



Universidade do Estado do Rio de Janeiro

Centro de Educação e Humanidades

Instituto de Letras

Thiago Quintanilha da Silva

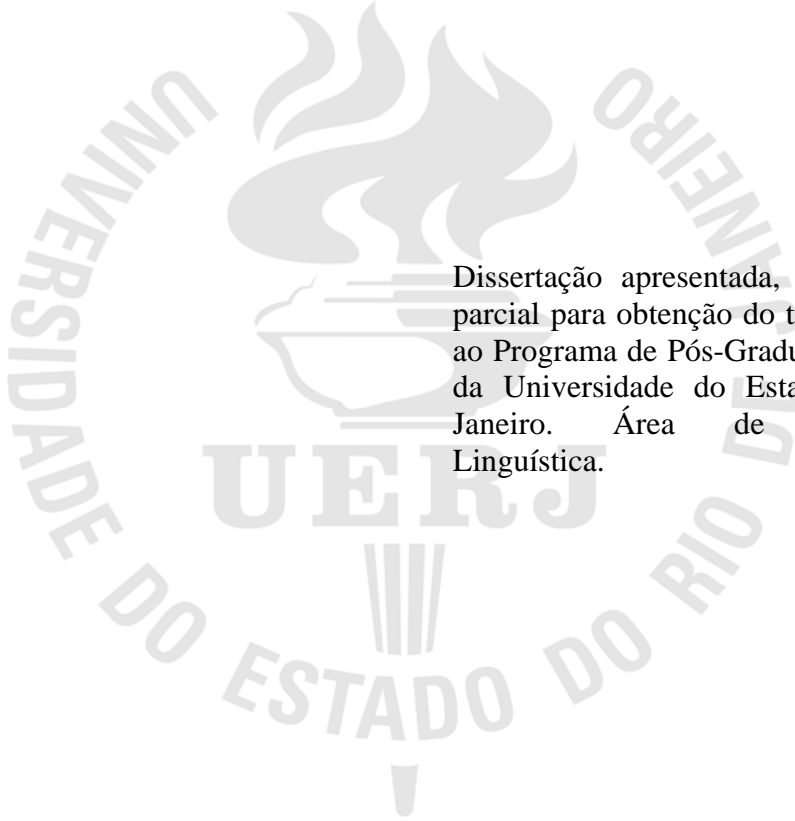
**Diagnóstico dos mecanismos disfluentes
em orações relativas do português brasileiro**

Rio de Janeiro

2012

Thiago Quintanilha da Silva

**Diagnóstico dos mecanismos disfluentes
em orações relativas do português brasileiro**



Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Letras da Universidade do Estado do Rio de Janeiro. Área de concentração: Linguística.

Orientador: Prof. Dr. Ricardo Joseh Lima

Rio de Janeiro

2012

CATALOGAÇÃO NA FONTE
UERJ/REDE SIRIUS/CEHB

S586 Silva, Thiago Quintanilha.
Diagnóstico dos mecanismos disfluentes em orações relativas do português brasileiro / Thiago Quintanilha Silva. – 2012.
98f.

Orientador: Ricardo Joseh Lima.
Dissertação (mestrado) – Universidade do Estado do Rio de Janeiro, Instituto de Letras.

1. Atos da fala (Linguística) - Teses. 2. Competência e desempenho (Linguística) - Teses. 3. Competência comunicativa – Teses. 4. Língua portuguesa – Brasil - Teses. 5. Gramática gerativa – Teses. 6. Gramática comparada e geral – Síntaxe - Teses. 7. Língua portuguesa – Orações - Teses. 8. Psicolinguística - Teses. I. Lima, Ricardo Joseh. II. Universidade do Estado do Rio de Janeiro. Instituto de Letras. III. Título.

CDU 801.4:159.955.652

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação desde que citada a fonte

Assinatura

Data

Thiago Quintanilha da Silva

**Diagnóstico dos mecanismos disfluentes
em orações relativas do português brasileiro**

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação Letras, da Universidade do Estado do Rio de Janeiro. Área de concentração: Linguística.

Aprovada em 13 de abril de 2012

Banca Examinadora:

Prof. Dr. Ricardo Joseh Lima (Orientador)

Instituto de Letras da UERJ

Prof^a. Dra. Marina Rosa Ana Augusto

Instituto de Letras da UERJ

Prof^a. Dra. Eduardo Kenedy Nunes Arêas

Instituto de Letras da UFF

Rio de Janeiro

2012

AGRADECIMENTOS

Em primeiro lugar a Deus. Sem ele, não teria forças, não teria ânimo. Com Deus, aprendi que para vencer, é necessário perseverança.

À minha família, principalmente à minha mãe, por sua dedicação, carinho e amor; e também ao meu irmãozinho, um dos principais motivos por continuar buscando mais e mais na minha vida.

À minha Linda, pela paciência e pelo incentivo. Sempre companheira e prestativa, jamais deixou de caminhar comigo.

Sem sombras de dúvidas, ao meu orientador, professor, amigo e maior incentivador, Ricardo Joseh Lima. Ter acreditado em mim foi uma grande honra. Sem seu apoio constante, nada disso aconteceria.

Às professoras Marina Augusto e Letícia Corrêa pelas suas aulas e pela ajuda primordial em relação a esta dissertação. Sem a solicitude de ambas, ao nos passar os arquivos da pesquisa original, nada sairia do papel.

Aos amigos, todos aqueles que, por contato pessoal, ou por redes sociais, se preocuparam com a minha jornada. Entre esses amigos, agradeço à Queila de Castro Martins, amiga de jornada desde o primeiro dia de graduação até os momentos finais do Mestrado.

À CAPES, pelo financiamento por todo esse tempo de Mestrado.

A todos os meus professores da UERJ, responsáveis diretos pela minha formação pessoal e intelectual.

Ao pessoal da secretaria de pós-graduação, principalmente à Tânia, que sempre torceu e me ajudou muito durante o período do Mestrado.

Obrigado a todos por ajudarem a fazer de um sonho, uma realidade.

RESUMO

QUINTANILHA, Thiago. *Diagnóstico dos mecanismos disfluentes em orações relativas do português brasileiro*. 2012. 98f. Dissertação (Mestrado em Linguística) - Instituto de Letras, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2012.

Alguns estudos na área da Psicolinguística tratam sobre a interrupção da forma natural da fluência da fala. Essas suspensões são denominadas disfluências e podem ser de variados tipos como, por exemplo, repetições, substituições ou pausas. Entre alguns trabalhos sobre disfluências, destaca-se o realizado por Jaeger (2005), que tratou da omissão do relativizador *that* na língua inglesa e a possibilidade desse fator ser responsável pela criação de disfluências. Nessa mesma linha, objetivamos analisar as disfluências na camada dos relativizadores – camada CP – em língua portuguesa brasileira. Para isso, tomaram-se dados obtidos por Corrêa et al (2008) sobre produção de orações relativas. Nesse estudo, os autores, buscando verificar a complexidade da formulação de orações relativas, manipularam duas condições: plenamente planejada – que reduziria a carga de processamento e, conseqüentemente, levaria à maior produção de orações relativas padrões – e parcialmente planejada – a qual aumentaria a carga de processamento e, conseqüentemente, maior produção de orações relativas cortadoras e resumptivas. Diante disso, sabendo que o planejamento de estruturas mais complexas pode ser um fator que gere mais sentenças disfluentes, resta saber em qual das condições, estipuladas por Corrêa et al (2008), haveria maior caso de disfluências. Acreditamos que a condição parcialmente planejada resultaria mais casos de disfluências em comparação à condição inteiramente planejada. Para verificação de tal hipótese, a presente pesquisa focou suas atenções no banco de dados de Corrêa et al (2008). Após determinados critérios, privilegiamos sentenças de orações relativas genitivas, orações relativas de objeto indireto funcional e orações relativas de objeto indireto lexical, pois tais seriam construções sintáticas mais complexas e passíveis de terem mais casos de disfluências. Identificadas e quantificadas todas as falas de interesse, criamos tabelas para melhor visualização das disfluências em localizações específicas na sentença como, por exemplo, antes da oração relativa (AOR), na oração relativa (NAOR) e depois da oração relativa (DPOR). Com tais dados analisados, verificamos que eles corroboravam a hipótese levantada: na condição parcialmente planejada há mais casos de disfluências do que na condição inteiramente planejada, principalmente quando se é focalizada a localização NAOR. Entre os principais resultados, percebemos que as disfluências do tipo pausas preenchidas aparecem em grande quantidade na localização NAOR, fator que revela uma característica especial da oração relativa. Essa disfluência, nesse trecho da sentença, revela que os falantes não somente fazem uma procura lexical, como também, um planejamento discursivo. Tais resultados nos motivaram a pensar em outras determinadas situações de pesquisa como, por exemplo, a análise das disfluências em sentenças de voz passiva.

Palavras-chave: Disfluências. Fala espontânea. Complexidade sintática. Dificuldades de processamento.

ABSTRACT

Some studies in Psycholinguistics area deal with the interruption of the natural form of fluency speech. These suspensions are called disfluencies and can be of various types, for example, repetitions, substitutions or pauses. Among some works about disfluency, the one made by Jaeger (2005) deserves attention. His research dealt with the omission of the relativizer *that* in English and the possibility of this factor being the responsible for the creation of disfluencies. With this in mind, we aimed to analyze disfluencies in the layer of relativizers - CP layer - in Brazilian Portuguese. For this, data were taken by Corrêa et al (2008) on production of relative clauses. In this study, the authors, trying to verify the complexity of the formulation of relative clauses, manipulated two conditions: totally planned – which would reduce the processing load and, consequently, major production of standard relative clauses – and partially planned – which would increase the processing load and, therefore, major production of gap and resumptive relative clauses. So, knowing that the planning of more complex structures can be a factor that generates more disfluent sentences, the question in which of the conditions stipulated by Corrêa et al (2008), we would find major cases of disfluency. We believe that the condition partly planned results in more cases of disfluencies in comparison with the condition entirely planned. To investigate this hypothesis, this research has focused its attention on the database of Corrêa et al (2008). After certain criterions, were privileged sentences of genitive relative clauses, functional indirect object of relative clauses and lexical indirect object of relative clauses, because these would be syntactic constructions more complex and likely to have more cases of disfluencies. Once identified and quantified all kinds of speech that interested us, there were created tables in order to enable a better visualization of disfluencies in specific locations in sentences, for instance, before the relative clause (BRC), in relative clause (IRC) and after the relative clause (ARC). With such data analyzed, it was found that the data corroborate the hypothesis: in the partially planned condition there are more cases of disfluencies than there are in the totally planned condition, principally when the location IRC is focused. Among the main results, we noticed that the disfluencies such as filled pauses appear in large amounts in IRC location, a factor that reveals a special feature of the relative clause. This disfluency, in this stretch of the sentence, reveals that the speakers are not only looking for a lexical, but also a discursive planning. These results motivated us to think of some other research situations, for example, analysis of disfluencies in the passive voice sentences.

Keywords: Disfluencies. Speech spontaneous. Complexity syntactic. Difficulties of processing.

SUMÁRIO

	INTRODUÇÃO	10
1	QUADRO TEÓRICO	12
1.1	Gerativismo	12
1.2	Psicolinguística	14
1.3	Disfluências	17
2	REVISÃO DE LITERATURA	20
2.1	Automonitoramento em Levelt (1989)	21
2.2	Modelização das disfluências em Sriberg (1994)	25
2.3	Disfluências em uma análise de língua portuguesa (Moniz, 2006)	29
2.4	Conceito de complexidade gramatical em Clark e Wasow (1998)	39
2.5	Disfluências na camada CP	40
2.5.1	<u>Dialeto <i>Orust</i>: análise de Horne et al (2005)</u>	40
2.5.2	<u>Omissão de <i>that</i> na língua inglesa: análise de Jaeger (2005)</u>	42
2.6	Orações relativas no PB vistas por Corrêa et al (2008)	44
3	DETALHES DA PESQUISA DE CORRÊA ET AL (2008)	47
3.1	Estratégias diferenciais na produção de orações relativas restritivas (Corrêa et al, 2008)	47
3.2	Objetivos em Corrêa et al (2008)	47
3.3	Três estratégias diferentes de formação de orações relativas	48
3.4	Orações relativas em condições diferentes de planejamento	48

3.5	Experimento	51
3.6	Resultados e discussões finais	52
4	METODOLOGIA	55
4.1	Corpus	55
4.1.1	<u>Diretrizes da coleta dos dados</u>	56
4.1.2	<u>Especificações dos dados recolhidos</u>	59
4.2	Transcrições em Corrêa et al (2008)	62
4.3	Transcrições: nossas propostas	64
4.4	Disfluências do <i>corpus</i>	69
4.4.1	<u>Disfluências isoladas e combinadas</u>	69
4.4.2	<u>Localização das disfluências</u>	71
5	RESULTADOS	73
5.1	Complicações em AOR	73
5.2	Caso DPOR: mais disfluências na condição planejada	77
5.3	Análise minuciosa de NAOR	80
6	CONSIDERAÇÕES FINAIS	87
	REFERÊNCIAS	91
	ANEXO A – Sentenças AOR	94
	ANEXO B – Sentenças NAOR	95
	ANEXO C – Sentenças DPOR	98

INTRODUÇÃO

Preocupado em elaborar um modelo teórico formal inspirado na matemática e com capacidade de descrever e explicar abstratamente o funcionamento da linguagem humana, o gerativismo, desenvolvido por Noam Chomsky (1965), é uma linha de pesquisa que dá suporte a muitos estudos. Não só isso, as preocupações impostas por esse modo de pensar a língua suscitaram diferentes abordagens ao longo da história. Um outro modo de pensar a linguagem é a psicolinguística, que tem como preocupação a relação entre os processos mentais e a fala.

De fato, se não fosse o desenvolvimento teórico nessas duas áreas, hoje haveria uma grande dificuldade em explicar a fala entre os seres humanos. Pelo senso comum, a fala é considerada apenas como uma faculdade responsável pela emissão de sons, palavras, ou discursos. Além disso, é um objeto que, fatalmente, cria conceitos, ou preconceitos, em determinados falantes. É muito comum juízos de valores relacionados à fala de determinada pessoa em um momento de discurso oral. Avalia-se tudo, de riqueza vocabular à postura, passando, inclusive, pelo grau de fluência da fala. É exatamente nesse momento que os estudos psicolinguísticos, de modo geral, podem contribuir para esclarecer aspectos da fala, evitando, assim, possíveis preconceitos em relação à capacidade oral de determinado falante.

Em relação à fluência da fala em discursos espontâneos, conceito que engloba o de disfluência (com exceção à disfemia, conhecido popularmente como gagueira), muito pode ser explicado mediante estudos psicolinguísticos. Primeiramente, é necessário deixar claro que a fala é produto de um complexo sistema cognitivo. Há etapas no estado mental que antecedem ao evento da fala, como por exemplo, a conceptualização (momento em se gera mensagens pré-verbais), a formulação (momento em que as mensagens pré-verbais são traduzidas em estruturas linguísticas) e a articulação da mensagem (quando é executado o plano fonético da mensagem) (Levelt, 1989). Na fase de conceptualização, há um elemento chamado de monitor, cuja responsabilidade é a de controlar a produção discursiva e detectar e corrigir os possíveis erros. Desse modo, já é possível vislumbrar explicações plausíveis em relação à fluência da fala, ou a falta dela, isto é, a momentos disfluente. Parece-nos mais claro que ser disfluente como, por exemplo, substituir algum elemento durante o discurso oral, como no exemplo “*essa questão, esse pronome, até às vezes dizem qual pronome*” (Quintanilha et al, 2010), é uma espécie de estratégia utilizada pelo falante para alcançar um

objetivo bem específico em relação ao seu interlocutor: estabelecer, da melhor maneira possível, a comunicação.

É importante ressaltar que o aprofundamento das observações relativas aos momentos disfluentes revela outras estratégias dos falantes em variadas estruturas sintáticas. A complexidade das estruturas, a localização sintática das disfluências, entre outros aspectos, têm sido alvo de vários estudos (ver Horne 2005, Jaeger 2005, Clark e Wasow, 1998). De forma geral, esses autores salientam o fato de haver estruturas sintáticas mais complexas em que o falante, provavelmente, suspenderá a fluência de sua fala, entre elas, as construções com pronomes relativos.

Esse tipo de sentença será o alvo deste trabalho, principalmente em sentenças no português brasileiro. Os estudos de Corrêa et al (2008) versaram sobre estruturas sintáticas construídas com orações relativas. Tinham como objetivo verificar os custos computacionais associados à produção de orações relativas. A base de dados permite verificar a ocorrência de disfluências, nessas construções sintáticas, com dados do PB, a fim de se contribuir para a discussão sobre as dificuldades do formulador sintático, que podem ser observados por meio da produção de disfluências.

Para tal, esta dissertação seguirá determinadas diretrizes: um primeiro capítulo que terá como base o quadro teórico que norteará este trabalho; um segundo capítulo com toda a literatura consultada para elaboração desta pesquisa; um terceiro capítulo que explicitará a metodologia empregada pela pesquisa que serve de base para a nossa (Corrêa et al, 2008); um quarto capítulo com todos os detalhes da metodologia empregada por nós; um quinto capítulo em que se discutem os resultados obtidos durante a pesquisa; e um capítulo conclusivo em que haverá uma retomada da discussão e as considerações finais.

1 QUADRO TEÓRICO

Como explicar a seguinte sentença em um discurso espontâneo “*essa questão, esse pronome, até às vezes dizem qual pronome*” (Quintanilha et al, 2010)? Nervosismo? Timidez? Insegurança? A interrupção da fluência da fala, isto é, a disfluência, revela ser algo bem mais complexo do que apenas constatações emocionais.

Para se chegar a tal assertiva, alguns detalhes importantes devem ser observados. O exemplo demonstra um momento de fala espontânea. Mesmo assim, essa fala não pode ser dissociada de uma estrutura. Por isso, há necessidade de uma teoria linguística que dê base para dar conta dessa estrutura. Esse apoio teórico será encontrado no Gerativismo.

Não só isso, é necessário mais informações sobre o comportamento linguístico dos indivíduos. Para isso, outro ponto de apoio é requerido. Esse ponto se encontra na Psicolinguística. Afinal, por se tratar de fala, elemento que envolve fatores como tempo, condição, entre outros, é necessário que se verifique a própria produção da fala.

Por fim, a sentença demonstra, claramente, interrupções, ou seja, é uma fala espontânea disfluente. Se há disfluência, imagina-se, então, que há uma dificuldade no processamento do formulador sintático. Eis a necessidade, portanto, de ter base sobre estudos que envolvam a fala e a mente.

