como no C-ORAL-BRASIL (MELLO et al., 2012).

## 2.6 O alinhamento texto-som

O alinhamento texto-som é o processo de associar cada enunciado de uma transcrição à porção de áudio correspondente, em um programa específico, permitindo que o pesquisador possa ouvir um enunciado ao clicar sobre o texto correspondente.

Em corpora com transcrições não alinhadas, o pesquisador deve procurar manualmente, no arquivo de áudio, os enunciados desejados. Compreensivelmente, é frequente que opte por fazer todas as suas análises baseando-se somente na transcrição dos textos. Um dos problemas associados a isso é o fato de uma transcrição estar sempre sujeita a in-

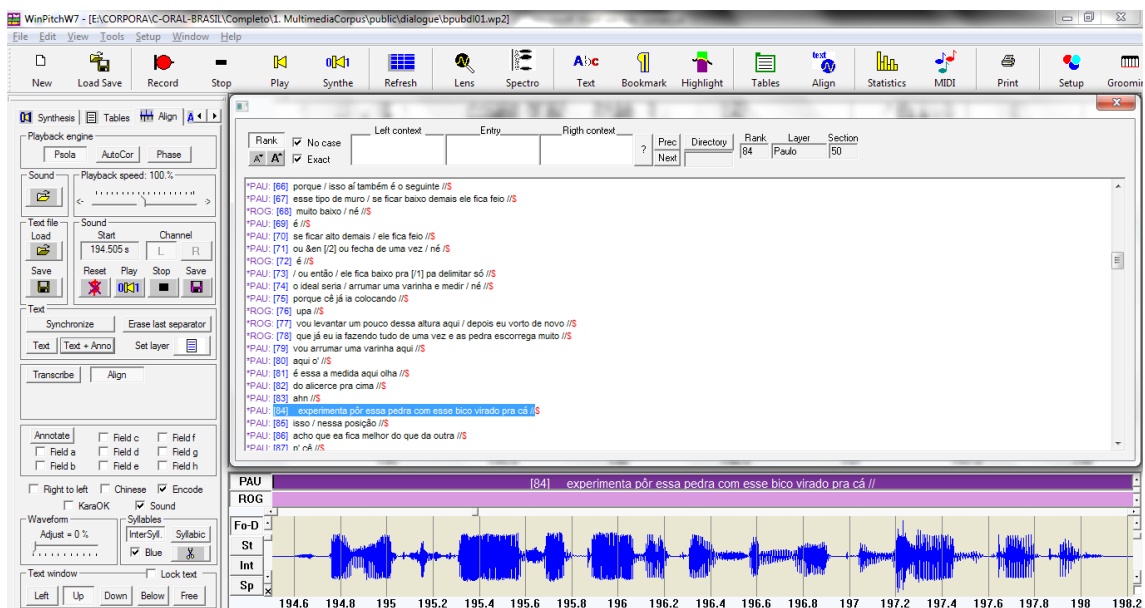
---

<sup>4</sup> Levantamentos no corpus LABLITA dão conta de que 38,1% dos enunciados é formado por SNs, SPs, SAdj ou interjeições, sem verbos (CRESTI; GRAMIGNI, 2004). No C-ORAL-BRASIL, 22,12% dos enunciados de textos monológicos e 29,5% dos enunciados de textos dialógicos também não possuem formas verbais (RASO; MITTMANN, 2012). A *Longman Grammar* registra 38% de enunciados para o mesmo fenômeno (CRESTI; GRAMIGNI, 2004). Em uma análise de três textos do *Corpus of Spoken Israeli Hebrew* (CoSIH), Izre’el (2011) reporta que cerca de 50% das unidades tonais dos enunciados não possui estrutura oracional.

interpretações – e eventuais enganos – de quem a produziu, de modo que o acesso direto ao material acústico pode ser decisivo para uma análise adequada. Mais importante ainda, convém lembrar que uma transcrição é uma representação de uma mensagem verbal que retira o que lhe é mais próprio: a prosódia. Por mais que muitos sistemas de transcrição tentem reproduzir, de alguma maneira, certas propriedades prosódicas dos enunciados, há de se reconhecer que essas tentativas são sempre simplificações extremamente limitadas da informação original. Assim, ainda que sejam úteis para certos fins, não substituem a riqueza da informação acústica. Por esses motivos, atualmente, mais do que uma comodidade, o alinhamento é considerado um recurso indispensável para a pesquisa em corpora de fala.