Dessa maneira, o caminho para comprovar a complexidade da disfluência está traçado. Na seção 1.1, haverá a apresentação dos aspectos da teoria gerativa relevantes para se entender disfluências. Já a seção 1.2 contará com as contribuições psicolinguísticas para o tema aqui em questão. A seção 1.3 tratará das disfluências em si, como, por exemplo, o encaminhamento teórico utilizado para análise.

1.1 Gerativismo

Tratar de fala espontânea, entre outras abordagens, é tratar de estruturas. Por isso, é importante ter como base uma teoria linguística que dê conta delas. Assim, evidencia-se a necessidade de uma teoria gerativa para dar solo firme a essa primeira abordagem estrutural de uma fala espontânea.

Fundado por Chomsky, o gerativismo tem como concepção “a capacidade humana de falar e entender uma língua (pelo menos), isto é, o comportamento linguístico dos indivíduos deve ser compreendido como o resultado de um dispositivo inato, uma capacidade genética e, portanto, interna ao organismo humano” (Kenedy, 2008). Essa capacidade genética, biologicamente, deve estar na relação cérebro/mente da espécie, sendo que, tem como função, construir a competência linguística de um falante. Essa disposição inata para a competência linguística é o que ficou conhecido como faculdade da linguagem. No entanto, há ainda algumas informações das quais necessitamos para alcançarmos uma compreensão mais satisfatória dos elementos gerativistas que serão usados neste texto.

Primeiramente, ao longo da história gerativa, houve uma preocupação, entre outras, de criar um modelo que pudesse descrever como os constituintes das sentenças eram gerados e como estes se transformavam em outros, exemplo disso foi a Gramática Transformacional desenvolvida durante as décadas de 60 e 70. Para tal, utilizou-se um diagrama arbóreo, denominado, também, de árvore sintática. Seguindo a lógica, cada parte da árvore recebe um nome, como exemplificado na figura 1:

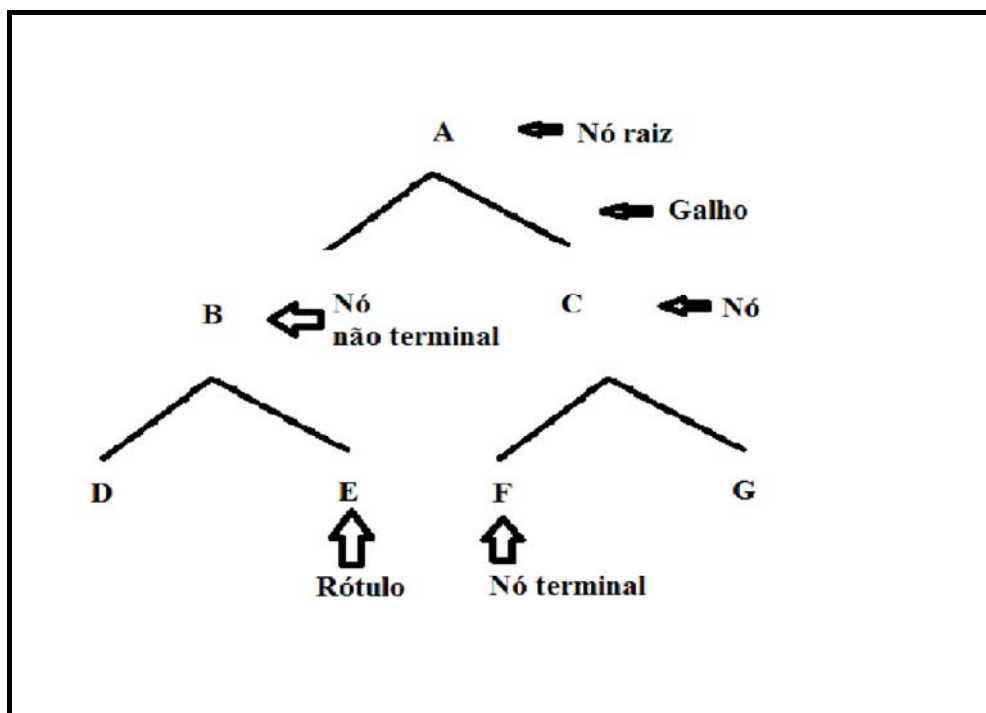


FIGURA 1: Árvore sintática (Curso de Linguística da Frase, ministrado pela professora Marina Augusto no primeiro semestre de 2010 na Universidade do Estado do Rio de Janeiro).

Em cada nó se encontram categorias lexicais e funcionais. As lexicais são classificadas em Nome, Verbo, Preposição, Adjetivo e Advérbio – *Noun (N)*, *Verb (V)*, *Preposition (P)*,

Adjective (A) e Adverb (Adv) em língua inglesa e suas siglas, respectivamente. As categorias funcionais são denominadas de Flexão, Negação e Determinante – *Inflectional (I), Negative (Neg) e Determiner (D)* em língua inglesa e suas siglas, respectivamente. Como são os núcleos e responsáveis pela nomeação dos sintagmas (*Phrases – P –*, em língua inglesa e sua sigla respectivamente), formam, desse modo, as categorias sintagmáticas, tanto lexicais (NP, VP, PP, AP e AdvP), como funcionais (IP, NegP e DP). Entre os funcionais, há uma categoria para qual são movidos, por exemplo, os elementos topicalizados e as palavras interrogativas e onde são gerados, por exemplo, os pronomes relativos. Tal camada é denominada de CP (*Complementizer Phrase* no inglês).

De acordo com o Programa Minimalista, versão mais recente do gerativismo, a camada CP é para onde os elementos Qu – assim, também, chamados pronomes relativos – se movimentam na derivação da sentença. Segundo Chomsky (1976), um elemento Qu é alçado à posição de especificador (Spec) do sintagma complementizador (CP).

Diante dessas possibilidades, a teoria linguística cumpre bem seu papel até certo ponto. Fica claro que uma teoria voltada à estrutura das sentenças é importante para uma análise mais fundamentada dos fenômenos linguísticos. Verificar a derivação da sentença nas árvores sintáticas, também é outro fator essencial para maior compreensão de tais fenômenos. No entanto, somente essas informações estruturais não seriam suficientes para dar conta do tema alvo dessa pesquisa: as disfluências. Para isso precisamos de uma abordagem que tenha a produção da fala como preocupação primordial. Em outras palavras, um estudo em que se evidencie uma preocupação da produção da fala em sua formulação sintática no âmbito mental, os elementos que estão presentes na formulação das sentenças e as dificuldades que podem ser solucionadas por esse modo de ver a fala.

1.2 Psicolinguística

As complexidades, em alguns nós frasais da árvore sintática, como, por exemplo, a camada CP, podem acarretar dificuldades de produção para os falantes. Essas dificuldades podem ser identificadas por meio de algumas estratégias adotadas pelos mesmos. Entre elas, estão as disfluências, alvos deste texto. Desse modo, argumentar sobre o fenômeno da disfluência é também argumentar sobre produção, e, logicamente, as dificuldades no formulador sintático acarretadas por tal.

Primeiramente, é importante destacar que a fala é produto de um complexo sistema cognitivo. Segundo Levelt (1989), há algumas etapas, no estado mental, que antecedem ao evento da fala como, por exemplo, a conceptualização, a formulação e a articulação da mensagem (ver figura 2)

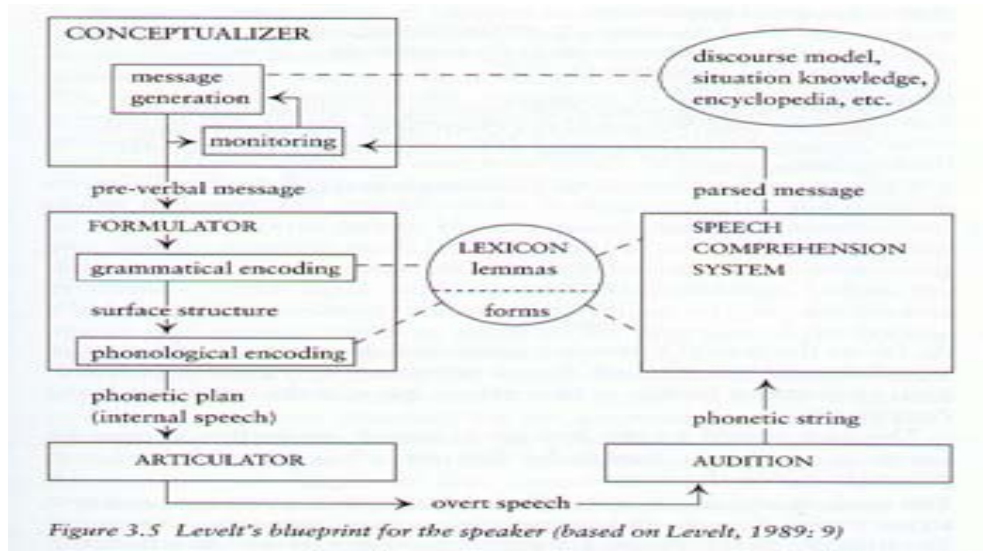


FIGURA 2: Modelo de processamento de Levelt (1989) – fonte <http://homepage.ntlworld.com/vivian.c/SLA/Bilingualism%20Models%20and%20Memory.htm>.

De acordo com Levelt (1989), a conceptualização da mensagem é um momento em que o conceptualizador gera mensagens pré-verbais. Não somente isso, esse estágio é responsável pela ocorrência de vários fatores ao mesmo tempo como, a intenção da fala, a seleção das informações relevantes para a realização da fala, a ordenação dessas informações para assim serem expressas, a memorização do que já foi dito pelo falante anteriormente e a monitoração do que o próprio falante está dizendo. Não obstante, para a mensagem ser codificada, é necessário, ainda, acessar dois tipos de conhecimentos: o procedimental e o declarativo. O conhecimento procedimental é resultado da memória de trabalho, a qual contém toda informação acessível aos falantes. Já o conhecimento declarativo, representado dentro de círculos, tem disponibilizado a memória de longo-termo (isto é, o conhecimento que o falante constrói ao longo da vida) e o registro de discurso (ou seja, possibilidade de guardar o que foi dito pelo próprio falante e pelos interlocutores).

Em relação à formulação, Levelt (1989) afirma ser o momento em que uma estrutura conceitual (mensagem pré-verbal) é traduzida em uma estrutura linguística. Essa tradução

ocorre devido a dois passos que são a codificação gramatical e a codificação fonológica. A codificação gramatical é um procedimento que acessa lemas¹ e construções sintáticas². Quando o lema combina com a parte da mensagem pré-verbal, torna disponível, então, sua sintaxe, ativando procedimentos de construção sintática. Com todos os lemas acessados e todos os procedimentos sintáticos feitos, o codificador gramatical produz uma estrutura de superfície, que é um conjunto ordenado de lemas agrupados em sintagmas e sub-sintagmas de vários tipos. Passada a codificação gramatical, acontece, então, a codificação fonológica, cuja função é reter ou construir um plano fonético ou articulatório para cada lema ou enunciado. O codificador fonológico acessa a forma lexical, encontrando informações morfológicas e fonológicas. Cabe lembrar que essas informações ainda não são o plano articulatório (a fala explícita), mas a representação interna (fala interna) de como o enunciado planejado deverá ser articulado. Por fim, a articulação, segundo Levelt (1989), é o componente de processamento que executa o plano fonético pela musculatura dos sistemas respiratório, laríngeo e sub-laríngeo. O produto final da articulação é a fala explícita.

Parecem ser esses procedimentos, da concepção da mensagem à produção da fala, perfeitos. No entanto, ocorrem, como já citado anteriormente, algumas hesitações do próprio falante ou trocas de palavras por outras, entre outros casos. Essas situações nos revelam que, em alguns momentos, há algum ajuste na formulação da mensagem. Além disso, o mais curioso, ainda, é que, em alguns casos, o próprio falante percebe o erro na mensagem, depois ou até mesmo antes da própria mensagem ser produzida. Levelt (1989), então, explicita que no conceptualizador da mensagem se encontra o monitor, cuja responsabilidade é a de controlar a produção discursiva e detectar e corrigir os possíveis erros. Sendo assim, o falante tem, então, essa capacidade de autocorrigir, ou melhor, automonitorar seu próprio discurso como no exemplo “*eles estão, eles têm muitas divisões ta*” (Quintanilha et al, 2010). Aliás, de acordo com esse exemplo, há uma ocorrência disfluente, devido à interrupção da fluência da fala feita pelo falante, pois, de acordo com o monitor no conceptualizador, há uma informação que necessita ser substituída. Dessa forma, então, quando há substituição de termos (estratégia que também é utilizada de outras formas), entre outros momentos, há o fenômeno da disfluência.

¹ Informações semânticas ou sintáticas que o falante tem armazenado em seu léxico mental, ou seja, seu conhecimento declarativo.

² Informações como, por exemplo, o sujeito selecionado é um nome contável ou o verbo escolhido tem o agente como sujeito.

Todavia, ter conhecimento de aspectos do formulador sintático não implica totalmente no conhecimento das disfluências. É necessário ir mais a fundo nesse tema para, assim, ser possível tecer argumentações sobre quais dificuldades o formulador sintático atravessa no momento de uma construção frásica disfluyente complexa. Entretanto, para chegar a essa conclusão, mister é entender as disfluências em seu âmbito mais básico para, assim, chegar ao seu momento mais complexo.

1.3 Disfluências

Basicamente, as disfluências são entendidas como reparos feitos pelo próprio falante, em situação de fala, a fim de que haja progressão da mensagem pretendida, como, por exemplo, “*essa questão, esse pronome, até às vezes dizem qual pronome*” (Quintanilha et al, 2010). Além disso, verifica-se que as disfluências podem ocorrer tanto no discurso espontâneo – seja em situações formais, seja em informais –, como em uma fala preparada, ou seja, em situações de falas planejadas estrategicamente.

No entanto, pelo senso comum, uma pessoa que produziu um discurso disfluyente em alguma situação de uso da fala espontânea, pode acabar sofrendo preconceito, ou melhor, seu discurso poderá ser considerado inferior. Enfim, uma sequência oral sem interrupções, de acordo com o senso comum, é considerada uma forma de expressão de alguém que “fala bem”. Contudo, os estudos das disfluências apontam para outro caminho. O falante disfluyente (retirando, logicamente, os casos patológicos), na verdade, está em um processo cognitivo altamente complexo, com o intuito de aprimorar sua mensagem o máximo possível, para que o interlocutor tenha o melhor entendimento da mensagem.

Portanto, se há um esforço do falante em transmitir uma mensagem da melhor maneira possível para seu interlocutor, variadas estratégias, para tal fim, são empregadas. Devido a isso, é possível criar um sistema de anotação que nomeie as estratégias disfluyentes utilizadas pelos falantes, baseada no comportamento dos ouvintes diante de uma disfluência (ver Shriberg, 1994 e Moniz, 2006). Diante disso, entre as categorias diferentes de disfluências, se encontram em destaque:

- Pausas preenchidas – Consistem em vocalizações que preenchem um espaço de silêncio na comunicação.

Ex.: “*esse túnel atolado aa fica um bocadinho antes da estrada das minas antigas*” (Moniz, 2006).

- Repetições – Remetem à repetição de material linguístico.

Ex.: “*porque o castigo das suspensões não vai, não vai educar o aluno nesse sentido*” (Moniz, 2006).

- Substituições – Correspondem a alterações de segmentos ou palavras com correspondência sintática e/ou semântica com a zona a corrigir.

Ex.: “*o inquirito, os inquiritos forma feitos há cerca de vinte anos*” (Moniz, 2006).

- Marcadores de edição – Correspondem a vocábulo ou expressões que transmitem informação de correção de um item ou de uma sequência produzida, como melhor, quer dizer, não, etc.

Ex.: “*é, é, quer dizer, é a lei do castigo*” (Moniz, 2006)

- Truncamentos de palavras - Correspondem a apagamentos de material linguístico, isto é, abandono por completo do material linguístico produzido.

Ex.: *pronomes pro- pessoais* (Moniz, 2006)

É curioso notar que, dependendo da localização da disfluência, é possível interpretar qual a dificuldade de processamento encontrada pelo falante. Por exemplo, se ocorrem pausas preenchidas em início de oração, essas disfluências indicam que o falante está em um momento de planejamento do discurso. No entanto, se essas pausas são encontradas em meio ao momento discursivo, pode se inferir que o falante está em uma procura lexical. Outro exemplo são os truncamentos de palavras. A presença dessas disfluências realiza um monitoramento interno praticado pelo falante em um momento de discurso espontâneo.

Desse modo, a base para o entendimento das disfluências, envolvendo orações relativas, está organizada. Uma teoria linguística, encontrada no gerativismo, dá sua contribuição para esclarecer a estrutura de sentenças da língua. A Psicolinguística oferece como contribuição, análises da produção da fala espontânea, a qual pode apresentar disfluências. Mesmo que essa organização vislumbre uma saída para o tema em questão deste texto, ainda há um pouco mais a se fazer. É necessário ver as disfluências em uma sentença complexa, para isso é preciso conceituar complexidade. Verificado esse fator, o próximo é analisar uma camada que é considerada complexa: a camada CP, que, em algumas línguas pode ser omitida, tamanha a sua complexidade. Por fim, se faz importante submeter sentenças com pronomes relativos da língua portuguesa brasileira em condições diferentes de produção para que se entenda com mais clareza as disfluências. Para maior precisão, o ideal é percorrer por estudos anteriores sobre tais características. Eis então, os nossos próximos passos, uma revisão cuidadosa da literatura.

2. REVISÃO DE LITERATURA

As seções 1.1, 1.2 e 1.3 do capítulo 1 trataram de alguns momentos do gerativismo, da psicolinguística e de aspectos introdutórios da disfluência relevantes para o tema tratado neste texto: as disfluências envolvidas em orações relativas e as dificuldades que essas sentenças geram ao formulador sintático. Todos esses pontos são essenciais para um entendimento inicial do problema em questão.

Em primeiro lugar, as disfluências necessitam ser explicadas de modo mais aprofundado. Com base em uma revisão de literatura dos estudos mais importantes – para esta dissertação, – sobre disfluências como, por exemplo, os estudos de Levelt (1989), de Shriberg (1994) e de Moniz (2006). Por conta disso, esses textos serão abordados na seção 2.1, 2.2 e 2.3 deste capítulo.