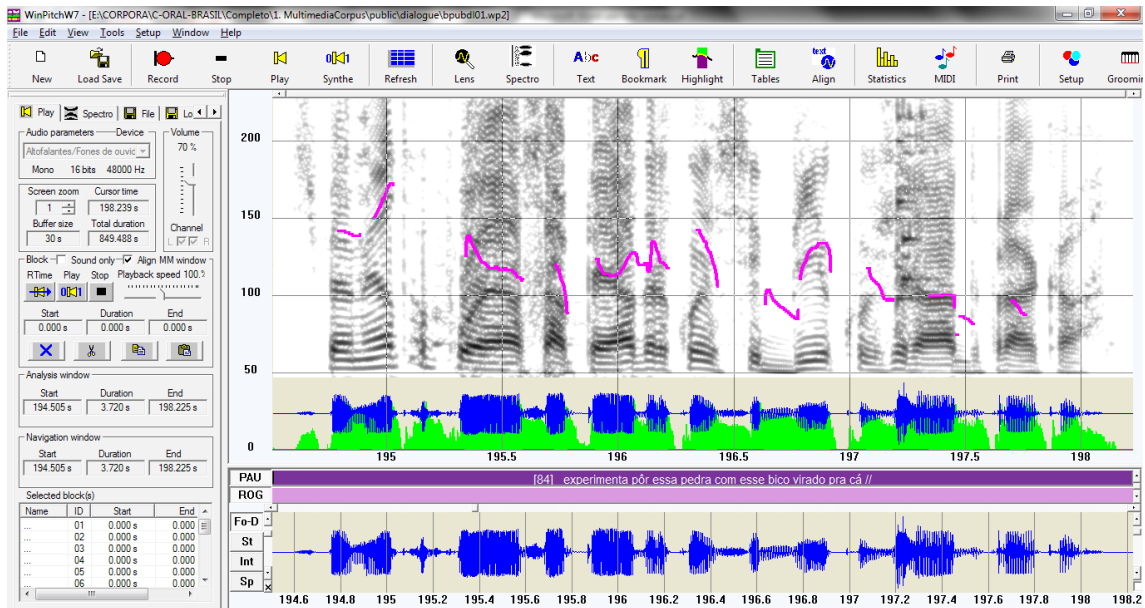
Seguindo o exemplo dos corpora C-ORAL, a unidade básica de alinhamento do C-ORAL-ESQ será o enunciado, delimitado por quebras prosódicas terminais. Para tanto, será usado o *WinPitch* (MARTIN, 2004), um programa desenvolvido para alinhamento e análise acústica de corpora orais. Nesse programa, ao clicar sobre um ou mais enunciados de uma transcrição, o pesquisador pode ouvi-los e também visualizar o oscilograma, o espectrograma e o curso de  $f_0$  desses enunciados. Além disso, o *WinPitch* permite que uma quantidade qualquer de enunciados alinhados seja exportada ou importada no formato .TextGrid, compatível com o *Praat* (BOERSMA; WEENINK, 2011), preservando a transcrição original. As imagens a seguir mostram telas do programa *WinPitch* durante uma pesquisa com um texto do C-ORAL-BRASIL (RASO; MELLO, 2012).

**Figura 1** – Tela do *WinPitch* mostrando a seleção de um enunciado de um texto do C-ORAL-BRASIL.



**Figura 2** – Tela do *WinPitch* com formantes, curva de  $f_0$ ,

oscilograma e intensidade de um enunciado de um texto do C-ORAL-BRASIL.



## 2.7 Equipes de trabalho

O C-ORAL-ESQ congrega grupos de trabalho de diferentes instituições, divididos em duas grandes equipes:

- I. Equipe com competência em estudos médicos relativos à esquizofrenia e no tratamento psiquiátrico de pacientes esquizofrênicos, composta por:
  - a. Grupo do Departamento de Morfologia do ICB/UFMG, sob a coordenação de João Salgado;
  - b. Grupo de Pesquisas Transtornos Mentais e Cognição (CNPq), sob a coordenação de João Salgado;
  - c. Instituto Raul Soares - IRS / Fundação Hospitalar do Estado de Minas Gerais - FHEMIG, sob a coordenação de Marco Antônio de Rezende Andrade;
  - d. Grupo da Duke University Medical Center (Durham, EUA), sob a coordenação de Richard Keefe.
- II. Equipe com competência na produção e estudo de corpora de fala de pessoas com e sem transtornos mentais, composta por:
  - a. LEEL (Laboratório de Estudos Empíricos e Experimentais da Linguagem, FALE/UFMG), sob a coordenação de Tommaso Raso;
  - b. Grupo de Pesquisas Incógnito (CNPq), sob a coordenação de Heliana Mello e Tommaso Raso;
  - c. Grupo da Università di Napoli Federico II, sob a coordenação de Francesca Dovetto.

### 3 A ESTRUTURA INFORMACIONAL DA FALA DE PACIENTES ITALIANOS

Essa seção dedica-se a reportar estudos de Cresti, Dovetto e Rocha (CRESTI; DOVETTO; ROCHA, 2015; DOVETTO; CRESTI; ROCHA, 2015) sobre a estrutura informacional da fala de pacientes com esquizofrenia, a partir da *Language into Act Theory* (CRESTI, 2000; MONEGLIA; RASO, 2014; RASO, 2012b). Por meio dessa exposição, busca-se destacar uma das possibilidades de uso do C-ORAL-ESQ para estudar uma dimensão da prosódia ainda pouco explorada na literatura médica e linguística sobre o tema. Os conceitos chave da L-AcT serão apresentados em (3.1), ao passo que a sua aplicação para a compreensão da fala de pacientes será mostrada em (3.2).

#### 3.1 Fundamentos da *Language into Act Theory* (L-AcT)

A L-AcT (CRESTI, 2000; MONEGLIA; RASO, 2014; RASO, 2012b) é uma teoria desenvolvida ao longo de mais de 30 anos de trabalho com corpora de fala espontânea. O seu objetivo é compreender a estrutura informacional da fala, explorando a relação existente entre o domínio prosódico e o plano pragmático da comunicação. A teoria fundamenta-se na observação de que as unidades com autonomia prosódica identificáveis no fluxo discursivo correspondem a atos de fala (AUSTIN, 1962) que o falante realiza sobre o interlocutor. Nesse sentido, a L-AcT identifica duas unidades básicas de estruturação da fala: o enunciado (sequência linguística prosodicamente autônoma que veicula uma única ilocução) e a *stanza* (sequência linguística com autonomia prosódica que realiza duas ou mais ilocuções). Os enunciados e as *stanzas* são *sequências terminais* (RASO et al., no prelo), dotadas de autonomia prosódica e pragmática (CRESTI, 2000).

No exemplo (11), observam-se 8 enunciados e 1 *stanza* (em negrito), separados por quebras terminais (//).

Exemplo 11: KAT e SIL fazem faxina na casa em que trabalham.

\*KAT: o quê //

\*SIL: copos // copos de Urano / que tem aí //

\*KAT: copos de quê //

\*SIL: Urano //

\*KAT: Urano //

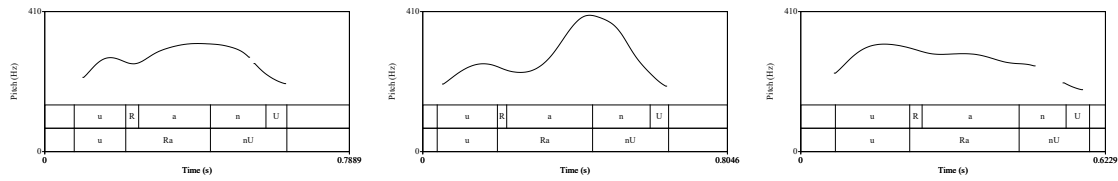
\*SIL: é // Urano // Urano // **é um negócio que tem / que es fazem na Itália / que custa caríssimo //**

Alguns enunciados apresentam somente uma unidade tonal e, por isso, são chamados de *enunciados simples*. Esse é o caso de “o quê //” e “copos //” [ex-11a e ex-11b], ao início de (11), que podem ser interpretados pragmaticamente como ilocuções de pergunta parcial e de resposta. Outros possuem uma unidade tonal ilocucionária e uma ou mais unidades tonais não ilocucionárias separadas por quebras não terminais (/), sendo denominados *enunciados complexos*. A sequência “copos de Urano / que tem aí //” é um enunciado complexo com uma unidade ilocucionária (“copos de Urano /”) e uma não ilocucionária (“que tem aí //”), dependente prosodicamente e funcionalmente da anterior. A unidade ilocucionária de um enunciado chama-se *comentário*. Nos enunciados complexos, o comentário é integrado por unidades não ilocucionárias, com as quais forma um *padrão informacional*.

As *stanzas* também possuem, necessariamente, mais de uma unidade tonal. São formadas por unidades ilocucionárias sem autonomia prosódica (*comentários ligados*) e por um comentário. A *stanza* de (11) tem dois comentários ligados (“é um negócio que tem /” e “que es fazem na Itália /”) e um comentário (“que custa caríssimo //”) [ex-10c a ex-10f]. Adicionalmente, as *stanzas* podem ter unidades não ilocucionárias que formam subpadrões informacionais com os comentários ligados. *Stanzas* expressam o desenvolvimento de um pensamento em curso, em que o falante se concentra não tanto na realização de uma ação, mas sobretudo na elaboração textual (CRESTI, 2000). Apesar de serem encontradas em qualquer tipo de interação verbal, são especialmente frequentes e complexas em monólogos.

O valor ilocucionário de um enunciado e de uma *stanza* é codificado linguisticamente por meio das propriedades prosódicas de um número reduzido de sílabas do comentário e dos comentários ligados (CRESTI, 2018; RASO; ROCHA, 2015; ROCHA, 2016). Tais sílabas correspondem sempre a uma proeminência prosódica, constituindo o *foco informacional* do comentário e dos comentários ligados (CRESTI, 2011). Voltando a (11), observam-se três enunciados formados por uma só palavra (“Urano”) que, apesar de possuírem a mesma estrutura morfossintática, realizam ilocuções diferentes [ex-11g a ex-11i]. O elemento que garante a variação ilocucionária é a prosódia. Essas sequências também reforçam que enunciados e *stanzas* não necessariamente são formados por frases, como discutido na seção 2.6.





**Figura 3** – Curso de  $f_0$  e segmentação em fones dos enunciados “Urano //”, na ordem em que aparecem na gravação, com ilocuções de confirmação (esquerda), incredulidade (centro) e conclusão (direita).

As unidades não ilocucionárias são realizadas para guiar a interpretação da ilocução do padrão de que fazem parte, seja em um enunciado, seja em *stanzas*. A partir da análise de corpora de fala espontânea, a L-AcT identificou diferentes tipos de unidades não ilocucionárias, divididos em dois grandes grupos:

- *unidades textuais* (tópico, apêndice de tópico, apêndice de comentário, parentético, introdutor locutivo): integram o texto da sequência terminal;<sup>5</sup>
- *unidades dialógicas* (fático, incipitário, expressivo, alocutivo, conativo, conector discursivo): regulam a interação com os interlocutores<sup>6</sup>.

Uma das unidades textuais é o *apêndice de comentário*, que se localiza após o comentário a que se refere, integrando-o com informações adicionais ou retomando informações já ditas. No enunciado complexo de (11), “que tem aí //” trata-se de um apêndice de comentário. Outra unidade textual é o *tópico*, cuja função é estabelecer um domínio cognitivo para a ilocução (e, portanto, vem sempre antes de uma unidade ilocucionária). Em (12), a seguir, “a orientadora /” é um tópico. Esse exemplo também contém um *fático* (“então /”), unidade de auxílio dialógico com a função de manter o canal de comunicação aberto e posição livre na sequência terminal [ex-12].

Exemplo 12: SHE relata problemas da carreira docente.