Como o tema trata das disfluências em um contexto complexo, ainda é preciso esclarecer o conceito de complexidade. Para isso, a seção 2.4 contará com a revisão do estudo de Clark e Wasow (1998), autores que estudaram os constituintes complexos. A complexidade pode ser encontrada em algumas estruturas da língua, um exemplo disso é a camada CP na árvore sintática. Tal ponto, por conta dessa complexidade, pode ser um lugar ideal para a ocorrência de disfluências. Algumas pesquisas como, por exemplo, a de Jaeger (2005) e a de Horne et al (2005), trataram sobre as disfluências na camada CP. Por conta disso, a seção 2.5 detalhará os estudos feitos por tais autores, sendo que 2.5.1 será a seção dedicada a Horne et al (2005) e 2.5.2. a Jaeger (2005).

Com a definição de disfluências, dos constituintes complexos e das dificuldades da camada CP, basta agora transferir esses conhecimentos para uma abordagem na língua portuguesa brasileira. Embora não trate das disfluências em si, Corrêa et al (2008) abordam sobre estruturas sintáticas construídas com orações relativas no PB. Uma vez feita uma análise sobre as orações relativas, um estudo sobre as disfluências ocorridas nessas estruturas se torna um bom caminho. No entanto, antes de tal estudo, é importante uma revisão dos achados de Corrêa et al (2008). Isso ocorrerá na seção 2.6 deste texto

2.1 Automonitoramento em Levelt (1989)

Os textos produzidos referentes à disfluência, no âmbito da fala e da mente, se baseiam, de certa forma, em um texto primordial para início de discussão: “*Self-monitoring and Self-repair*” (1989) de Levelt. Basicamente, o autor parte de uma visão que comportaria alguns aspectos, como, por exemplo, o monitoramento do falante, como e quando ocorre a interrupção do discurso e como é feita a correção de tal discurso.

Recorrendo a pesquisas anteriores sobre fenômenos de autocorreção (nomenclatura usada antes de o termo disfluências ser usado pela literatura crítica), principalmente em tarefas em que consistiam em descrever padrões de pontos de cor que estavam ligados entre si por linhas horizontais ou linhas verticais, Levelt (1989) percebeu que o falante monitora o seu discurso e os seus diferentes aspectos, com o auxílio das autocorreções, não apenas com o intuito de corrigir os erros. Ele verificou que essas autocorreções serviam, também, como meios de aprimorar ou especificar a mensagem para que ela se conforme à situação comunicativa em questão, como é perceptível nos exemplos “*a noro- não a sudoeste*” (Moniz, 2006) e “*Como um problema de poluição, poluição pelo som*” (Moniz, 2006). Os dois exemplos realçam essas duas perspectivas: a de correção do erro e a de refinamento da mensagem, respectivamente.

De acordo com Levelt, essas autocorreções obedecem a várias estratégias com intuito de ajuste à situação comunicativa. Elas podem ser determinadas por fatores gramaticais, atuando assim nos níveis lexical, sintático-semântico morfológico e fonético-fonológico ou, simplesmente, na percepção do falante de que sua fala precisa ser reajustada para maior compreensão daquele que o ouve. Exemplos dessa constatação do autor se encontram nas sentenças seguintes:

1) “*we go straight on, or – we enter via red, than go straight on to green*”(Levelt, 1989).

Esse exemplo revela que o falante concebeu a mensagem, no entanto, ele percebe que suas informações podem ser confusas para o ouvinte, por isso, acrescenta “*we enter via red*” para clarificar, ainda mais, sua mensagem. Desse modo, verificamos que houve problemas na concepção da mensagem.

2) “*to the right is yellow, and to the right further to the right is blue*” (Levelt, 1989).

Para assegurar a coesão com o discurso anterior, o falante acrescenta “*ainda mais para a direita*”, após perceber que já tinha dito “*para a direita*”. Desse modo, denotam-se problemas na estruturação formal da mensagem.

3) “*left to pink – er straight to pink*” (Levelt, 1989).

Ao perceber que sua orientação não corresponde à mensagem pretendida, o falante substitui a expressão “*esquerda*” por “*direita*”. Assim, ocorre um erro de ordem lexical.

4) “*what things are this kid – is this kid going to say incorrectly?*” (Levelt, 1989).

Nessa sentença exemplo, o falante percebe um erro sintático, ou seja, percebe a falta de concordância entre sujeito e predicado, alterando, assim, na sequência, a forma verbal. Dessa maneira, há, segundo Levelt (1989), um erro sintático ou morfológico.

5) “*a unt – unit from the yellow dot*” (Levelt, 1989).

No exemplo 5, ocorre um erro fonológico. O falante percebe o erro de pronúncia da vogal e a substitui na sequência.

6) “*ainda por cima a coisa veio falada como, veio exposta como um problema de poluição*” (Moniz, 2006).

Nesse exemplo, o falante expõe um modo de falar, mas faz uma troca de termos, ao perceber que o contexto comunicativo em que se encontrava, não condizia com aquela expressão. Em outras palavras, houve uma inadequação à situação comunicativa.

Um segundo aspecto abordado por Levelt é a questão da interrupção. Por meio dos dados analisados, percebeu-se que quando o falante se dá conta de que a produção do discurso possui algum problema, ele poderá interrompê-lo de duas maneiras diferentes: de modo explícito (*overt*), por exemplo, “*Não vêm a procura de servimos os estrangeiros com comidas... com molhos lá ao modo deles*” (Moniz, 2006) e de modo não explícito (*covert*), por exemplo, “*disseram que era apenas para não haver tanto barulho na, nos corredores*” (Moniz, 2006). Os dois exemplos clarificam perfeitamente as definições *overt* e *covert* de Levelt (1989). Enquanto que, no primeiro exemplo, o elemento problemático é produzido, no segundo exemplo, talvez somente o contexto possa auxiliar na tarefa de descobrir que expressão problemática seria produzida. Segundo Levelt (1989), as autocorrekções explícitas

ocorrem mais frequentemente do que as implícitas (75% x 25%, respectivamente, in Moniz, 2006).

Não somente isso, o falante, dependendo do tempo em que ele leva para perceber a inadequação da mensagem, pode interromper sua fala dentro da própria palavra problemática (situação que ocorreu em 18% dos casos na pesquisa de Levelt, 1989), logo a seguir ao item problemático (51% dos casos) e uma ou duas mais palavras depois da palavra problemática (31% dos casos). Essa possibilidade de interromper dentro da própria palavra, sem respeitar fronteiras sintáticas, palavras ou sílabas, é evidência forte em relação à hipótese do automonitoramento proposto por Levelt (1989). Entretanto, essas interrupções ocorrem em autocorreções de erros e não de especificações. Em erros de especificações, o falante tende a um único comportamento: completa as palavras problemáticas, antes de corrigi-las.

Quando se trata da reparação do erro, Levelt (1989) expõe características comuns aos falantes. Eles inserem em seu discurso, pausas preenchidas e marcadores de edição como “*er*”, “*rather*”, “*no*”, “*thus*”, “*I mean*” (expressões em português como, por exemplo, “*ééé*”, “*ããã*” para pausas preenchidas e “*isto é*”, “*quer dizer*” para marcadores de edição). Se comparada à interrupção, que não respeita fronteiras de constituintes em qualquer nível, a reparação ocorre obedecendo a várias restrições linguísticas.

Em nível sintático, a correção “respeita as fronteiras de frase ou de constituintes e a regra de boa formação frásica” (Moniz, 2006), como por exemplo, “*straight on to green – to red*” (Levelt, 1989). A autocorreção, nesse caso, foi sintaticamente bem formada, uma vez que o constituinte foi respeitado e reconstituído. Segundo Levelt (1989), sentenças como a do exemplo são evidências de que as autocorreções são processos sintáticos regulares que seguem regras de coordenação sintática. Importante ressaltar que o constituinte original e o constituinte referente à correção “são justapostos de uma forma idêntica à que se observa nas estruturas coordenadas” (Moniz, 2006) como, no exemplo “*straight on to green or to red*” (Levelt, 1989).

Outra informação importante se encontra no fato de o falante, de acordo com Levelt (1989), utilizam variadas estratégias de autocorreção, podendo ser instantâneas, antecipatórias ou reestruturadas. Em relação às estratégias instantâneas, a autocorreção começará sempre com a palavra que toma o lugar do elemento inadequado, como no exemplo “*what things are this kid – is this kid going to say incorrectly?*” (Levelt, 1989) em que o trecho que corresponde à correção é iniciado por “*is*” que substitui o item inadequado “*are*”. Já em

relação às estratégias antecipatórias, o falante faz uma recuperação de palavras anteriores ao erro, como no exemplo “*straight on to green – to red*” (Levelt, 1989) em que “*to*”, apesar de não ser o elemento problemático, é recuperado na sentença que corresponde à correção. Quanto às estratégias reestruturadas, as autocorreções feitas pelos falantes apresentam material linguístico que não fazia parte da frase interrompida, como no exemplo “*from yellow down to brown – no – that’s red*” (Levelt, 1989) em que, além da autocorreção, ainda apresenta uma clarificação maior da mensagem ao ouvinte. Aliás, essa última estratégia é o que Levelt (1989) afirma ser, agora de modo mais claro, de *especificações*.

Embora o falante utilize diversas estratégias de autocorreção, é necessário saber se o ouvinte as compreende e as integra aos constituintes originais. Segundo Levelt (1989), a primeira palavra da autocorreção é pista primordial para que haja compreensão por parte do ouvinte. Em casos como das estratégias de autocorrekções antecipatórias, o ouvinte interpreta a mensagem como uma continuação a partir da palavra que está sendo repetida. Quanto a estratégias instantâneas, quando o ouvinte percebe que a categoria sintática da primeira palavra da autocorreção é a mesma do constituinte original, ele a interpretará como uma continuação a partir daquela categoria.

Tais estratégias de autocorreção, quando comparadas entre as línguas, apresentam semelhanças, principalmente ao que concerne ao modo dessas estratégias serem efetuadas. Desse modo, os processos de autocorrekções parecem seguir princípios universais, sendo assim, passíveis de modelização. De acordo com o modelo de automonitoramento de Levelt, as mensagens seriam criadas no Conceptualizador, o Formulador seria o responsável em dar forma à mensagem pretendida e o Articulador selecionaria os músculos necessários para produção da mensagem. Inserido no Conceptualizador estaria o Monitor, responsável por controlar a produção discursiva e detectar e corrigir os possíveis erros. Por isso é que se afirma, mais uma vez, que o falante consegue monitorar o discurso, antes mesmo de produzi-lo (*Internal Loop*) ou após a produção (*External Loop*).

Logicamente, é compreensível que o modelo de Levelt (1989) não contemple explicações mais claras sobre como apenas um único mecanismo de retorno interno pode explicar porque há correção de determinados tipos de erros mais frequentemente do que outros. Mesmo assim, não se pode negar que a contribuição de Levelt (1989) é o ponto de partida, mesmo em seus pontos controversos, para os estudos das disfluências, principalmente ao que concerne ao lançamento de bases metodológicas para o tratamento das disfluências em

três pontos de vistas distintos, entretanto, complementares: a detecção, a interrupção e a reparação.

Com a apresentação do modelo de Levelt (1989) e a possibilidade de modelização das disfluências, parece-nos ser o próximo passo a verificação de uma pesquisa que concentra seus esforços em busca da modelização das disfluências. É essa a preocupação primordial de Shriberg (1994), próximo texto a ser analisado.

2.2 Modelização das disfluências em Shriberg (1994)

O estudo de Levelt (1989) foi importante na construção de reflexões acerca do monitoramento interno da fala. Com esse estudo também se comprovou a existência das disfluências e a sua relação direta com o automonitoramento da fala. Tal estudo não contemplou a modelização das disfluências, algo compreensível, uma vez que o foco não era esse. Eis o foco dessa seção, verificar estudos referentes à modelização das disfluências.

Partindo dos estudos de Levelt (1989), Shriberg (1994) elaborou um sistema de anotação das disfluências simulando o que julga ser verdadeiramente o comportamento dos ouvintes diante de uma sentença disfluente. Para tal, a autora analisou conversas telefônicas entre estranhos do corpus *Switchboard*. Tal corpus consiste em gravações telefônicas de conversas de indivíduos, que não se conheciam, sobre um tópico pré-determinado. Após a análise desse corpus, várias contribuições foram feitas por Shriberg (1994), entretanto, para fins dessa pesquisa, a principal contribuição da autora foi a criação do sistema de anotação das disfluências, que contribuiu para a generalização da adoção de um sistema de anotação e da terminologia associada a tal sistema.

Essencialmente, as análises de Shriberg (1994) consistem em verificar que a zona a ser reparada pelo falante corresponde à sequência do enunciado que deve ser apagada para que, dessa maneira, a mensagem faça sentido. Afinal, o que deve ser realmente apagado é o material disfluente. Desse modo é possível simular o comportamento do ouvinte. Devido a esse modo de analisar a disfluência, Shriberg (1994) procurou, então, criar uma terminologia, que pudesse dar conta, de forma neutra, dos fenômenos disfluentes e que permitisse fazer comparações entre as demais línguas, além de, logicamente, facilitar a compreensão dos fenômenos disfluentes e enquadrá-los em uma escala maior. Devido a isso, a autora levantou oito categorias de fenômenos disfluentes:

TABELA 1: Tipologia das disfluências de Shriberg (1994).

DISFLUÊNCIA	DEFINIÇÃO	EXEMPLO
Pausas preenchidas	Vocalizações ortografadas que preenchem um espaço de silêncio na comunicação	<i>“esse túnel atolado <u>aa</u> fica um bocadinho antes da estrada das minas antigas”</i> (Moniz, 2006)
Alongamentos disfluentes	Materiais linguísticos alongados que podem ocorrer tanto em palavras funcionais, como em lexicais	<i>“ela estava a estudar mais as amigas no quarto – <u>e...</u> cheguei lá estavam três senhoras”</i> (Moniz, 2006)
Repetições	Remetem à repetição de material linguístico	<i>“porque o castigo das suspensões <u>não vai</u>, não vai educar o aluno nesse sentido”</i> (Moniz, 2006)
Fragmentações	Consiste no fato de o material linguístico ter sido fragmentado	<i>“enfim por <u>vá(...)</u>, vários motivos”</i> (Moniz, 2006)
Substituições	Aterações de segmentos ou palavras com correspondência sintática e/ou semântica com a zona a corrigir	<i>“<u>o inquerito</u>, os inqueritos forma feitos há cerca de vinte anos”</i> (Moniz, 2006)
Apagamentos	Abandono por completo do material linguístico produzido	<i>“[...] aliás, uma das coisas graves que me parecem que têm, duma maneira geral, os latinos, sobretudo do sul da Europa e portugueses, espanhóis e sua da Itália – e eu não</i>

		<i>falo na, nos gregos, porque os gregos são(...), não conheço a Grécia.”</i> (Moniz, 2006)
Inserções	Correspondem a acréscimos de material linguístico	“ <i>uma época, uma longa época, digamos, de escassa iniciativa</i> ” (Moniz, 2006)
Marcadores de edição	Correspondem a vocábulo ou expressões que transmitem informação de correção de um item ou de uma sequência produzida, como <u>melhor</u> , <u>quer dizer</u> , <u>não</u> , etc	“ <i>é, é, quer dizer, é a lei do castigo</i> ” (Moniz, 2006)

Todas essas disfluências já haviam sido analisadas por alguns autores anteriormente, algumas com mais ênfase, outras com pouco entusiasmo. Uma dessas disfluências, largamente analisada, é a pausa preenchida. Levelt (1983; 1989), Clark & Fox Tree (2002), Mata (1999), Zhao & Jurafsky (2005), O’ Shaughnessy (1992); Shriberg (1994; 1998); Eklund (2001; 2004), Eklund & Shriberg (1998), Brennan (2000) e Brennan & Schober (2001) (in Moniz, 2006), discutiram exaustivamente sobre as pausas preenchidas e chegaram a algumas conclusões bastante pertinentes: as pausas preenchidas são bem mais longas do que as mesmas vogais, quando elas estão em contextos fluentes, ou até mesmo mais longas do que os alongamentos disfluentes. Analisando as línguas, as pausas preenchidas têm duração que parece ser um traço que não depende da língua específica.

Outro dado importante, visto por Eklund & Shriberg (1998 in Moniz, 2006), é a localização de ocorrência da pausa preenchida. Esses autores verificaram que as pausas preenchidas ocorrem mais frequentemente em início de uma frase do que no meio. Eles afirmam que, quando ela ocorre em início de frase, estaria havendo, por parte do falante, um planejamento global do enunciado. Quando há a ocorrência em meio à frase, ocorreria, então, uma procura lexical.

Quando as pausas preenchidas são analisadas pelo viés da percepção, a afirmação a que se chega é a de que as pausas preenchidas são como mecanismos que auxiliam ou preparam o ouvinte para sequências discursivas complexas ou longas, ou seja, uma espécie de alarme com intuito de avisar ao ouvinte que ele necessita de mais atenção ou concentração em relação à mensagem.

Em relação aos alongamentos disfluentes, algumas observações podem ser tecidas. Alongamentos disfluentes e pausas preenchidas se confundem por possuírem traços em comum, entretanto, devido a algumas propriedades, podem ser distinguidos. Os alongamentos são mais curtos do que as pausas e ocorrem, na maioria das línguas, em final de palavra, diferentemente das pausas que frequentemente ocorrem no início de constituinte e antes de uma outra disfluência.

Quando a discussão se direciona para as repetições, Levelt (1989) e Clark e Wasow (1998) as consideram formas eficientes de continuação da mensagem. Heike (1981) e Shriberg (1995; 2001) (in Moniz, 2006) afirmam, de forma geral, que as repetições podem ser estratégias de ganho de tempo para preparar o que vem a seguir na mensagem, tanto em nível lexical como sintático. Dessa maneira, pausas preenchidas e repetições são disfluências que se assemelham quanto a suas características: estão associadas a planejamento das estruturas que ainda serão produzidas.

Mesmo pouco explorada (Nakatani ; Hirschberg, 1994; Pallaud 2002; 2003 e Henry & Pallaud, 2003 in Moniz, 2006), a disfluência denominada fragmentação é um elemento com características contrárias às repetições. A ocorrência da fragmentação é uma espécie de pista na detecção de uma correção. Além disso, afeta mais frequentemente palavras lexicais e sintaticamente pode ocupar posições de sujeito, verbo e objeto, refletindo, dessa forma, uma procura lexical exatamente no domínio sintático específico.