\*SHE: então / a orientadora / ela nũ quer fazer o papel da coordenadora //

As funções não ilocucionárias também são codificadas prosodicamente. Todavia, a única unidade não ilocucionária que, assim como o comentário e os comentários ligados, possui foco informacional é o tópico.<sup>7</sup> Nas demais, a função é assinalada pela totalidade da unidade e não por uma porção prosodicamente saliente.

<sup>5</sup> O comentário e o comentário ligado também são unidades textuais, mas não estão nessa lista por terem valor ilocucionário.

<sup>6</sup> Para uma descrição detalhada das unidades dialógicas em português brasileiro veja Raso (2014).

<sup>7</sup> Para uma descrição do tópico em diferentes em português brasileiro, italiano, inglês norte-americano, veja-se Raso, Mittmann e Cavalcante (2017).

Até agora, foi observado que enunciados e *stanzas* possuem uma ou mais unidades ilocucionárias e podem apresentar unidades não ilocucionárias com funções específicas. Na comunicação cotidiana, os falantes se valem de um amplo arcabouço de funções informacionais para estruturar enunciados e *stanzas* em padrões informacionais, de modo a atingir objetivos comunicativos variados. A seguir, será visto como essas noções podem ser usadas em uma tentativa de compreender a fala de pacientes com esquizofrenia.

### **3.2 Uma aplicação da L-AcT para a análise da fala de pacientes italianos**

Nessa seção, serão apresentados sucintamente os trabalhos de Cresti, Dovetto e Rocha (CRESTI; DOVETTO; ROCHA, 2015; DOVETTO; CRESTI; ROCHA, 2015) sobre a fala espontânea de pacientes italianos com esquizofrenia, feitos a partir de dados do corpus CIPPS (DOVETTO; GEMELLI, 2012). Para esses estudos, os autores selecionaram amostras de aproximadamente 100 enunciados e *stanzas* de três dos quatro pacientes do corpus (identificados como ‘PAA’, ‘PAB’ e ‘PAD’) e fizeram observações gerais sobre sua estrutura informacional com base na L-AcT (CRESTI, 2000; MONEGLIA; RASO, 2014; RASO, 2012b). Tratam-se de pesquisas exploratórias, as quais objetivam apontar direções de investigação para compreender melhor as particularidades da fala dos pacientes, sem com isso pretender chegar a conclusões definitivas.

#### **3.2.1 Aspectos convergentes entre a fala de pacientes e de pessoas sem transtornos mentais**

O primeiro aspecto investigado foi a capacidade dos pacientes com esquizofrenia de segmentar a fala em sequências terminais. Essa habilidade diz respeito à manipulação de complexos padrões de parâmetros prosódicos para produzir fronteiras que sinalizem limites de enunciados e *stanzas* ou indiquem suas unidades internas. Trata-se de uma capacidade que pode estar comprometida, em diferentes graus e por motivos distintos, em algumas patologias, como certos tipos de síndrome de Down (SORIANELLO, 2012) e de afasia de Broca (MAROTTA, 2009).

Com relação aos dados analisados, os autores notam que os pacientes criam fronteiras prosódicas reconhecíveis como indicadoras de limites de sequências terminais e de suas unidades tonais, como em (13) a (15). O primeiro exemplo reporta três enunciados simples, cujas quebras terminais são facilmente perceptíveis mesmo por falantes de PB [ex-13 a ex-13c]. A sensação de independência dos enunciados não é dada pelas pausas entre eles, como pode ser notado em uma edição que remove as pausas [ex-13d]. Em (14)

e (15) observam-se enunciados complexos formados por uma unidade não ilocucionária seguida do comentário, separadas por uma quebra não terminal.<sup>8</sup> Nos dois enunciados, a unidade não ilocucionária não possui autonomia prosódica [ex-14a, ex-15a].

Exemplo 13: PAB relata ao médico que era perseguido por ter ideias inovadoras.

\*PAB: io tenevo le idee //<sup>COM</sup> hai capito //<sup>COM</sup> quello era il problema //<sup>COM</sup>

\*PAB: eu tinha as ideias // entendeu // esse era o problema //

%ill: [1] conclusão; [2] pedido de confirmação; [3] conclusão

Exemplo 14: O médico pergunta a PAB como foi a sua semana.

\*PAB: dopo quest'incidente /<sup>TOP</sup> tutto bene //<sup>COM</sup>

\*PAB: depois desse incidente / tudo bem //

%ill: [1] resposta

Exemplo 15: PAD diz algo interessante, mas o médico não toma notas.

\*PAD: dutto' /<sup>ALL</sup> scrivitt' coccos' 'ncopp' 'o computer //<sup>COM</sup>

\*PAD: doutor / escreve alguma coisa no computador //

%ill: [1] ordem

Esses exemplos também ilustram a capacidade de usar a prosódia para sinalizar o tipo ilocucionário do comentário, bem como o valor das unidades não ilocucionárias. (13) é formado por uma conclusão, um pedido de confirmação e uma conclusão. Em (14), observa-se uma unidade não ilocucionária textual de tópico seguida do comentário, que veicula uma resposta. O exemplo (15) apresenta uma unidade de auxílio dialógico de alocutivo seguida do comentário, com a ilocução de ordem.

Tão importante quanto isso, esses exemplos mostram casos em que os pacientes são capazes de planejar e executar a fala para que a estrutura informacional dos enunciados responda às necessidades pragmáticas da situação. Em (13), PAB relata que era perseguido por ter ideias de destaque na empresa. A ilocução de pedido de confirmação é usada para assegurar-se que o médico esteja acompanhando seu raciocínio. Em (14), o paciente usa o tópico para contextualizar a ilocução de resposta. Em (15), o paciente diz algo que considera interessante e percebe que o médico não toma notas. Assim, realiza a ordem para que o médico escreva algo em seu computador. A unidade de alocutivo, cuja função é indicar ao interlocutor que o enunciado é direcionado a ele, também é usada apropriadamente.