Substituições, apagamentos, inserções e marcadores de edição são disfluências que ainda não receberam vasta discussão. Shriberg (1994) e Levelt (1989) foram os autores que teceram as principais contribuições sobre tais disfluências. Geralmente, de acordo com os autores, as substituições se limitam a mesma categoria sintática do elemento corrigido. Em relação aos apagamentos, nada mais é acrescentado do que a sua relação com a reestruturação sintática. Já as inserções e os marcadores de edição, além de uma frequência de ocorrência muito baixa – o que não permite uma clareza maior em relação ao reconhecimento das autocorreções –, quando ocorrem isoladas, a sequência das palavras é identificada

corretamente pelos modelos de língua, quando juntas a outras disfluências, formam sequências complexas, tornando tais sequências objeto das análises.

Embora haja muitas contribuições em relação às disfluências, voltamos a bater na mesma tecla: a importância de uma terminologia para dar conta das disfluências. Como visto, o trabalho de Shriberg (1994) tem grande valor por conta dessa necessidade terminológica. Tal terminologia foi aplicada e vista em variadas línguas como o inglês, o sueco e o mandarim, como os exemplos revelaram. Verificar o fenômeno das disfluências, fazendo uso dessas terminologias, na língua portuguesa parece ser um caminho interessante a ser seguido. Eis, então, o próximo passo, apresentamos uma pesquisa em língua portuguesa sobre disfluências a qual adota as terminologias de Shriberg (1994), tópico que será tratado na próxima seção.

2.3 Disfluências em uma análise na língua portuguesa (Moniz, 2006)

Consultando a base de dados do trabalho de Mata (1995, 1999) *Corpus de Português Europeu Falado por Adolescentes em Contexto Escolar* (CPE FACES), Moniz (2006) utilizou a gravação de 2 horas, 10 minutos e 3 segundos de exposições preparadas e espontâneas de uma professora (MA) e de quatro alunos (AA, SN, AX e PD) de uma aula de língua portuguesa em Portugal.

Os alunos tinham que apresentar *O Velho e o Mar* de Ernest Hemingway ou o *Auto da Índia* de Gil Vicente. Essas apresentações representam as falas preparadas dos alunos. A professora deu uma aula sobre Camões e a contextualização histórico-literária de *Os Lusíadas* como forma de caracterizar sua fala preparada. Para detecção da fala espontânea dos participantes, foram pedidos a eles relatos de uma experiência pessoal agradável. Do tempo inicial de gravação, Moniz (2006) separou 1 hora, 11 minutos e 16 segundos de tempo de exposição dos falantes sem interrupções ou sobreposições. Utilizando o programa *wavesurfer* (<http://www.speech.kth.se/wavesurfer/>), a autora transliterou e alinhou o sinal acústico em diferentes níveis. Já para as anotações, foram feitas três fileiras, correspondendo, cada uma, respectivamente, à identificação do falante, a anotação ortográfica e a descrição das disfluências (ver figura a seguir):

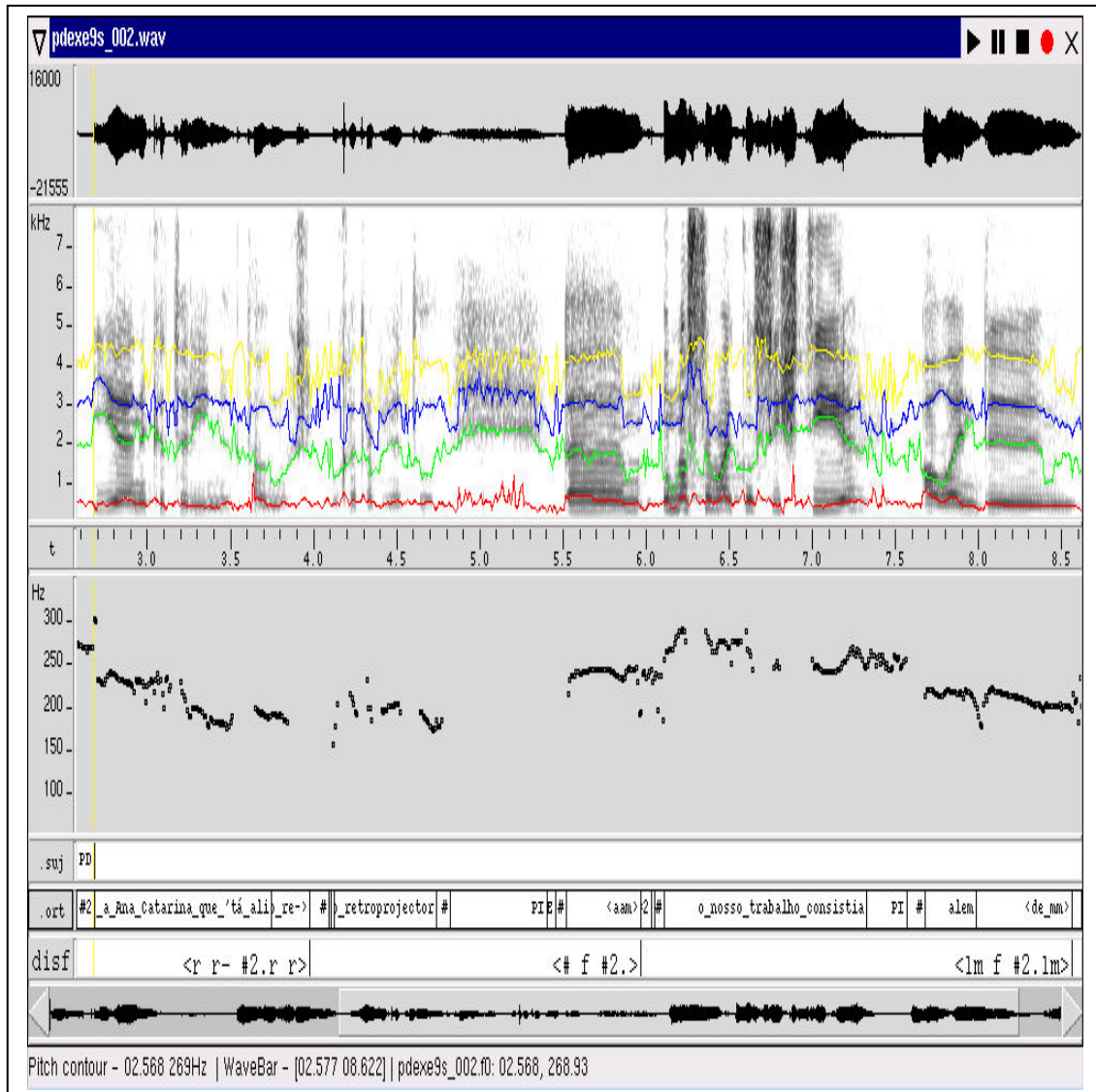


FIGURA 3: Exemplo de um arquivo gravado pelo programa *wavesurfer* (Moniz, 2006).

De acordo com Moniz, na fileira de identificação do sujeito é inserido o nome do falante, sendo que o início da tomada da palavra é identificado por dois pontos. A fileira ortográfica corresponde a seguinte sequência: “#2 e a Ana Catarina que ‘tá ali <no re-> # CLIQUE # no retroprojector # P1 CLIQUE # <aam> # 2 CLIQUE # o nosso trabalho consistia PI # além de muitas outras propostas” (MONIZ, 2006). Nessa fileira, apesar da aparição de pausas e sequências disfluente, não ocorre uma fileira miscelânea, pois, de acordo com Moniz (2006), “uma vez que, com base nas informações introduzidas, é possível construir automaticamente as fiadas palavra a palavra e miscelânea” (MONIZ, 2006). Na terceira fileira, houve um preenchimento de acordo com as anotações de Shriberg (1994), as quais consistem em classes de disfluência que são resultados do mapeamento da zona a

reparar e a reparação, exceto as pausas preenchidas e as expressões de edição que não são fruto das duas zonas citadas, mas são produzidas no momento de edição do discurso.

Em relação às disfluências catalogadas por Moniz (2006), algumas pequenas diferenças foram feitas se comparados aos originais de Shriberg (1994). As categorias pausas preenchidas, alongamentos disfluentes, repetições, substituições, apagamentos, inserções e marcadores de edição (ver tabela 1) foram mantidas. As distinções se encontram em duas categorias com nomes novos, mas com funções similares a outras. Por exemplo, Moniz (2006) categoriza como disfluências as palavras mal pronunciadas e o truncamento de palavras.

As palavras mal pronunciadas ocorrem quando alguma palavra não é produzida corretamente, por exemplo, “*metarmorfo- metamorfose do homem*” (Moniz, 2006). No exemplo, há a má produção inicial da palavra e logo substituída pela pronúncia correta. Como dito, há uma substituição, por conta disso, a própria autora revela que os casos de palavras mal pronunciadas são incorporados aos casos de substituição em sua pesquisa.

O truncamento de palavras, segundo a autora, consiste na palavra cuja produção é interrompida, por exemplo, “*huma- humanidade vem de quê*” (Moniz, 2006). De acordo com Moniz (2006), essa categoria ocorre em simultâneo ao fragmento e incorporada nos casos de repetições.

Todo o tempo de gravação foi dividido em 11,04% de sequências disfluentes, 65,11% de sequências fluentes e 23,85% de tempo de pausas. Na tabela a seguir, Moniz (2006) relata a frequência de ocorrência das categorias disfluentes no interior de uma sequência e a frequência da categoria de forma isolada ou com outro da mesma categoria.

TABELA 2: Frequência de ocorrência de cada item disfluyente no interior de uma sequência e frequência de ocorrência de um fenômeno isolado ou com outro da mesma categoria (Moniz, 2006).

Categoria disfluyente	Frequência de ocorrência do item disfluyente no interior de uma sequência	Frequência de ocorrência de um fenômeno isolado ou com outro da mesma categoria
Alongamentos	497 – 31,68%	288 – 45,93%
Repetições	485 – 30,85%	94 – 14,99%
Pausas preenchidas	274 – 17,53%	198 – 31,58%
Substituições	177 – 11,28%	23 – 3,67%
Apagamentos	112 – 7,14%	24 – 3,83%
Marcadores de edição	20 – 1,27%	0 – 0%
Inserções	4 – 0,25%	0 – 0%
Total	1569 – 100%	627 – 100%

Esses dados revelam que os alongamentos disfluents e as pausas preenchidas ocorrem, de forma mais expressiva, sem outras disfluências do que as outras categorias, que são produzidas mais significativamente em sequências complexas. Além disso, três categorias disfluents ocorrem mais vezes do que as outras: alongamentos, repetições e pausas preenchidas. Todas essas categorias apontam uma tendência do *corpus* à seleção de classe das disfluências associadas a planejamento discursivo.

Em pesquisas como de Shriberg (1994), Candea (2000) e Eklund (2004), as pausas preenchidas apareceram como categoria mais frequente, tendência essa que não se confirmou na pesquisa de Moniz (2006). Os dados indicam uma estigmatização das pausas preenchidas em contexto escolar e preferência por palavras, mesmo alongadas ou repetidas, que permitam, ao falante, uma preparação das unidades subsequentes. Uma das preocupações de Moniz (2006) foi a análise dos truncamentos em material linguístico (ver tabela a seguir):

TABELA 3: Porcentagens de cada tipo de fenômeno em função de ser ou não truncado (Moniz, 2006).

Categoria disfluente	Truncada %	Não truncada %	Total no <i>corpus</i> %
Alongamentos	0,83	30,85	31,68
Repetições	5,48	25,37	30,85
Pausas preenchidas	0	17,53	17,53
Substituições	1,15	10,13	11,28
Apagamentos	0,76	6,37	7,13
Marcadores de edição	0	1,27	1,27
Inserções	0	0,25	0,25
Total	8,22	91,78	100

Os dados apontam para 8,22% das disfluências serem truncados e 91,78% não, como, por exemplo, nas ocorrências dos seguintes casos: “*Vocês acham que são impor- importantes para caracterizar o Renascimento e o Humanismo*” (retrata a repetição truncada) e “*acho que el- a Xana tinha um relógio*” (corresponde a substituição truncada).

As repetições é a categoria com maior porcentagem de truncamentos. Uma hipótese para tal fator, segundo a autora, é de que esses segmentos estariam a favor de uma função de suporte vocálico para construção de uma mensagem ou confirmação da seleção lexical feita.

Além disso, outra informação surge com bastante importância, os truncamentos não ocorrem desordenadamente. Shriberg (1994) verificou que há truncamentos em categorias disfluente que ocorrem na posição medial, descartando aquelas de início de frase. Não somente isso, as palavras lexicais, em posição medial, estão mais sujeitas a sofrerem fragmentações do que as palavras funcionais. De certa forma, é uma evidência de que os truncamentos dependem de vários fatores como, por exemplo, tarefa, *corpus*, agentes comunicativos e localização nas frases. As tabelas a seguir refletem dados de categorias disfluente truncadas e não truncadas em relatos espontâneos não espontâneos:

TABELA 4: Porcentagens de categorias disfluente truncadas e não truncadas na situação RLC (Relato espontâneo) para todos os falantes (Moniz, 2006)

Categoria disfluente	Truncada %	Não truncada %	Total no <i>corpus</i> %
Alongamentos	1,53	24,81	26,34
Repetições	7,25	30,92	38,17
Pausas preenchidas	0	15,25	15,25
Substituições	1,53	15,65	17,18
Apagamentos	0,76	1,53	2,29
Marcadores de edição	0	0	0
Inserções	0	0,76	0,76
Total	11,07	88,93	100

TABELA 5: Porcentagens de categorias disfluente truncadas e não truncadas na situação EXE (Exposição preparada) para todos os falantes (Moniz, 2006)

Categoria disfluente	Truncada %	Não truncada %	Total no <i>corpus</i> %
Alongamentos	0,69	32,06	32,75
Repetições	5,13	24,25	29,38
Pausas preenchidas	0	17,98	17,98
Substituições	1,07	9,03	10,10
Apagamentos	0,77	7,34	8,11
Marcadores de edição	0	1,53	1,53
Inserções	0	0,15	0,15
Total	7,66	92,35	100

Observando os dados, constata-se que o truncamento é mais produtiva quando em situação de relato espontâneo (11,07%) em face da exposição preparada (7,66%). A quantidade considerável de produção de disfluências na exposição preparada é argumento forte favorável ao uso desses fenômenos como automonitoração e autocorreção da mensagem. Essa afirmação faz sentido, uma vez que é nessa situação que o falante tem, realmente, de ajustar ou adaptar sua mensagem para que haja a clara e correta compreensão daquilo que ele quer transmitir.

Quando se observam os dados de Moniz (2006) em relação às pausas preenchidas, algumas constatações podem ser feitas. Em primeiro lugar, as pausas preenchidas que mais ocorrem no *corpus* são as formas *aa*, *aam* e *mm*. Contabilizando todas as pausas, verifica-se que *aa* ocorre em 78,47%, *aam* 17,15% e *mm* 4,38% no total do *corpus*. Quando isolados, ou seja, quando há a ocorrência das pausas preenchidas por si, há ocorrência de 79,7% de *aa* e 20,3% de *aam*. *Mm* nunca tem produção isolada, pois só ocorre no final de alongamentos disfluentes.

Se a análise se concentrar na ocorrência das pausas preenchidas em situação relato espontâneo e exposição preparada, há a porcentagem de 14,76% e 85,24% para cada um, respectivamente. O alto índice de pausas preenchidas na exposição preparada denota uma grande preocupação de monitorização dos falantes. As tabelas a seguir destacam informações bastante pertinentes sobre as localizações prosódicas das pausas preenchidas, tanto em relato espontâneo, como em exposição preparada:

TABELA 6: Porcentagens das localizações prosódicas dos diferentes tipos de pausas preenchidas na situação relato espontâneo (Moniz, 2006)

Tipos de pausas	Dentro de constituinte %	Após constituinte menor %	Após constituinte maior %	Total %
<i>Aa</i>	25	20	50	95
<i>Aam</i>	0	2,5	2,5	5
<i>Mm</i>	0	0	0	0
Total	25	22,5	52,5	100

TABELA 7: Porcentagens das localizações prosódicas dos diferentes tipos de pausas preenchidas na situação exposição preparada (Moniz, 2006)

Tipos de pausas	Dentro de constituinte %	Após constituinte menor %	Após constituinte maior %	Total %
<i>Aa</i>	15,81	11,97	47,86	75,64
<i>Aam</i>	0	1,28	17,95	19,23
<i>Mm</i>	1,28	1,28	2,56	5,13
Total	17,09	14,53	68,38	100

Nas duas situações, *aa* é a pausa preenchida que tem maior ocorrência, *aam* é pouco representativo na situação relato espontâneo e *mm* sequer tem produção nessa situação. De certo, as várias pausas preenchidas revelam diferentes graus de planejamento. O alongamento disfluyente quando ocorre com a pausa preenchida *mm* desempenha a função de procura lexical. Em contrapartida, *aa* e *aam* são pausas preenchidas que estão à disposição para planejamentos de constituintes maiores e mais complexos.

Quando o objeto de análise são os alongamentos disfluents, verifica-se que 58,22% são produzidos em palavras funcionais e 41,78% em palavras lexicais. Em relação à localização, os dados sugerem que alongamentos isolados ou no início de uma sequência podem estar associados ao planejamento “de unidades discursivas e, no interior da sequência, a micro unidades, como a procura de uma palavra específica para aquele contexto” (Moniz, 2006). Como alguns dados já indicaram, algumas disfluências ocorrem simultaneamente em uma mensagem. Os gráficos seguintes representam a ocorrência de combinatórias disfluents.

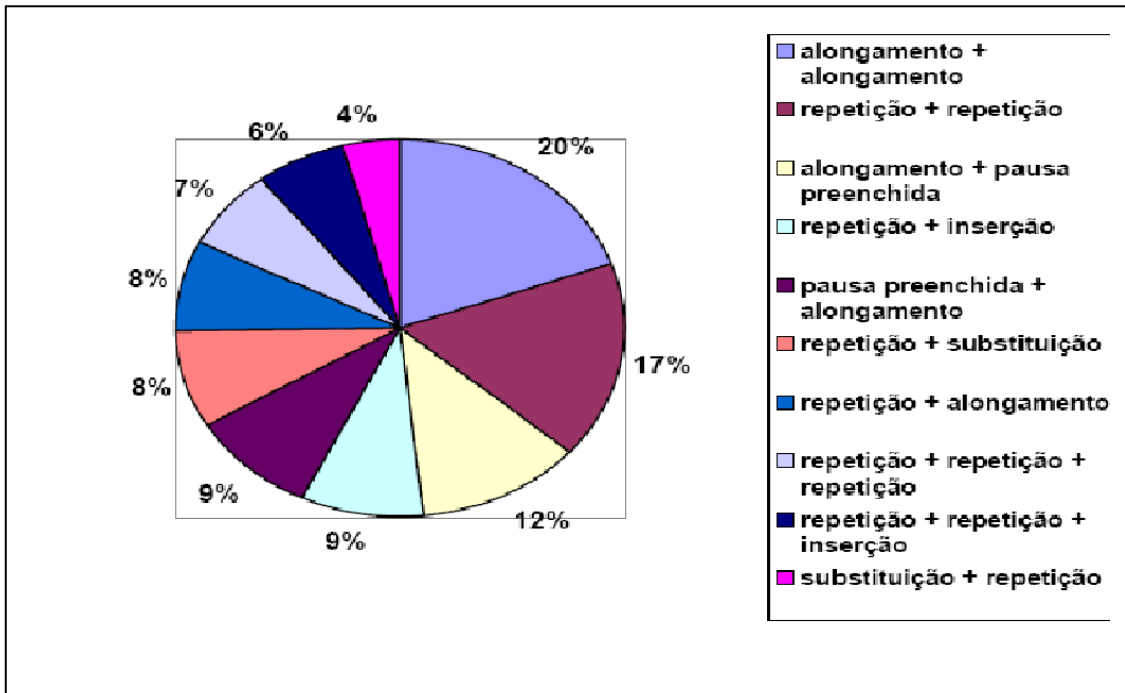


GRÁFICO 1: Porcentagens de categorias disfluentes mais frequentes no *corpus*, considerando a sua ordem de produção e a ocorrência com a mesma categoria (Moniz, 2006).

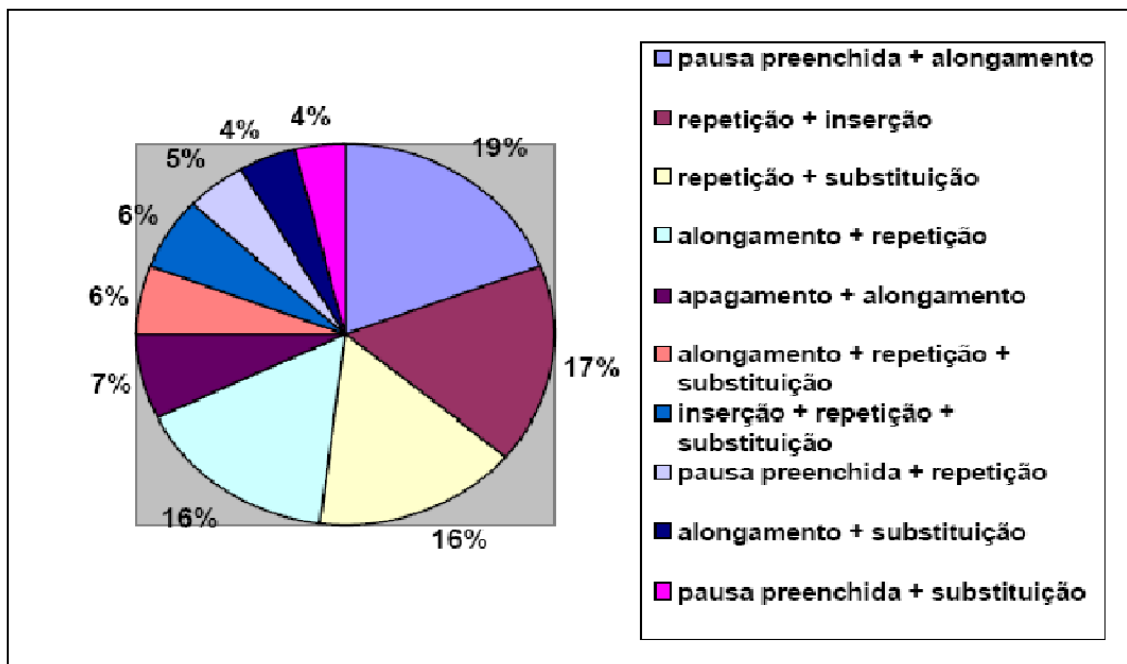


GRÁFICO 2: Porcentagens das dez combinatórias mais frequentes que ocorrem no *corpus* independentemente da sua posição na sequência disfluyente e de se combinarem com a mesma categoria (Moniz, 2006).

Os dados apontam para a maior ocorrência dos alongamentos disfluêntes, inclusive quando combinados com eles próprios (20% em maior número de vezes). A partir desses dados, algumas hipóteses podem ser esclarecidas, segundo Moniz (2006), o contexto escolar, devido a sua especificidade, parece influenciar a escolha das estratégias pelos falantes. Preocupados em cumprir as tarefas solicitadas, os falantes tem um compromisso maior com a mensagem. Desse modo, se esforçam para adaptar seu discurso aos parâmetros de avaliação, os quais prezam a clareza e a correção, usando, assim estratégias simultâneas de vocalização e duração.

Um último dado a ser analisado, entre outros, é a riqueza vocabular. Alguns estudos cogitam a hipótese de que, quanto mais rico o vocabulário do falante, maior a probabilidade de o falante utilizar certos tipos de disfluências. Compare com a figura que traz o número de vocábulos diferentes produzidos por falante e por situação.

% de palavras	AA		SN		AX		PD		MA	
	RLC	EXE	RLC	EXE	RLC	EXE	RLC	EXE	RLC	EXE
Disfluêntes	10,32	14,79	10,24	16,58	19,75	16,64	11,59	13,51	9,09	11,83
Fluêntes	89,68	85,21	89,76	83,42	80,25	83,36	88,41	86,49	90,91	88,17

FIGURA 4: Porcentagens de palavras disfluêntes e fluêntes (Moniz, 2006).

O *corpus* aponta para uma riqueza vocabular da professora MA, seguido, de longe, pelos alunos SN e PD. Em relação às disfluências, a professora é quem utiliza menor número de vocábulos disfluêntes e quem se aproxima mais é o aluno PD. Com essa base vista, agora é momento de aprofundar ainda mais as questões sobre as disfluências. Verificar disfluência em um contexto complexo é o objetivo. No entanto, esclarecer o que é realmente considerado

complexidade pode ser um obstáculo. Por isso, a próxima seção tem, como principal tema, a complexidade dos constituintes.

2.4 Conceito de complexidade gramatical em Clark e Wasow (1998)

Diante da sentença “*yes, I uh I wouldn’t be surprised at that, – – I really wouldn’t*” (Clark e Wasow, 1998) (“Sim, eu uh eu não seria surpreendido, [silêncio] eu realmente não seria”), não há motivos, a princípio, para que o falante repita as palavras. Clark e Wasow (1998) buscam, diante disso, investigar a origem das palavras repetidas. Consideradas por dois pontos de vista, as disfluências são definidas como um processo puro (tratadas como resultados de um processo que, uma vez iniciadas, correm sem intervenção) ou um processo estratégico (vistas como resultado de determinadas estratégias). Apesar de distintas, as duas visões se complementam, sendo que sua diferença se estabelece na recolha do material. Por isso, os autores focam suas atenções no processo estratégico, sem perder o olhar no processo puro.

Baseados em falas espontâneas de língua inglesa americana e britânica, Clark e Wasow (1998) apontam quatro fases para ocorrência das repetições: compromisso inicial, suspensão da fala, hiato e reinício do constituinte. O compromisso inicial pode ser entendido como a produção de uma primeira palavra o mais cedo possível, comprometendo, dessa forma, o constituinte, por exemplo, “*yes, I uh I wouldn’t be surprised at that, – – I really wouldn’t*” (Clark e Wasow, 1998). Já a suspensão da fala é o momento em que o falante suspende a fala em qualquer ponto de um enunciado, por exemplo, “*because you see I {- uh} some of our people, {. (clears throat)} who are doing LEs, {- - u:m} have to consider which paper {.} to do*” (Clark e Wasow, 1998). Os espaços vazios entre a suspensão da fala e a retomada são considerados hiatos, que não estão vinculados às repetições. A retomada do discurso a partir do último constituinte falado é o que se pode denominar de reinício do constituinte. Somente quando todas essas fases acontecem, ocorre, então, uma repetição.

Embora as repetições sejam estabelecidas nessas condições, ainda é necessário verificar quando o falante produz uma sentença considerada complexa. Por partes, os falantes, no momento da criação de um constituinte complexo, podem pausar seu discurso, porque podem levar mais tempo para criar planos de articulação nos constituintes complexos. Para que um constituinte, de acordo com os autores, seja mais complexo no nível sintático, maior

deverá ser o seu peso gramatical. O peso gramatical é entendido pela quantidade de informação expressa em um constituinte e, também, pode ser medido pelo número de palavras, nós sintáticos ou nós frasais no componente. Embora Clark e Wasow tenham pesquisado sobre a complexidade gramatical, uma ressalva se torna bastante importante: nada ainda indica quantos nós e quais nós gerariam complexidade. Além disso, também não está claro o quanto de informação é o suficiente para que haja um constituinte complexo.

Esclarecido o contexto da complexidade, é momento de verificar uma camada da árvore sintática também considerada complexa: a camada CP. Por conta de sua complexidade, as disfluências são esperadas. Basta agora verificar detalhadamente como isso ocorre em pesquisas focadas nesse tema.

2.5 Disfluências na camada CP

A camada CP, como já dito em momentos anteriores, revela ter uma grande complexidade. Sendo complexo, se espera disfluências. Foi o que Horne et al (2005) e Jaeger (2005) consideraram em suas pesquisas. Mesmo tratando de línguas distintas, esses pesquisadores verificaram as disfluências que ocorrem com relativizadores e os momentos de sua produção. Por serem pesquisas importantes para essa dissertação, esta seção será subdividida para que as duas pesquisas ganhem o destaque merecido. Na seção 2.5.1, haverá a exposição do texto de Horne et al (2005) e, na seção 2.5.2, a exposição do texto de Jaeger (2005)

2.5.1 Dialeto *Orust*: análise de Horne et al (2005)

Horne et al (2005) verificam as conjunções *att* e *och* do *Orust*, um dialeto sueco, equivalentes ao *that* em língua inglesa, em produções fluentes e disfluentes e os fragmentos do discurso que os seguem. Assumindo o *Phonological Loop* de Baddeley (1997) – subsistema na memória de trabalho associada ao processo da fala (plano fonético e plano articulatório) –, os autores afirmam que as informações ficariam armazenadas durante dois segundos no *Phonological Loop*, mas que poderiam ser retidas por longos períodos, desde que sejam refrescadas pelo controle do processo articulatório. As restrições do tempo do *Phonological Loop* na memória de trabalho, segundo os autores, desempenham um papel

fundamental na produção da fala e isso reflete diretamente na fragmentação da fala. Devido a isso, as atenções de Horne et al (2005) se voltam para as correlações prosódicas, sintáticas e, até, pragmáticas de 2 a 2,5 segundos de unidades produzidas e codificadas no *Phonological Loop*.

Seguindo a afirmação de Clark e Wasow (1998, in Horne et al, 2005), segundo a qual as conjunções *att* e *och* ocorrem antes de constituintes maiores, Horne et al (2005) mediram a duração desses segmentos em contextos fluentes e disfluentes. Os autores obtiveram os seguintes dados:

- Duração longa de *att* e *och* antes de hesitações (cerca de 130 ms ou mais);
- Duração das vogais em *att* e *och* é maior antes das hesitações (cerca de 20 ms);
- Duração de 20 ms para *att* e 10 ms para *och* antes de hesitações em finais de sintagmas de aspiração;
- Duração de 90 ms de *att* e *och* antes de uma hesitação em sintagmas de paradas oclusivas.

Em suma, concluíram que *att* e *och* são marcados por características segmentais antes das hesitações disfluentes, em relação as suas formas em contextos fluentes. Por conta disso, algumas hipóteses são fortalecidas como, por exemplo, as formas fonéticas das palavras *att* e *och*, junto com a pausa, podem prover informações sobre os fragmentos do discurso após a disfluência. Com o apoio da hipótese da complexidade, um caso de *att* foneticamente marcado e que ocorre antes da disfluência deverá ser seguido por termos mais pesados ou mais complexos do que formas de *att* em contextos fluentes. Para exemplificar tais afirmações, Horne et al (2005) utilizam sentenças do *Orust* (traduzidas para o inglês) de uma pesquisa realizada em 2003:

a) ‘*there was a lot when we were at elite camp ATT PAUSE [CP then was [TP there more [ConjP [CP one had relaxing before] and [CP the trainer talked] and [CP it had to be dark in the room]]]]’*

b) ‘*it was really ATT PAUSE [CP [CP when one is around (PAUSE can it be) fifteen sixteen] then get [TP all clubs PAUSE in Bohuslândia [VP [DP ∅]]]]’*

- c) *'just like not Autonova ATT PAUSE [CP there know [TP I [CP what happens]]]'*
- d) *'ATT INHALATION [CP I want [vP to be able [ConjP [VP to go to work] and INHALATION [VP to know in principle [CP what is happening]]]]] maybe'*
- e) *'but nevertheless ATT PAUSE [CP it can happen [QP så much [CP that you do not know [CP what you have to do or what SWALLOW PAUSE happens just then]]]]'*

(Horne et al, 2005)

Valorizando apenas a abordagem da estrutura sintática, em detrimento da estrutura pragmática, as sentenças demonstram que *att* disfluyente funciona como um sinal de que, no discurso, um fragmento complexo está sendo planejado. Devido a isso, orações fluentes dentro do fragmento complexo são esperadas, uma vez que os falantes tiveram tempo suficiente para planejar a oração complexa durante a hesitação. Os dados também demonstram que os fragmentos que seguem o *att* disfluyente parecem ter sido planejados como uma unidade, desde que tenham sido articulados durante um tempo de 2-2,5s.

2.5.2 Omissão de *that* na língua inglesa: análise de Jaeger (2005)

Seguindo, basicamente, a mesma linha de raciocínio de Horne et al (2005), Jaeger (2005) afirma que a omissão do relativizador *that* em orações complementos ou em orações relativas em que *that* não seja sujeito (NSRC, non-subject-extracted relative clauses, Jaeger, 2005), na língua inglesa, está correlacionada à dificuldade de produção e, conseqüentemente, uma evidência para disfluências. Para chegar a tal conclusão, algumas análises anteriores foram feitas. Em primeiro lugar, Jaeger (2005) discute se a presença ou não do pronome relativo *that* tem alguma função específica, como, por exemplo, a facilitação da compreensão por conta do destinatário. Nos exemplos *"I believe (that) my brother stole your bike"* (oração complemento) e *"I mean everything (that) they spray out in the fields"* (oração relativa), a omissão de *that* no primeiro exemplo criaria, em um primeiro momento, uma sentença ambígua. No entanto, nem toda sentença, em que há omissão de *that*, resultará em uma sentença que abriga ambigüidade.

Partindo disso, outras omissões de *that* são verificadas. De acordo com o autor e apoiado em Ferreira e Dell (in Jaeger, 2005), a omissão de *that* ocorre quando o material que o segue está prontamente disponível. Tal observação parece ser aceitável, uma vez que, pelo contrário, os falantes costumam não omitir *that* quando o sujeito encaixado é mais complexo e, portanto, levam mais tempo para planejar. Essas constatações levam à criação de duas hipóteses: *Signal Hypothesis* e *Alleviation Hypothesis*. Esta consiste em que o falante insere *that* para aliviar dificuldades de produção, ou seja, ganhar mais tempo para superar problemas na produção de orações encaixadas. Aquela tem como conceito a inserção de *that* (em contextos que ele pode ser suprimido) como um sinal de que o que será produzido traz dificuldades de produção ao emissor.

Apesar de parecerem ser muito próximas, as duas hipóteses diferem em relação às disfluências (tratadas como evidências sobre a dificuldade na produção). A *Alleviation Hypothesis* afirma que a menção de *that* no início de um complemento ou oração relativa reduz as disfluências na sentença. No entanto, a *Signal Hypothesis* afirma exatamente o oposto. Por causa dessa incerteza, Jaeger (2005) propõe testar ambas as hipóteses. Para tal, ele anota as disfluências de 650 pares de conversas telefônicas entre estranhos do corpus *Switchboard*. Contando todas as disfluências como, por exemplo, repetições e pausas, o corpus teve o total de 793 disfluências em orações relativas, sendo que desse total, 351 são as pausas preenchidas *uh/um* (61%) e *you know* (32%), como nos exemplos a seguir:

1)... *the nuclear [NSRC that, uh, they use] ...*

2) ... *things [NSRC that, you know, two people can do] ...*

3) ... *every time [NSRC I, uh, I spent money, I mean, cash] ...*

(Jaeger, 2005)

Entre algumas medidas, Jaeger (2005) controlou alguns aspectos da pesquisa para chegar a uma conclusão. Entre alguns métodos, ele normalizou a taxa de disfluência na oração relativa (número de palavra por palavra disfluente). Essa medida tem como previsão taxas de disfluências normalizadas para as orações relativas com um relativizador menores do que as

taxas de disfluências em orações relativas sem um relativizador, uma vez que, de acordo com a *Alleviation Hypothesis*, os relativizadores ajudam a aliviar a dificuldade de produção.

Após testes nas versões mais fracas e fortes da *Alleviation Hypothesis*, em linhas breves, Jaeger (2005) confirma que a presença do relativizador está relacionada à presença de disfluências na oração relativa, afirmação que vai de encontro a tal hipótese. Os relativizadores, de modo geral, não ajudam a aliviar a dificuldade de produção. Pelo contrário, a presença do relativizador correlacionada com as disfluências, mesmo depois do controle de alguns fatores, aponta para uma espécie de sinalização, fator que corrobora a *Signal Hypothesis*. Desse modo, se a presença dos relativizadores é significativa para uma taxa maior de disfluência, os ouvintes deveriam ser mais sensíveis a essa informação. No entanto, de acordo com Jaeger (2005), nada ainda comprova que os ouvintes interpretam a opção dos relativizadores como sinal que os falantes estão antecipando dificuldades na produção.

Após a demonstração de extensas e complexas pesquisas envolvendo disfluências em pronomes relativos nas línguas Orust (Horne et al, 2005) e inglesa (Jaeger, 2005), uma análise no português brasileiro é o próximo passo. Todavia, antes de tal verificação, é preciso observar sentenças com pronomes relativos no PB. Para isso, parece-nos importante uma esmerada revisão de uma pesquisa que, apesar de não tratar das disfluências, tem o custo de processamento na produção de orações relativas no PB como foco.

2.6 Orações relativas no PB vistas por Corrêa et al (2008)

Complexas e propícias às disfluências, as orações relativas já apontaram ser, durante boa parte da revisão de literatura. Apresentar mais evidências ou, pelo menos, contribuir para discussão, com dados do PB, é o objetivo. Entretanto, para alcançar tal meta, é necessário caminhar, um pouco mais, em pesquisas linguísticas envolvendo os pronomes relativos. Sendo assim, o próximo destino é a apresentação da pesquisa *Avoiding processing cost: differential strategies in the production of restrictive relative clauses* de Corrêa et al (2008).