---

<sup>8</sup> Todos os exemplos em italiano dessa sessão apresentam dados do CIPPS (DOVETTO; GEMELLI, 2012).

### 3.2.2 Aspectos divergentes entre a fala de pacientes e de pessoas sem transtornos mentais

As capacidades citadas anteriormente, que se manifestam em níveis diferentes nos pacientes analisados, são importantes pontos em comum entre a fala de portadores de esquizofrenia e a fala de pessoas sem transtornos mentais. A seguir, serão abordadas algumas peculiaridades da fala dos pacientes. Antes de passar à exposição, convém lembrar que os estudos aqui mencionados possuem natureza exploratória, e as características que serão abordadas podem variar expressivamente de um falante para outro.

#### *Baixa variação ilocucionária e inadequação contextual*

Todos os pacientes realizaram, nas amostras estudadas, mais de um tipo ilocucionário. A variação de ilocuições expressa a capacidade de conseguir ajustar o planejamento da fala às necessidades comunicativas. As consultas médicas documentadas pelo CIPPS exprimem um contexto de acionalidade mais restrita, em que é esperada uma presença expressiva de asserções, mas também um leque de outros tipos ilocucionários. Todavia, os autores notam que há pouca variabilidade na amostra de dois dos três pacientes estudados (PAA e PAD).

Nas 100 sequências terminais de PAD, por exemplo, encontram-se quase exclusivamente asserções, as quais são produzidas com uma atitude fortemente enfática, como em (16). Nesse exemplo, há também duas ilocuições de contraste (“&n un’ marron /” e “na casa verde /”). Na transcrição, a sigla ‘COM’ indica unidades de comentário, e ‘COB’ assinala comentários ligados. O exemplo possui dois enunciados simples e duas *stanzas*.

Exemplo 16: O paciente descreve para o médico um desenho que fez em sua casa.

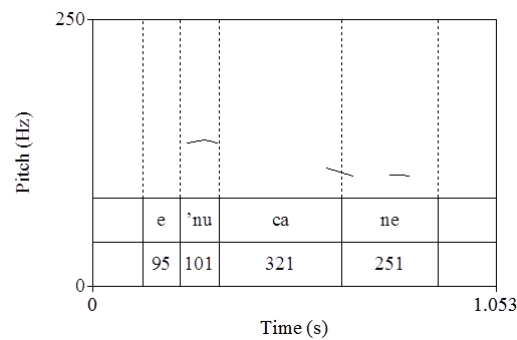
\*PAD: muntagne viola //<sup>COM</sup> viola //<sup>COM</sup> &n un’ marron /<sup>COB</sup> ‘na person’ umana //  
<sup>COM</sup> ‘n alber’ viola /<sup>COB</sup> ‘na casa verde /<sup>COB</sup> e ‘n persona umana //<sup>COM</sup> e ‘nu cane //  
COM

\*PAD: montanhas roxas // roxas // um marrom / uma pessoa humana // uma árvore  
roxa / uma casa verde / e uma pessoa humana // e um cão //

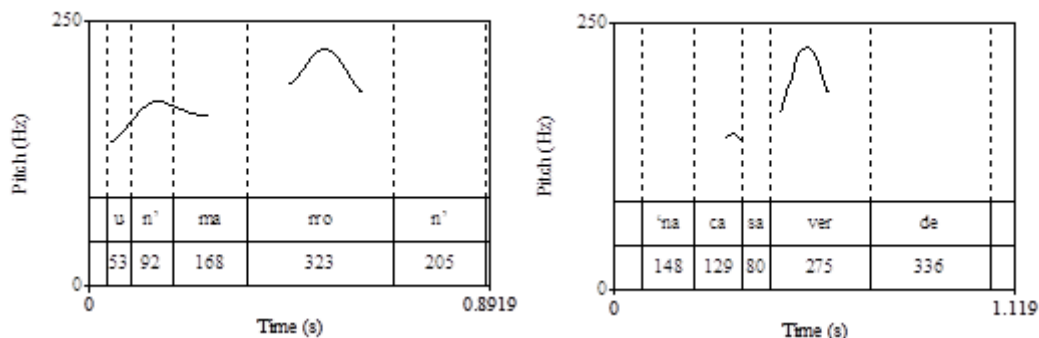
%ill: [1] asserção; [2] asserção; [3] contraste, asserção; [4] asserção, contraste, asserção; [5] asserção

Prosodicamente, essas asserções são realizadas por uma configuração ascendente-descendente de f0 alinhada à última tônica da unidade de comentário, com variação entre 80Hz e 150 Hz, aparente alongamento e voz crepitante (*creaky*) na tônica e na pos-tônica finais. Os enunciados de contraste caracterizam-se por uma configuração ascendente-descendente de f0 na tônica final do comentário, com valores altos de f0 (150Hz-

230Hz), aparente alongamento da tônica e postônica finais e voz crepitante especialmente na postônica final. A Figura 4 exibe a curva de f0 de “e ‘nu cane //”, com uma asserção de (16) [ex-16a]. Na Figura 5, podem ser observadas as curvas de f0 das duas ilocuções de contraste [ex-16b, ex-16c].