As autoras têm três objetivos específicos: verificar em que medida as estratégias de formação de orações relativas (OR), refletem as demandas do processamento interno da OR na produção de sentenças; argumentar que as distintas análises sintáticas das orações relativas refletem condições distintas de planejamento; e fornecer evidências para usar métrica de custos da produção da OR, no qual o genitivo e o *pied piping* (quando uma palavra *wh* traz

uma outra palavra junto com ela) são fatores que acarretam demandas pesadas de processamento.

Antes de confirmar tais hipóteses, Corrêa et al (2008) apresentam algumas definições importantes para sua pesquisa. Primeiramente, elas fazem uma apresentação das estratégias para formação das orações relativas que serão alvo da pesquisa. No total são três: padrão, cortadora e resumptiva. A padrão tem uma lacuna na posição do pronome relativo e o *pied piping* é requerido como, no exemplo, “*O aluno de (pied piping) quem a professora falou*”. Já a oração cortadora, também conhecida como cortadora, possui uma lacuna (um pronome nulo), mas não há preposição, nem *pied piping* ocupando o lugar, por exemplo, “*O aluno que a professora falou*”. Por fim, a oração resumptiva, conhecida, também, como copiadora, tem um pronome que preenche um espaço que deveria se encontrar vazio, por exemplo, “*O aluno que a professora falou dele*”.

Basicamente, as autoras analisam as orações relativas, seguindo dois critérios primários. O primeiro critério é a análise da oração relativa funcionando como um complemento. Já o segundo é a análise da oração relativa funcionando como adjunto. De forma breve, Corrêa et al (2008) analisaram, em duas condições distintas (condição inteiramente planejada e parcialmente planejada), as produções induzidas de 40 participantes entre 18 e 48 anos, falantes de língua portuguesa brasileira. Analisados os dados, uma das conclusões a que chegam as autoras é a de que as estratégias de formação da OR em PB são meios de evitar demandas pesadas de processamento na produção on-line das sentenças. Além disso, o caso do genitivo, o *pied piping* e a impossibilidade de uma preposição órfã adicionam custo computacional, causando estratégias alternativas sob pressão.

Em relação a uma das conclusões das autoras, uma importante observação pode ser feita. As estratégias alternativas, a que elas fazem referência, suspeitamos que podem ser as disfluências. Por conta disso, esta dissertação se concentrará nos dados de Corrêa et al (2008) para verificar a ocorrência das disfluências nas sentenças de orações relativas produzidas pelos participantes do teste. Não só isso, de acordo com as condições pré-estabelecidas pelas autoras, verificaremos se na condição parcialmente planejada haverá mais casos de disfluências do que na condição inteiramente planejada.

Mediante a importância da pesquisa de Corrêa et al (2008) para nossos objetivos futuros, não poderemos apenas citar partes dela. É necessário adentrar, com mínimos detalhes, nessa pesquisa, uma vez que ela é peça fundamental para verificação da tese que foi levantada

por nós. Por isso, abordaremos no próximo capítulo todos os detalhes da pesquisa de Corrêa et al (2008).

3 DETALHES DA PESQUISA DE CORRÊA ET AL (2008)

Pesquisa primordial para nossos objetivos nesta dissertação, o trabalho realizado por Letícia Corrêa Sicuro, Marina Augusto, Mercedes Marcilese e Fernanda Miranda intitulado *Avoiding processing cost: differential strategies in the production of restrictive relative clauses*, foi um pôster apresentado na universidade de Cambridge em 2008.

Como já exposto em outros momentos, a pesquisa realizada por tais autoras tem importância essencial para o desenvolvimento de nossa dissertação. Portanto, neste capítulo, faremos uma apresentação detalhada do trabalho de 2008.

3.1 Estratégias diferenciais na produção de orações relativas restritivas (Corrêa et al, 2008)

Esta dissertação, em todo momento, caminhou para uma discussão sobre as disfluências do PB em um momento de grande complexidade como é a camada CP. O trabalho realizado por Corrêa et al (2008), como já explicado em momentos anteriores, não concentra suas atenções às disfluências, porém faz uma grande contribuição em relação à dificuldade de processamento na produção de orações relativas (orações que são geradas na camada CP). Para que o tema levantado pelas autoras seja claro, um passo a passo da pesquisa é requerido.

3.2 Objetivos em Corrêa et al (2008)

Com a base de dados formada, o trabalho de Corrêa et al (2008) trata das dificuldades computacionais de produção das orações relativas no português brasileiro. Em princípio, as autoras traçam três objetivos específicos:

- Verificar em que medida as estratégias de formação de orações relativas (OR), refletem as demandas do processamento interno da OR na produção de sentenças;
- Argumentar que as distintas análises sintáticas das orações relativas refletem condições distintas de planejamento;

- Fornecer evidências para usar métrica de custos da produção da OR, no qual o genitivo e o *pied piping* (quando uma palavra *wh* traz uma outra palavra junto com ela) são fatores que acarretam demandas pesadas de processamento.

3.3 Três estratégias diferentes de formação de orações relativas

Uma das preocupações das autoras foi a definição das três diferentes maneiras de formação das orações relativas. De acordo com elas, há as orações relativas padrão, cortadora e resumptiva (também conhecida como copiadora).

Na oração relativa padrão, ocorre uma lacuna na posição do elemento relativo e o *pied piping* é requerido, como, por exemplo, “*o aluno de quem a professora falou*”. Neste exemplo, a preposição “de” é requerida pelo verbo “falar” e a localização para sua presença (antes do pronome relativo) é satisfatoriamente preenchida.

Já a oração relativa cortadora acontece quando há uma lacuna (um pronome nulo), mas não há uma preposição nem um *pied piping* como, por exemplo, “*o aluno que a professora falou*”. Neste exemplo, a preposição, que é requerida pelo verbo “falar”, não está presente na sentença.

Por fim, na oração relativa resumptiva, um pronome preenche um espaço que deveria se encontrar vazio. Por exemplo, “*o aluno que a professora falou dele*”. Aqui, a preposição “de” contraída com o pronome “ele”, ocupam um espaço que não precisava ser preenchido.

Como já exposto em 2.6, as autoras relatam que as análises sintáticas das orações relativas se concentram nos seguintes pontos de vistas: análise das orações relativas como complementos; e análise das relativas como adjuntos.

3.4 Orações relativas em condições diferentes de planejamento

As autoras estipulam dois tipos de condições de planejamento: condição inteiramente planejada (condição 1) e condição parcialmente planejada (condição 2). Na condição inteiramente planejada, a oração relativa é gerada no complemento do determinador e o nome núcleo se move para Spec de CP de uma posição argumento dentro deste CP. Desse modo, a

carga de processamento é reduzida e, conseqüentemente, haverá maior produção de orações relativas padrão, como exposto na figura seguinte:

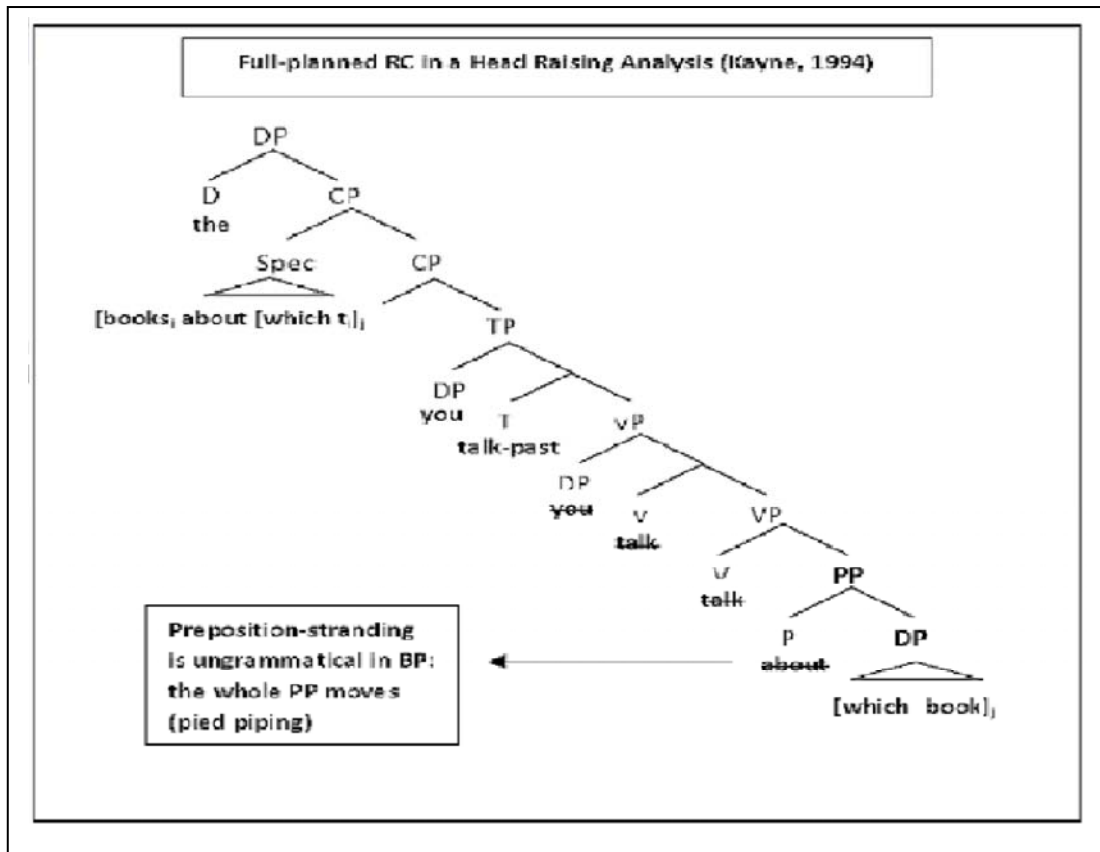


FIGURA 6: Árvore sintática: oração relativa construída na condição inteiramente planejada (Corrêa et al, 2008).

Já na condição parcialmente planejada, a oração relativa é concebida como em adjunção ao NP. O nome núcleo está fora do CP, é correlacionado com o elemento relativo e movido para um argumento com este CP para Spec C, limitado por um operador. Desse modo, a carga de processamento aumenta, o que pode causar mais produção de orações cortadoras e resumptivas.

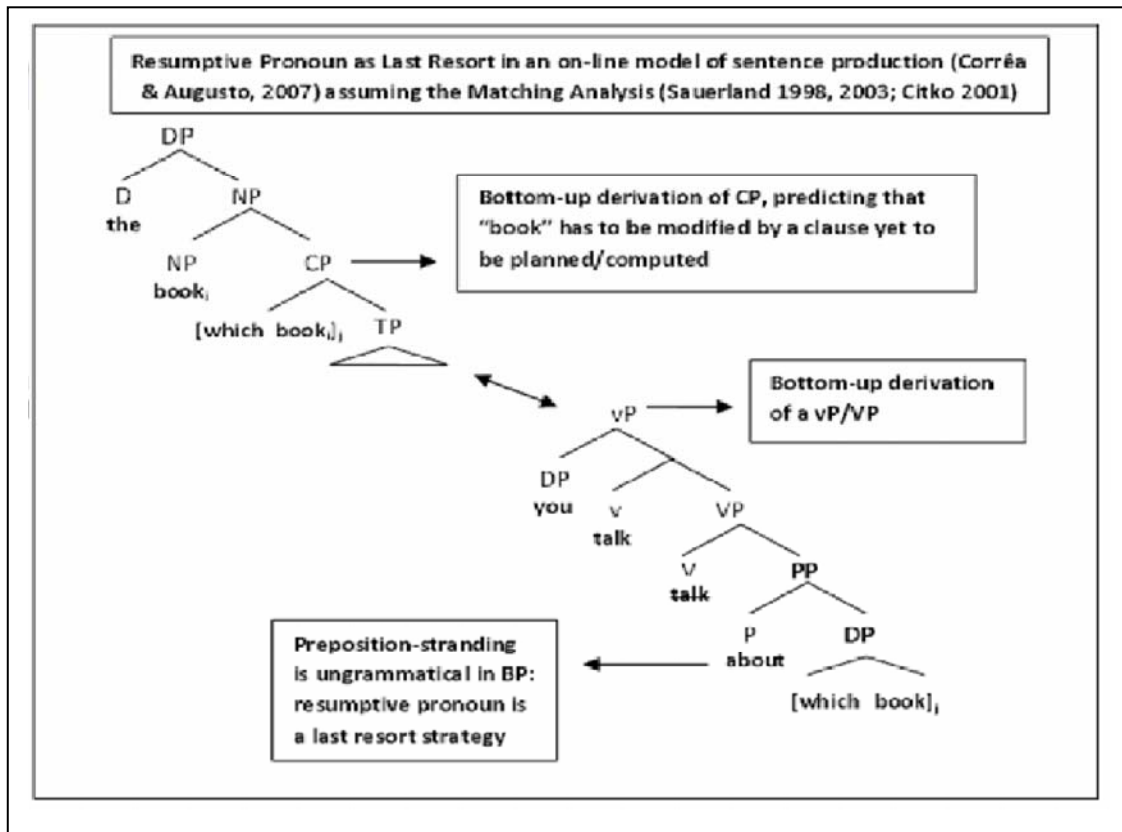


FIGURA 7: Árvore sintática: oração relativa construída na condição parcialmente planejada (Corrêa et al, 2008).

Assume-se que na condição inteiramente planejada haverá uma preferência pela produção de orações relativas padrão. Pelo contrário, na condição parcialmente planejada, espera-se maior produção de orações relativas cortadora e resumptiva (desde que não haja tempo para uma computação de uma reformulação sintática).

Sobre os dois últimos tipos de orações, Corrêa et al (2008) afirmam que as orações relativas resumptivas são uma espécie de último recurso estratégico a fim de evitar ruídos na computação sintática ou a relativização de um TopP (sintagma de tópico, como, por exemplo, “Esse menino eu não conheço”) em línguas tópico orientadas. Tal estratégia, no modelo de produção de oração relativa, seria uma maneira de contornar demandas de processamento. Em relação às orações relativas cortadoras, elas são passíveis de acontecer apenas em línguas de objeto nulo, assim como é a língua portuguesa.

3.5 Experimentos

Para alcançar tais objetivos, Corrêa et al (2008) fizeram um experimento bastante engenhoso, demonstrado de forma simplificada em seu pôster. A princípio será apresentada a forma como as autoras demonstram a metodologia em seu pôster, o que não significa dizer que será a única vez que tal metodologia será apresentada nesta dissertação. Em capítulo propício, a metodologia empregada por Corrêa et al (2008) será alvo de maiores detalhes.

O experimento contou com 40 participantes adultos falantes da língua portuguesa. Como função, eles tinham de ler alto um preâmbulo (correspondente a uma oração principal) na presença de duas figuras idênticas ou quase idênticas de uma menina ou de um menino. Depois disso, o participante deveria continuar a sentença de tal maneira que o referente do seu objeto (correspondente à figura) pudesse ser distinguido. A informação que permite distinguir as figuras é provida de forma adiantada ou on-line (5 segundos depois do preâmbulo ser apresentado), aproximadamente quando o objeto da sentença começa a ser produzido. Ao apresentar a informação de forma adiantada, assume-se que a sentença pode ser planejada inteiramente e sua produção terá, relativamente, baixa demanda de processamento (Ver figura 5). Pelo contrário, ao apresentar as informações de forma on-line, assume-se que a sentença será parcialmente planejada, ocasionando, dessa forma, altas demandas de processamento.

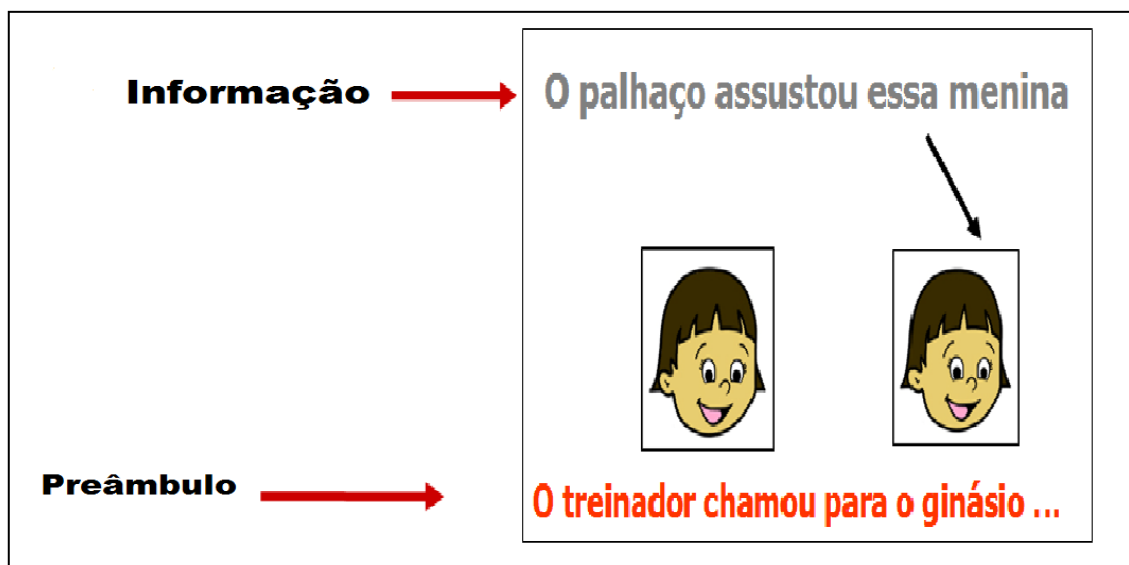


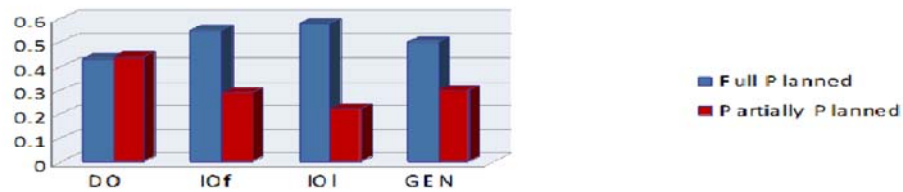
FIGURA 5: Exemplo de teste estímulo (Corrêa et al, 2008)

Para finalizar esta seção, as autoras apresentam, também, suas variáveis independentes. O foco se concentra nas orações relativas de objeto direto (OD), objeto indireto com preposição lexical (Oil), objeto indireto com preposição funcional (Oif) e orações relativas de genitivo (Gen), exemplificados pelas seguintes sentenças: “*O treinador levou para o ginásio o menino que o palhaço assustou*” (OD), “*O síndico abordou no corredor o menino para quem o escritor telefonou*” (Oil), “*A juíza convocou para a audiência a menina de quem o estudante precisou*” (Oif) e “*a arrumadeira fechou no banheiro o menino cujo tio serviu o exército*” (Gen) (Corrêa et al 2008). As demandas de processamentos são representadas por alta demanda e baixa demanda de processamento.