**Figura 4** – Curso de f0, divisão em sílabas e duração (ms) da última asserção de (16): “e ‘nu cane //”. O curso de f0 da tônica e postônica finais não aparecem por completo na imagem em função da voz crepitante.



**Figura 5** – Cursos de f0, divisão em sílabas e duração (ms) das ilocuções de contraste de (16): “un’ marron’ //” (esquerda) e “na casa verde //” (direita).

Analisados isoladamente, os perfis prosódicos das asserções não são anômalos. Por outro lado, chama atenção a atitude fortemente enfática com a qual PAD realiza não somente as asserções de (16), mas grande parte das sequências terminais da amostra estudada. O uso excessivo de tal atitude motiva a percepção de inadequação dos enunciados à situação comunicativa, bem como de monotonia prosódica ainda que o paciente produza amplas excursões de frequência fundamental nessas sequências.

### *Estrutura informacional pouco articulada*

Os exemplos observados até aqui mostram que os pacientes são capazes de produzir enunciados simples e complexos, bem como *stanzas*. Antes de prosseguir, cumpre lembrar que enunciados complexos e *stanzas* são formados por mais de uma unidade tonal, sendo que uma delas é necessariamente ilocucionária (o comentário). Além dela, podem conter:

- unidades ilocucionárias adicionais, como as *stanzas* de (16);
- unidades não ilocucionárias com função textual, como (14);
- unidades não ilocucionárias de auxílio dialógico, como (15);
- combinações entre as possibilidades anteriores.

Apesar de Cresti, Dovetto e Rocha não terem se focado em uma análise quantitativa dos dados, observaram que alguns pacientes realizam um número restrito de unidades não ilocucionárias (especialmente PAD). Com isso, seus enunciados expressam uma estrutura informacional pouco articulada se comparada à de falantes sem transtornos mentais. O exemplo (16), exibido anteriormente, ilustra essa situação: trata-se de uma sequência de 4 enunciados e 2 *stanzas* sem unidades textuais ou de auxílio dialógico.

Outra característica marcante na fala dos pacientes estudados é a pequena variação de *tipos* de unidades não ilocucionárias usadas por cada um. Dentre o amplo repertório de unidades textuais e de auxílio dialógico disponíveis, cada paciente tende a realizar somente alguns deles, e não outros. PAD, por exemplo, produz poucos tópicos, que, na fala de pessoas sem transtornos mentais, é a unidade textual não ilocucionária mais frequente (CRESTI, 2000; SIGNORINI, 2005). Por outro lado, realiza um número expressivo de apêndices de comentário. Já PAB, em seus 100 primeiros enunciados, produz muitos tópicos, mas nenhum apêndice de comentário. Outras unidades de auxílio dialógico bastante comuns, como o fático (cf. seção 3.1), ocorrem em número muito reduzido nas amostras dos três pacientes.

### *Ecolalia e palilalia*

A ecolalia, considerada um fenômeno típico na fala de pacientes com esquizofrenia, é “a repetição da última ou das últimas palavras que o entrevistador (ou alguém no ambiente) falou ou dirigiu ao paciente”, realizada “sem planejamento ou controle” (DALGALARRONDO, 2009, p. 239). A palilalia, por sua vez, seria a repetição das últimas palavras que o próprio paciente produziu. Cresti, Dovetto e Rocha notam que a ecolalia e a palilalia são sempre realizadas em unidades tonais a parte, sem valor ilocucionário, logo após o enunciado que retomam. Possuem perfil com valores baixos de f0 e intensidade. Em (17) [ex-17], que traz enunciados de PAA, “la volta scorsa //” é uma ecolalia (sigla ‘ECO’) da fala do médico [ex-17a], ao passo que o fragmento “&different //” é uma palilalia (sigla ‘PAL’) [ex-17b].

Exemplo 17: O médico pergunta ao paciente PAA o que ele escreveu durante a semana.

\*MED: la volta scorsa /<sup>TOP</sup> che cosa /<sup>SCA</sup> scrivesti //<sup>COM</sup>

\*PAA: la volta scorsa //<sup>ECO</sup> in quella sensorialità //<sup>COM</sup> cioè /<sup>PHA</sup> percezione //<sup>COM</sup>  
specificare tutte le differenze //<sup>COM</sup> &differen //<sup>PAL</sup>

\*MED: da última vez / que coisa / você escreveu //

\*PAA: da última vez // naquela sensorialidade // quer dizer / percepção // especificar todas as diferenças // &diferen //

Os autores enfatizam que a palilalia e a ecolalia não devem ser analisadas como enunciados novos que retomam as palavras do anterior para fins retóricos, mas sim como unidades sem valor informacional, com um perfil prosódico idiossincrático e não tipificado. Visto que esses fenômenos constituem repetições automáticas e voltadas para si mesmo, colocam em evidência “uma distinção entre ‘falar’, que depende de uma capacidade expressiva, e ‘comunicar’, que depende de uma capacidade afetiva e pragmática” (DOVETTO; CRESTI; ROCHA, 2015).

#### 4 Observações finais

Os trabalhos de Cresti, Dovetto e Rocha apontam para o fato de que a elaboração informacional restrita, em conjunto com a baixa variação ilocucionária e a inadequação situacional de alguns enunciados – características que não são compartilhadas por todos os pacientes e podem se apresentar em medidas diferentes entre eles – são pontos importantes para compreender os impactos da esquizofrenia sobre a fala dos pacientes. De acordo com os autores, a repetição excessiva de um conjunto limitado de módulos prosódicos idiossincráticos na forma e no uso resulta em padrões pouco comunicativos, que chamam a atenção do ouvinte e causam uma sensação de monotonia prosódica. Apesar de ainda muito embrionários, esses trabalhos parecem indicar que a análise da estrutura informacional da fala pode contribuir para compreender aspectos relevantes – e pouco estudados – da manifestação linguística de pacientes com esquizofrenia. Nesse contexto, o corpus C-ORAL-ESQ apresenta-se como um recurso importante possibilitando não somente esse tipo de estudo, mas também outros tipos de investigações sobre a fala espontânea de pacientes com esquizofrenia.

#### Referências

ALPERT, M. et al. A comparison of clinical ratings with vocal acoustic measures of flat affect and alogia. *Journal of Psychiatric Research*, v. 36, n. 5, p. 347-353, 2002.

- AUSTIN, J. *How to do things with words*. Oxford: Oxford University Press, 1962.
- BACH, D. R. et al. Altered lateralisation of emotional prosody processing in schizophrenia. *Schizophrenia Research*, v. 110, n. 1-3, p. 180-187, 2009.
- BEDWELL, J. S. et al. Speech prosody abnormalities and specific dimensional schizotypy features are relationships limited to male participants? *Journal of Nervous and Mental Disease*, v. 202, n. 10, p. 745-751, 2014.
- BICK, E. A anotação gramatical do C-ORAL-BRASIL. In: RASO, T.; MELLO, H. (Eds.). *C-ORAL-BRASIL I: Corpus de referência do português brasileiro falado informal*. 1. ed. Belo Horizonte: Editora UFMG, 2012. p. 223-254.
- BOERSMA, P.; WEENINK, D. *Praat: doing phonetics by computer*, 2011. Disponível em: <http://www.praat.org/>. Acesso em: 1 de junho de 2019.
- COHEN, A. S.; KIM, Y.; NAJOLIA, G. M. Psychiatric symptom versus neurocognitive correlates of diminished expressivity in schizophrenia and mood disorders. *Schizophrenia Research*, v. 146, n. 1-3, p. 249-253, 2013.
- CRESTI, E. *Corpus di Italiano parlato*. Firenze: Accademia della Crusca, 2000.
- CRESTI, E. The definition of focus in Language into Act Theory (LAcT). In: MELLO, H.; PANUNZI, A.; RASO, T. (Eds.). *Pragmatics and Prosody: Illocution Modality, Attitude, Information Patterning and Speech Annotation*. Firenze: Firenze University Press, 2011.
- CRESTI, E. The illocution-prosody relationship and the Information Pattern in spontaneous speech according to the Language into Act Theory (L-AcT). *Linguistik online*, v. 88, p. 33-62, 2018.
- CRESTI, E.; GRAMIGNI, P. *Per una linguistica corpus based dell'Italiano parlato: le unità di riferimento*. In: ALBANO LEONI, F. et al. (Eds.). *Il parlato italiano*, Napoli, 13-15 febbraio 2003. *Anais...* 2004.
- CRESTI, E.; MONEGLIA, M. (EDS.). *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam / Philadelphia: John Benjamins Publishing Company, 2005.
- CRESTI, E.; DOVETTO, F. M.; ROCHA, B. Schizophrenia and prosody. First investigations. Em: MANFREDI, C. (Org.), *Models and analysis of vocal emissions for biomedical applications - 9th international workshop - September 2-4, 2015*. Firenze: Firenze University Press, 2015.
- DALGALARRONDO, P. *Psicopatologia e semiologia dos transtornos mentais*. Porto Alegre: Artmed Editora, 2009.
- DOVETTO, F. M. Schizofrenia e deissi. *Studi e saggi linguistici*, v. 2, p. 101-132, 2014.
- DOVETTO, F. M.; GEMELLI, M. *Il parlar matto*. Schizofrenia tra fenomenologia e linguistica. Il corpus CIPPS. Roma: Aracne, 2012.
- DOVETTO, F. M.; CRESTI, E.; ROCHA, B. Schizofrenia tra prosodia e lessico. Prime