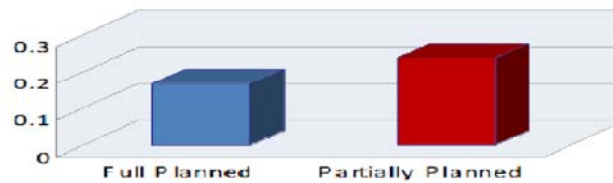
3.6 Resultados e discussões finais

Verificadas e analisadas as sentenças produzidas pelos participantes, Corrêa et al (2008) chegaram aos seguintes resultados:

Mean S responses as a function of planning condition and focus



Mean G responses as a function of planning condition



Mean R responses as a function of planning condition and focus

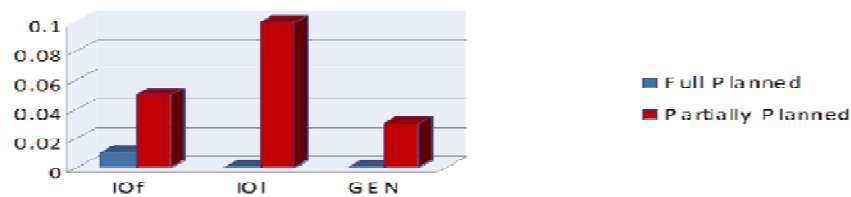


GRÁFICO 3: Resultados em Corrêa et al (2008).

De acordo com os gráficos, *Mean S* (média de respostas que utilizaram a relativa padrão), foi utilizado com mais frequência na condição inteiramente planejada (barra em azul). Em contrapartida, *Mean G* e *Mean R* (média de respostas que utilizaram as relativas cortadora e resumptiva, respectivamente) ocorrem mais vezes na condição parcialmente planejada (barra em vermelho).

Diante desses resultados algumas considerações são tecidas. Primeiramente, os dados apontam uma controvérsia na literatura a respeito da análise sintática da oração relativa poder ser resolvida se cada análise for associada a uma condição de planejamento.

Não somente isso, além de afirmarem que há adição de custo computacional nos casos do genitivo, o *pied piping* e a impossibilidade de uma preposição órfã (elementos que causam estratégias alternativas sob pressão, entre tais estratégias, suspeitamos as disfluências), as

autoras também relatam que as estratégias de formação da oração relativa em PB são meios de evitar demandas pesadas de processamento na produção on-line das sentenças.

Enfim, esses trajetos percorridos em duas pesquisas nos ambientam sobre a produção de orações relativas no PB. Como é sabido, esta dissertação busca investigar a produção de elementos disfluentes em orações relativas no PB. É momento, então, de verificar na base de dados de Corrêa et al (2008), o comportamento dos elementos disfluentes. No entanto, uma metodologia cuidadosa é requerida, por isso, o próximo capítulo reserva todos os procedimentos tomados por esta pesquisa.

4 METODOLOGIA

Muitas teses são erguidas com o apoio de textos de grandes pensadores. Esta dissertação não foge a essa regra. Sua hipótese principal se encontra apoiada na pesquisa de Corrêa et al (2008) de forma bem mais forte do que a usual. Isso porque esta dissertação busca verificar se, em uma das condições da pesquisa de Corrêa et al (2008), há maior ocorrência de disfluências do que em outra verificando, assim, que tal condição tem maior custo computacional do que a outra.

No entanto, para que tudo isso fique mais claro, este capítulo iniciará com um detalhamento maior da metodologia de Corrêa et al (2008). Logo em seguida, com base na pesquisa, apresentaremos a metodologia empregada por nós nesta dissertação. Para clareza deste capítulo, seguiremos com as seguintes seções: 4.1 *corpus* em Corrêa et al (2008) e suas subseções; 4.2 transcrições de Corrêa et al (2008) e as subseções correspondentes; 4.3 Transcrições propostas por este trabalho; e 4.4 disfluências do *corpus*.

4.1 *Corpus*

Como já exposto anteriormente, o *corpus* desta dissertação é o banco de dados da pesquisa de Corrêa et al (2008). Devido a isso, verificar detalhadamente todo o banco de dados é de suma importância.

O *corpus* foi formado por meio de gravações. Utilizando imagens em um programa de computador, o grupo de pesquisadoras coletou falas de 40 participantes, de idade entre 18 e 48 anos, todos eles falantes de língua portuguesa brasileira. Desses 40 participantes, metade participou de testes da condição inteiramente planejada (condição 1) e a outra parte de testes da condição parcialmente planejada (condição 2).

Eis aqui um detalhe importante: a distinção das condições de produção estabelecida pelas autoras. Como visto na seção 3.1.2, de acordo com as autoras, a condição 1 consiste em uma oração relativa que é gerada no complemento do determinador. O nome núcleo se move para Spec de CP, advindo de uma posição argumento dentro do CP. Nesse tipo de condição, assume-se que a carga de processamento é reduzida e, por conta disso, uma produção maior de orações relativas padrão. Já na condição 2, a oração relativa é gerada com adjunção ao NP.

Limitado por um operador, o nome núcleo se encontra fora do CP, correlaciona-se com o elemento relativo e se move para um argumento com este CP para Spec C. Assume-se que, nessa condição de planejamento, há carga maior de processamento, causando, dessa forma, produção maior de orações abertas e resumptivas.

4.1.1 Diretrizes da coleta dos dados

Para que os participantes produzissem orações relativas de acordo com as duas condições pré-estabelecidas, Corrêa et al (2008) seguiram algumas diretrizes importantes. Após a apresentação de uma figura do teste estímulo (ver figura 6), os participantes, de cada condição, deveriam obedecer a certos critérios para que as diferenças de produção pudessem ficar bem mais marcadas.



FIGURA 6: Exemplo de teste estímulo (Corrêa et al, 2008)

Para as sentenças de condição 1, as autoras apresentavam aos participantes, por meio de um programa de computador, uma figura com dois personagens. As sentenças testes eram sempre constituídas com personagens idênticos. Sentenças que serviram para distratores tiveram figuras que apresentavam personagens idênticos ou com características distintas como sexo, idade ou qualquer propriedade que levaria o participante a não produção de uma oração relativa.

Após esse primeiro momento, um dos personagens da figura é apontado por uma seta. A essa figura é dada uma informação a qual será a base de formação da oração relativa, no caso das sentenças de teste. Em seguida, o participante terá 5 segundos para análise das figuras. Com o término desse tempo, uma frase surgirá na tela com um sinal sonoro, tal frase será a oração principal. Na figura 6 temos esse momento de forma muito clara: a sentença em azul “o pai dessa menina passou mal” (fonte para construção da oração relativa) surge na tela e é indicada por uma seta para a figura alvo, que no caso das sentenças testes são idênticas; 5 segundos se passam e surge a sentença em vermelho “o vendedor atraiu para a loja” (a oração principal), juntamente com o sinal sonoro. Com todos esses elementos expostos, o participante deve dizer em voz alta o início da frase e continuá-la, fazendo referência à figura apontada pela seta. Como essa frase busca a produção de uma oração relativa genitiva (mais informações em 4.1.2), o participante deveria dizer a seguinte sentença “o vendedor atraiu para a loja a menina cujo pai passou mal”.

Quase semelhante à condição 1, os procedimentos de coleta de dados para a condição 2 variam em determinados momentos. Por exemplo, observemos a fig. 7:



FIGURA 7: Exemplo de teste estímulo na condição 2 (Corrêa et al, 2008)

Do mesmo modo da condição 1, o participante é exposto às duas figuras, sempre idênticas no caso das sentenças estímulos e nem sempre nos casos das sentenças distratoras. Uma das figuras é apontada por uma seta. Eis a primeira diferença, após esse momento, o participante tem 2 segundos para analisar as figuras, só então depois desse tempo é que surge uma frase na tela, juntamente com um sinal sonoro. Essa frase é a oração principal, em nosso exemplo a frase vermelha “o cantor incluiu na banda”. Após o sinal sonoro, é o momento de o participante ler em voz alta o início da frase e continuá-la fazendo referência à figura apontada pela seta. Por volta do momento em que o participante estiver dizendo a referência alvo, por exemplo, “o menino” ou “a menina”, será liberada mais informações sobre a referência alvo. Essa nova informação é a fonte da oração relativa, em nosso exemplo, a sentença em preto “o apresentador gosta deste menino”. Como esse exemplo de teste estímulo visa à produção de uma oração relativa de objeto indireto funcional (mais informações em 4.1.2), o participante deveria produzir a seguinte sentença “o cantor inclui na banda o menino de quem o apresentador gosta”. Todos os participantes foram submetidos a 6 frases estímulos por condição – 3 com alvo menino e 3 com alvo menina – e 44 distratores. A seguir, um quadro resumitivo dos procedimentos empregados por cada condição.

TABELA 8: Procedimentos dos testes para a condição 1 e condição 2.

Condição 1	Condição 2
O participante é exposto a duas figuras, que podem ser iguais ou não	O participante é exposto a duas figuras, que podem ser iguais ou não
Uma das figuras é apontada por uma seta	Uma das figuras é apontada por uma seta
Uma informação é dada sobre a figura apontada (essa informação será a base da oração relativa)	O participante tem 2 segundos para analisar as figuras
O participante tem 5 segundos para analisar as figuras	O início de uma frase surge na tela com um sinal sonoro (essa frase é a base da oração principal)
O início de uma frase surge na tela com um sinal sonoro (essa frase é a base da oração principal)	Após o sinal sonoro, o participante deve ler em voz alta o início da frase e continuá-la, fazendo referência à figura apontada pela seta
Após o sinal sonoro, o participante deve ler em voz alta o início da frase e continuá-la, fazendo referência à figura apontada pela seta	Após dizer a referência, por exemplo, “menino” ou “menina”, mais informações serão liberadas sobre a figura apontada pela seta (essa informação será a base da oração relativa)

4.1.2 Especificações dos dados recolhidos

Diante de todos os procedimentos, os participantes do teste produziram sentenças de orações relativas de várias formas. Além das classificações, já exemplificadas em 2.6, Padrão (Ex.: *O aluno de quem a professora falou*), Cortadora (*O aluno que a professora falou*) e Resumptiva (*O aluno que a professora falou dele*) (Corrêa et al, 2008), as autoras também definiram as orações relativas de acordo com determinações sintáticas distintas (Como visto em 3.1.2): Oações relativas com objeto direto (OD), Oações relativas genitivas (Gen), Oações relativas com objeto indireto funcional (Oif) e Oações relativas com objeto indireto lexical (Oil). Para obter orações relativas com objeto direto, Corrêa et al (2008), apresentavam figuras como as de 8, por exemplo.



FIGURA 8: Teste estímulo para produção de oração relativa com objeto direto (Corrêa et al, 2008).

Nesse caso, para haver uma oração relativa de OD, espera-se que o participante produza a seguinte sentença “o treinador levou para o ginásio o menino que o palhaço assustou”. Em orações relativas genitivas, um exemplo de teste estímulo se encontra na figura 9.



FIGURA 9: Teste estímulo para produção de oração relativa genitiva (Corrêa et al, 2008).

Aqui se espera que o participante produza a seguinte sentença “a arrumadeira fechou no banheiro o menino cujo tio serviu o exército”. Para orações relativas de objeto indireto funcional, figuras como as de 10 foram apresentadas.



FIGURA 10: Teste estímulo para produção de oração relativa de objeto indireto funcional (Corrêa et al, 2008).

Haverá uma Oif quando o participante produz a sentença “a juíza convocou para a audiência a menina de quem o estudante precisou”. Finalmente, para a produção de orações relativas de objeto indireto lexical, figuras como as de 11 foram apresentadas aos participantes.



FIGURA 11: Teste estímulo para produção de oração relativa de objeto indireto lexical (Corrêa et al, 2008).

Nessa situação, uma Oil acontecerá caso o participante produza a seguinte sentença “o síndico abordou no corredor o menino para quem o escritor telefonou”.

Logicamente, no calor do teste, o participante poderá criar sentenças em que apareçam os traços distintivos das figuras como, por exemplo, “o síndico abordou no corredor o menino moreno para quem o escritor telefonou”. Essas modificações não são suficientes para descaracterizar uma oração relativa, por isso, pequenas alterações do que se espera do ideal não são considerados erros na produção de orações relativas.

4.2 Transcrições em Corrêa et al (2008)

As transcrições dos arquivos de áudio são importantes para uma análise apurada dos fenômenos que se pretende verificar nesta dissertação. Devido a isso, esta seção tem a função de esclarecer o processo de transcrição, primeiramente em Corrêa et al (2008) e, depois, o que foi feito nesta pesquisa.

Todas as sentenças produzidas pelos participantes foram gravadas e separadas em arquivos de áudio. Em outras palavras, cada arquivo de áudio é uma sentença de cada participante. No entanto, nas gravações das sentenças de Oil, houve uma falha: o programa gravou uma sentença por cima da outra, ou seja, em vez de 6 preâmbulos gravados, 5 foram salvos. Todos os áudios foram transcritos como exemplificado na figura 12:

Transcrições	
A. P.	
Gen	<ol style="list-style-type: none"> 1. O atacante agarrou pelas pernas o menino.... ham.... O menino que o instrutor... Não sei. 2. O salva-vidas arrastou para a praia.... a menina.....meu deus, a menina... Não sei. 3. A arrumadeira fechou no banheiro o menino..... (perdi a informação....) 4. O vendedor atraiu pra loja a menina..... Não sei. 5. A universitária ajudou na tarefa.....é.... a menina do qual o professor saiu de férias. 6. A taxista largou na escola a menina..... cujo o irmão pegou gripe.
OD	<ol style="list-style-type: none"> 1. O apresentador chamou para o palco..... o menino que foi treinado.....ái... 2. A professora repreendeu no passeio a menina que o enf.... o menino que o enfermeiro visitou. 3. O treinador levou para o ginásio o menino ...o qual o palhaço assustou. 4. A bailarina convidou para a equipe a meninaque o cantor elogiou. 5. O nadador empurrou para a piscina, a menina.....é...a menina que o ciclista empurrou. 6. A trapezista puxou para a rede a menina..... com quem o domador.....se encantou, sei lá. Não consegui ler.
Oif	<ol style="list-style-type: none"> 1. O cantor incluiu na banda o ...que o apresentador gosta. 2. A enfermeira carregou para a maca o menino..... que o fotógrafo riu. Não, de quem o fotógrafo riu. 3. O palhaço entreteve na festa o menino.... cujo o repórter discordou. 4. A juíza convocou para a audiência.... a meninacom quem o estudante... precisou desta.... 5. O mágico escondeu dentro da caixa..... a menina ...que foi levada pelo turista. 6. A arquiteta trouxe para a reunião a menina..... cuja o faxineiro reclamou.
Oil	<ol style="list-style-type: none"> 1. O mestre promoveu para a faxina azul (?) o menino..... com quem o gerente brigou. 2. A redatora entrevistou para o jornalum menino com quem o advogado concordou. 3. O síndico abordou um/no corredor um menino.....éum menino..... que o escritor telefonou. 4. A psicóloga ouviu com atenção o pipo.... Não sei. 5. ? 6. O motorista esqueceu na escolaa menina... com quem o monitor viajou.

FIGURA 12: Exemplo de transcrição (Dados gentilmente cedidos pelas autoras).

Logo abaixo do título transcrições, há uma sigla A.P., feita por nós para, por questão ética, não revelar o nome do participante da pesquisa. Todas as sentenças são separadas pelas siglas Gen, OD, Oif e Oil, sendo 6 de cada tipo e 5 do caso Oil, como exposto anteriormente. As reticências significam uma espécie de pausa do participante ao falar a sentença. As informações que aparecem em parênteses são grifos das próprias autoras.

4.3 Transcrições: nossas propostas

De posse de todos os dados de Corrêa et al (2008), o primeiro passo do trabalho foi a familiarização com a pesquisa e audição das falas de todos os 40 participantes. Como o nosso foco é a ocorrência das disfluências nas orações relativas e sua relação com as dificuldades de processamento, nem todos os participantes produziram uma quantidade satisfatória de sentenças que privilegiasse a primeira condição de nossa pesquisa. Por exemplo, nos casos dos participantes J.C. e M.F., há algumas propriedades em suas falas, que não serviriam ao propósito deste texto, como exposto na tabela a seguir:

TABELA 9: Sentenças de participantes excluídos (Corrêa et al, 2008).

Sentença esperada ou próxima da esperada após a apresentação do teste estímulo	Sentença produzida pelo participante J.C	Sentença produzida pelo participante M.F
<i>“O atacante agarrou pelas pernas o menino cujo instrutor pediu demissão”</i>	<i>“O atacante agarrou pelas pernas e o instrutor do menino pediu demissão”</i>	<i>“O atacante agarrou pelas pernas o zagueiro intempestivo”</i>
<i>“O vendedor atraiu para a loja a menina cujo pai passou mal”</i>	<i>“O vendedor atraiu para a loja a menina da esquerda e o pai preocupado passou mal”</i>	<i>“O vendedor atraiu para a loja a menina assustada”</i>
<i>“O cantor incluiu na banda o menino do qual o apresentador gosta”</i>	<i>“O cantor incluiu na banda muitas crianças, o apresentador gosta desse menino da esquerda”</i>	<i>“O cantor incluiu na banda o exímio guitarrista”</i>
<i>“A juíza convocou para a audiência a menina da qual o estudante precisou”</i>	<i>“O juiz a convocou para a audiência muitas crianças, o estudante precisou desta menina da direita”</i>	<i>“A juíza convodou para a audiência a menina inteligente”</i>

Como são perceptíveis, muitas sentenças desses participantes não possuem, sintaticamente, formação de oração relativa. Sequer há a produção de uma oração relativa não

esperada, por exemplo, espera-se a produção de uma oração relativa genitiva e o participante produz uma sentença de oração relativa de objeto direto. Há casos em que os participantes até produzem orações relativas, mas em quantidade ínfima. Embora isso aconteça, para que não houvesse um desequilíbrio na análise das sentenças nas duas condições propostas por Corrêa et al (2008), consideramos as falas de 13 participantes por condição, totalizando 26 participantes, o que não quer dizer que todos esses participantes tenham produzido sempre orações relativas. Pelo contrário, há produção de orações não esperadas, mas a quantidade de orações relativas, produzidas por esses participantes, é satisfatória. Apesar da grande produção de orações relativas de distintas funções sintáticas (Gen, OD, Oif e Oil), os participantes produziram sentenças não esperadas, como já explicadas anteriormente. Tais sentenças são objetos de nossa análise e as classificamos como sentenças “Erros”.