- analisi, *Studi Italiani di Linguistica Teorica e Applicata*, v. 3, p. 486-507, 2015.
- FLEISS, J. L. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, v. 76, n. 5, p. 378-382, 1971.
- FROMKIN, V. A. A linguist looks at “a linguist looks at ‘schizophrenic language’”. *Brain and Language*, v. 2, n. C, p. 498-503, 1975.
- HOEKERT, M. et al. Impaired recognition and expression of emotional prosody in schizophrenia: Review and meta-analysis. *Schizophrenia Research*, v. 96, n. 1-3, p. 135-145, 2007.
- IZRE’EL, S. Intonation Units and the Structure of Spontaneous Spoken Language : A View from Hebrew. IDP05 International Symposium on Discourse-Prosody Interfaces. *Anais...*, p. 1-20, 2011.
- KAY, S. R.; FISZBEIN A.; OPLER L. A. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull.* 13 (2), p. 261-76, 1987.
- LANGUUPES, C. A Linguist looks at “Schizophrenic” Language. *Brain and Language*, v. 1, p. 257-276, 1974.
- LEITMAN, D. I. et al. Sensory contributions to impaired prosodic processing in schizophrenia. *Biological Psychiatry*, v. 58, n. 1, p. 56-61, 2005.
- LEWIS, D. A.; LEVITT, P. Schizophrenia as a disorder of neurodevelopment. *Annual Review of Neuroscience*, v. 25, n. 1, p. 409-432, 2002.
- LIDDLE, P. F. et al. Thought and Language Index : an instrument for assessing thought and language in schizophrenia Thought and Language Index : an instrument for assessing thought and language in schizophrenia. *British Journal of Psychiatry*, p. 326-330, 2002.
- MAROTTA, G. Aspetti fonologici e prosodici nell’afasia di Broca. In: FAVILLA, M. E.; FERRONI, L. (Eds.). *Neurolinguistica e disturbi del linguaggio*. Perugia: Guerra, 2009. p. 19-38.
- MARTIN, P. WinPitch Corpus: A text to speech alignment tool for multimodal corpora. LREC. *Anais...* Lisboa: 2004. Disponível em: <http://www.lrec-conf.org/proceedings/lrec2004/pdf/780.pdf>. Acesso em: 1 de junho de 2019.
- MELLO, H. et al. Transcrição e segmentação prosódica do corpus C-ORAL-BRASIL: critérios de implementação e validação. In: RASO, T.; MELLO, H. (Eds.). *C-ORAL-BRASIL I: Corpus de referência do português brasileiro falado informal*. 1. ed. Belo Horizonte: Editora UFMG, 2012. p. 125-176.
- MITTMANN, M. M.; BARBOSA, P. A. An automatic speech segmentation tool based on multiple acoustic parameters. *Chimera*, v. 3, n. 2, p. 133-147, 2016.
- MONEGLIA, M.; RASO, T. Notes on Language into Act Theory (L-Act). In: RASO, T.; MELLO, H. (Eds.). *Spoken Corpora and Linguistic Studies*. Amsterdam / Philadelphia: John Benjamins Publishing Company, 2014. p. 468-495.

- NENCIONI, G. Parlato-parlato, parlato-scrito, parlato-recitado. In: NENCIONI, G. (Ed.). *Di scritto e di parlato*. Discorsi linguistici. Bologna: Zanichelli, 1983.
- RASO, T. et al. Segmentation and analysis of the two English excerpts: the Brazilian team proposal. In: *In search for a reference unit of spoken language: a corpus driven approach*. Amsterdam / Philadelphia: John Benjamins Publishing Company, no prelo.
- RASO, T. O corpus C-ORAL-BRASIL. In: RASO, T.; MELLO, H. (Eds.). *C-ORAL-BRASIL I: Corpus de referência do português brasileiro falado informal*. 1. ed. Belo Horizonte: Editora UFMG, 2012a. p. 55-90.
- RASO, T. O C-ORAL-BRASIL e a Teoria da Língua em Ato. In: RASO, T.; MELLO, H. (Eds.). *C-ORAL-BRASIL I: Corpus de referência do português brasileiro falado informal*. 1. ed. Belo Horizonte: Editora UFMG, 2012b. v. 1p. 91-123.
- RASO, T. Prosodic constraints for discourse markers Tommaso. In: RASO, T.; MELLO, H. (Eds.). *Spoken Corpora and Linguistic Studies*. Amsterdam / Philadelphia: John Benjamins Publishing Company, 2014. p. 411-468.
- RASO, T.; MELLO, H. C-ORAL-BRASIL II: Corpus de referência do Português Brasileiro falado formal, em preparação.
- RASO, T.; MELLO, H. (Eds.). *C-ORAL-BRASIL I: Corpus de referência do Português Brasileiro falado informal*. Belo Horizonte: Editora UFMG, 2012.
- RASO, T.; MITTMANN, M. M. As principais medidas da fala. In: RASO, T.; MELLO, H. (Eds.). *C-ORAL-BRASIL I: Corpus de referência do português brasileiro falado informal*. 1. ed. Belo Horizonte: Editora UFMG, 2012. p. 177-221.
- RASO, T.; ROCHA, B. Como a categoria de atitude condiciona a metodologia para o estudo de ilocuções. *Diadorim*, v. 17, n. 2, 2015.
- RASO, T.; MITTMANN, M. M.; CAVALCANTE, F. A. Prosodic forms of the Topic information unit in a cross-linguistic perspective: a first survey. In: De MEO, A.; DOVETTO, F. *Proceedings from GSCP La comunicazione parlata*. Canterrano: Aracne Editrice, 2017.
- RASO, T.; MITTMANN, M. M.; MENDES, A. C. O. O papel da pausa na segmentação prosódica de corpora de fala. *Revista de Estudos da Linguagem*, v. 23, n. 3, p. 883-922, 2015.
- ROCHA, B. *Uma metodologia empírica para a identificação e descrição de ilocuções e a sua aplicação para o estudo da Ordem em PB e Italiano*. (Tese de Doutorado em Estudos Linguísticos) - Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2016.
- SALGADO, J. V. Neuropsicologia da esquizofrenia. In: FUENTES, D. et al. (Eds.). *Neuropsicologia: Teoria e prática*. Porto Alegre: Art, 2008. p. 273-287.
- SCHULTZ, S. K.; ANDREASEN, N. C. Schizophrenia. *Lancet*, v. 353, n. 9162, p. 1425-1430, 1999.

SIGNORINI, S. Topic e soggetto in corpora di italiano parlato spontaneo. 2005. PhD Thesis (Dottorato di ricerca in linguistica italiana) – Dipartimento di Linguistica, Università degli Studi di Firenze, Firenze, 2005.

SORIANELLO, P. *Linguaggio e sindrome di Down*. Milano: Franco Angeli, 2012.

VIEIRA, M. A.; RASO, T.; OLIVEIRA, E. *Métodos automáticos de avaliação da qualidade acústica*, em preparação.



Data de submissão: 15/03/2018

Data de aceite: 29/01/2019