De acordo com a conclusão de Corrêa et al (2008), de que os casos de genitivos e de *pied piping*, entre outros, causam maior custo computacional e forçam estratégias alternativas nos momentos de produção, essa pesquisa, então, resolveu privilegiar as sentenças que teria custo computacional alto como as relativas genitivas, relativas de objeto indireto lexical e objeto indireto funcional. Em outras palavras, essas orações foram escolhidas por serem mais complexas e, assim, possivelmente, acarretarem mais casos de disfluências. Por conta dessa adaptação, em relação aos participantes e às orações relativas alvos, e dos problemas na gravação das autoras da condição Oil, foram analisadas 17 sentenças de cada um dos 26 participantes, totalizando 442 sentenças analisadas. Diante dessas primeiras medidas, nossas primeiras transcrições seguiram o modelo que pode ser melhor visualizado na figura a seguir:

Transcrições	
A. P.	
Gen	
1.	O atacante agarrou pelas pernas o menino.... ããã O menino que o instrutor... Não sei. E
2.	O salva-vidas arrastou para a praia.... a menina.....meu deus, a menina... Não sei. E
3.	A arruma, a arrumadeira fechou no banheiro o menino..... (perdi a informação....) E
4.	O vendedor atraiu pra loja a menin ããã Não sei. E
5.	A universitária ajudou na tarefa..... ééé a menin ããã do qual o professor saiu de férias. E
6.	A taxista largou na escola a menina..... cujo o irmão pegou gripe. P
P 1 0/ C 0 0/ R 0 0/ E 1 4	
Oif	
1.	O cantor incluiu na banda o ...que o apresentador gosta. E
2.	A enfermeira carregou para a maca o menino..... que o fotógrafo riu. Não, de quem o fotógrafo riu. P
3.	O palhaço entreteve na festa o menino.... cujo o repórter discordou. E
4.	A juíza convocou para a audiência.... a meninacom quem o estudante... precisou desta.... R
5.	O mágico escondeu dentro da caixa..... a menina ...que foi levada pelo turista. E
6.	A arquiteta trouxe para a reunião a menina..... cuja o faxineiro reclamou. E
P 0 1/ C 0 0/ R 1 0/ E 4 0	
OII	
1.	O mestre promoveu para a faxina azul (?) o menino..... com quem o gerente brigou. P
2.	A redatora entrevistou para o jornalum menino com quem o advogado concordou. P
3.	O síndico abordou um/no corredor um menino..... ééé um menino que o escritor telefonou. C
4.	A psicóloga ouviu com atenção o pipo Não sei. E
5.	?
6.	O motorista esqueceu na escolaa menina ééé ... com quem o monitor viajou. P
P 2 1/ C 0 1/ R 0 0/ E 0 1	

FIGURA 13: Transcrição adaptada para nossa pesquisa.

Em comparação à base de dados original, acrescentamos detalhes importantes para a nossa pesquisa. Por partes: o que está em vermelho em meio às frases como, por exemplo, *ããã*, *ééé*, *a arruma arrumadeira*, é o que consideramos disfluências (mais informações na seção 4.2.3).

Ao final de cada frase, designamos quatro siglas P, C, R e E (Padrão, Cortadora, Resumptiva e Erro, respectivamente). Em sentenças padrão, há o exemplo encontrado na sentença 2 de Oif da fig. 13 “*A enfermeira carregou para a maca o menino...que o fotógrafo riu. Não de quem o fotógrafo riu.*”. Em um primeiro momento, o participante produz uma sentença não esperada, no entanto, ao perceber a inadequação, rapidamente faz uma substituição do material inadequado pelo esperado. O novo material “*não de quem o*

fotógrafo riu”, juntamente com o primeiro trecho formam uma sentença padrão, desse modo, consideramos uma produção padrão, mesmo com a presença da disfluência. Nesse caso, a disfluência ajudou o participante a substituir uma sentença inadequada por uma sentença padrão.

Para sentenças cortadoras, há o exemplo 3 de Oil da figura 13 “*O síndico abordou um, no corredor um menino... ééé...um menino... que o escritor telefonou*”. A frase apresenta muitos elementos disfluentes como “um, no” e “ééé”, mas não deixa de ser cortadora, uma vez que não surge uma preposição requerida pelo verbo *telefonar*, nesse caso, a sentença relativa deveria ser “*para quem o escritor telefonou*”, ou algo próximo disso, desde que se mantivesse a estrutura de objeto indireto lexical.

O exemplo 4 de Oif da figura 13 “*a juíza convocou para a audiência... a menina...com quem o estudante...precisou desta*”. A frase não apresenta disfluências, mas tem a presença de um elemento “*desta*” que retoma a referência alvo “*menina*”. Apesar da pouca coerência na sentença produzida, existe uma retomada do referente, o que caracteriza uma frase resumptiva.

Sentenças como a do exemplo 5 de Gen da figura 13 “*a universitária ajudou na tarefa...ééé...a menina ããã do qual o professor saiu de férias*” são consideradas erros, pois são produções inesperadas. Apesar de a sentença ser uma tentativa de produção de uma oração relativa, a frase não alcança a expectativa frasal “*a universitária ajudou na tarefa a menina cujo professor saiu de férias*”. Desse modo, sentenças como essas são consideradas erros.

Outro detalhe, abaixo de cada tipo de oração relativa (Gen, Oil e Oif), há a contagem de sentenças fluentes e de sentenças disfluentes dos tipos frasais (P, C, R e E) (ver figura 18). O primeiro número está relacionado a sentenças fluentes ocorridas naquela seção e o segundo número se relaciona a sentenças disfluentes. Observando a fig. 13, na seção Oil, há três sentenças Padrão, porém as frases 1 e 2 são sentenças fluentes, daí a contagem de número 2. Já a frase 6 apresenta disfluência, por isso, a contagem de número 1. Para finalizar a análise da contagem da seção Oil, há a presença de uma sentença cortadora com elementos disfluentes e uma sentença erro, também, com elementos disfluentes.

4.4 Disfluências no *corpus*

Antes de se aprofundar nos dados encontrados, uma seção demonstrando os elementos disfluentes adotados nesta pesquisa é fundamental. Por isso, esta seção será um passo a passo das disfluências analisadas aqui.

4.4.1 Disfluências isoladas e combinadas

Em todo o *corpus*, há elementos disfluentes que ocorrem sozinhos ou combinados. Na tabela a seguir, há a apresentação das disfluências que ocorreram sozinhas:

TABELA 10: Disfluências que ocorrem isoladas

Disfluência (sigla)	Definição	Exemplo
Alongamento (alo)	Material linguístico alongado que podem ocorrer em palavras lexicais ou funcionais	“O cantor incluiu na banda <u>áá</u> .. o menino que o apresentador gosta” (Corrêa et al, 2008)
Pausa preenchida (pp)	Vocalizações que preenchem um espaço de silêncio	“O mágico escondeu dentro da caixa a menina <u>ééé</u> de cabelos pretos da qual o turista falou” (Corrêa et al, 2008)
Repetição (r)	Repetição de material linguístico	“O síndico abordou no corredor o menino <u>que o, que o</u> escritor telefonou” (Corrêa et al, 2008)
Substituição (s)	Alteração de segmentos que tem correspondência sintática ou semântica com a zona corrigida	“O vendedor atraiu para a loja o pai <u>da, desta</u> menina o <u>qual</u> passou mal” (Corrêa et al, 2008)
Truncamento (trun)	Palavra interrompida de forma abrupta	“A juíza convocou para a audiência <u>a menina q, que</u> o estudante precisou dela”

		(Corrêa et al, 2008)
--	--	----------------------

Já a tabela a seguir reúne todas as disfluências que ocorrem de forma combinada e seus exemplos correspondentes.

TABELA 11: Disfluências que ocorrem de forma combinada

Disfluências (sigla)	Exemplo
Pausa preenchida + alongamento (ppalon)	“A <i>universitária</i> ajudou na tarefa..... <i>ééé...</i> <u>a <i>meninããã</i> do qual o professor saiu de férias</u> ” (Corrêa et al, 2008)
Pausa preenchida + substituição (pps)	“A enfermeira carregou para a maca, <u><i>o fotógrafo do qual, o o o menino do qual o fotógrafo tinha rido</i></u> ” (Corrêa et al, 2008)
Substituição + repetição + pausa preenchida (srpp)	“O síndico abordou <u><i>um, no corredor um menino.....éééum menino.....</i></u> que o escritor telefonou” (Corrêa et al, 2008)
<u>Substituição + substituição (ss)</u>	“ <u>O mestre promoveu para a faixa azul o menino...que brigou o ge... O mestre promoveu para a faixa azul o menino.....o menino com o qual o gerente brigou</u> ” (Corrêa et al, 2008)

Como um dos objetivos desta dissertação é a ocorrência das disfluências em determinados momentos da sentença, faz-se necessário a adoção de um critério de localização das disfluências, especificados na próxima seção.

4.4.2 Localização das disfluências

Todas as transcrições das sentenças disfluentes, produzidas pelos participantes da pesquisa, receberam destaques em relação à localização. Nessa pesquisa, adotamos 4 designações distintas AOR (antes da oração relativa), NAOR (na oração relativa), DPOR (depois da oração relativa) e SMOR (sem oração relativa) (ver figura 19).

Transcrições	
Gen	
1.	A taxista largou na escola <u>a menina</u> que, cujo irmão está com gripe. P/naor/s.
2.	a arrumadeira fechou na banheira <u>o menino...ããã</u> cujo pai serviu o exército. P/naor/pp
3.	a universi... a universitá... a universitária ajudou na tarefa a menina <u>cujo professor saiu de férias.</u> P/ aor/ s
4.	A arrumadeira fechou no banheiro <u>o menino</u> cujo tio serviu e, ao exército. P / dpor/ s
Oil	
1.	O motorista esqueceu na escola,..... <u>a menina, ...que</u> havia viajado com o monitor.... e não tinha, che, chegado na hora. C/ dpor/ s
Oif	
1.	A juíza convocou para audiên, para a audiência.... <u>a menina da qual</u> o estudante precisou....para atestar seu ponto de vista. P/ aor/s

FIGURA 14: Coletânea de exemplos de transcrições e localizações das disfluências.

Uma sentença será AOR, para nossos fins, quando a disfluência ocorrer antes da construção substantivo + pronome relativo + substantivo, como ocorre na sentença 3 de Gen na figura 14 “*a universi... a universitá... a universitária ajudou na tarefa a menina cujo professor saiu de férias*”. Nesse exemplo, a disfluência “*a universi... a universitá... a universitária*” ocorre antes do substantivo “*menina*” e do pronome relativo “*cujo*” e do substantivo mais próximo “*professor*”, caracterizando, então, uma sentença AOR.

NAOR sempre ocorrerá quando a disfluência se encontrar exatamente na construção substantivo + pronome relativo + substantivo, como, por exemplo, na primeira sentença de Gen, na figura 14, “*a taxista largou na escola a menina que, cujo irmão está com gripe*”.

Nessa sentença, a disfluência “que, cujo”, pronomes relativos, fazem parte da construção substantivo + pronome relativo + substantivo. Assim, consideramos uma sentença NAOR.

Seguindo a lógica, DPOR só ocorrerá depois de tal construção, como assim acontece na sentença 4 de Gen da figura 14 “*A arrumadeira fechou no banheiro o menino cujo tio serviu e, ao exército*”. A disfluência “e, ao” segue a construção alvo, dessa vez exemplificada por “menino cujo tio”. Logo, essa sentença é definida como DPOR.

Por fim, uma sentença SMOR será quando o participante produzir uma sentença sem que haja a presença de um pronome relativo, como, por exemplo, “*a redatora entrevistou para o jornal... o menino...o menino.... de cabelo verme, castanho*” (Corrêa et al, 2008). Mesmo repleto de elementos disfluentes, o participante não produz nada que se assemelhe a uma oração relativa, caracterizando, dessa forma, uma sentença SMOR.

Com todos os cuidados tomados, eis que chega o momento mais importante para nossa pesquisa: a verificação de nossa hipótese. Uma vez consciente das dificuldades de produção dos falantes e das condições estabelecidas por Corrêa et al, 2008 (condição parcialmente planejada e condição inteiramente planejada), uma dúvida nos parece pertinente: se os participantes da pesquisa de Corrêa et al, 2008, em meio ao teste, produziram estratégias alternativas, entre elas, as disfluências, em qual das duas condições haverá mais disfluências? A essa pergunta, nossa hipótese nos parece ser uma boa resposta: **a condição parcialmente planejada terá maior ocorrência de disfluências do que na condição inteiramente planejada.**³

A hipótese está lançada, agora é hora de se debruçar sobre os dados e esclarecer que tal afirmação tem consistência. Eis o que nos espera no próximo capítulo.

³ Com o auxílio da professora Marina Augusto, outra hipótese nos foi sugerida. Sabendo que as disfluências na região do CP indicam que uma estrutura de alto custo será produzida (Horne et al, 2005) e que Corrêa et al (2008) sugerem que a condição planejada propicia maior número de sentenças relativas padrão (com custo computacional mais elevado) e a não-planejada, a ocorrência de sentenças relativas cortadoras ou resumptivas (custo computacional menos elevado), como poderíamos cruzar essas informações com as disfluências? Diante disso, duas hipóteses são sugeridas: quando há a produção de sentenças relativas padrão, espera-se maior número de disfluências, não importando se a condição é planejada ou não-planejada, o custo da relativa padrão se sobrepõe à natureza do planejamento. Essa natureza é essencial para prevenir-se mais disfluências, não importando o tipo de relativa produzida, se padrão ou não-padrão.

Para tal hipótese ser confirmada, é importante lançar mão de um cruzamento dos dados com os tipos de relativas (padrão, cortadora ou resumptiva), podendo, dessa forma, trazer evidências como antecipação de uma estrutura complexa, entre outros. Essa forma de cruzamento de dados é uma sugestão importante que será tratada com mais cuidados em trabalhos futuros.

5 RESULTADOS

Para cumprir nossos objetivos nesta pesquisa, destacaremos as ocorrências das disfluências no *corpus*. Não só isso, focaremos a nossa análise, principalmente nas disfluências em localizações específicas das sentenças. Acreditamos que, com isso, poderemos confirmar nossa hipótese de que na condição parcialmente planejada haverá maior ocorrência de disfluências do que na condição inteiramente planejada.

Devido a isso, seguiremos por certos caminhos. Na seção 5.1, abordaremos as disfluências na localização AOR (antes da oração relativa). Na seção 5.2, avançaremos e trataremos das disfluências na localização DPOR (depois da oração relativa). Por fim, na seção 5.3, por conta de sua importância em nossos achados, analisaremos as disfluências na localização NAOR (na oração relativa).

5.1 Complicações em AOR

Ao separarmos a ocorrência das disfluências em localizações, pudemos observar diferentes comportamentos do formulador sintático. Quando as disfluências ocorrem antes da oração relativa (AOR), algumas análises interessantes podem ser feitas. Primeiramente, apresentamos a tabela geral de ocorrência das disfluências em AOR:

TABELA 12: Ocorrência das disfluências em AOR.

Disfluências	Total
Alongamento	1 (5,88%)
Pausa preenchida	1 (5,88%)
Pausa preenchida + substituição	1 (5,88%)
Repetição	2 (11,76%)
Substituição	10 (58,84%)
Substituição + substituição	1 (5,88%)
Truncamento	1 (5,88%)
Total	17 (100%)

De acordo com os dados, houve 17 ocorrências de disfluências em AOR como, por exemplo, “*a universi... a universitá... a universitária ajudou na tarefa a menina cujo professor saiu de férias*”. Neste exemplo, a disfluência (nesse caso, uma substituição) ocorre antes da construção substantivo + pronome relativo + substantivo (neste exemplo, seria o trecho “*a menina cujo professor*”).

Quando há a separação das disfluências do *corpus* nas condições 1 e 2 (inteiramente planejada e parcialmente planejada, respectivamente), há maior número de disfluências na condição 2 do que na condição 1:

TABELA 13: Ocorrência das disfluências na condição 1 em AOR.

Disfluências	Condição 1
Alongamento	1 (20%)
Pausa preenchida	0 (0%)
Pausa preenchida + substituição	0 (0%)
Repetição	1 (20%)
Substituição	3 (60%)
Substituição + substituição	0 (0%)
Truncamento	0 (0%)
Total	5 (100%)

TABELA 14: Ocorrência das disfluências na condição 2 em AOR.

Disfluências	Condição 2
Pausa preenchida	1 (8,33%)
Pausa preenchida + substituição	1 (8,33%)
Repetição	1 (8,33%)
Substituição	7 (58,35%)
Substituição + substituição	1 (8,33%)
Truncamento	1 (8,33%)
Total	12 (100%)

Importante destacar nesses dados é a grande ocorrência das substituições em AOR. Essa situação fica melhor visualizada no quadro comparativo a seguir:

TABELA 15: Comparação entre as disfluências na condição 1 e 2 em AOR.

Disfluências	Condição 1	Condição 2	Total
Alongamento	1 (100%)	0 (0%)	1 (100%)
Pausa preenchida	0 (0%)	1 (100%)	1 (100%)
Pausa preenchida + substituição	0 (0%)	1 (100%)	1 (100%)
Repetição	1 (50%)	1 (50%)	2 (100%)
Substituição	4 (36,37%)	7 (63,63%)	11 (100%)
Substituição + substituição	0 (0%)	1 (100%)	1 (100%)
Truncamento	0 (0%)	1 (100%)	1 (100%)
Total	5 (29,41%)	12 (70,59%)	17 (100%)

Com exceção do alongamento, todas as disfluências que ocorrem na condição 2 empatam ou superam a ocorrência de disfluências na condição 1. Um exemplo é o que ocorre nos casos da disfluência substituição. Há 4 ocorrências de substituições na condição 1 (36,37%), por exemplo, a sentença “*A redado, a redatora entrevistou para o jornal,... o menino,... com o qual... o advogado concordou*”; enquanto ocorrem 7 casos (63,63%) na condição 2, por exemplo, a sentença “*A juíza convocou para audiên, para a audiência.... a menina da qual o estudante precisou....para atestar seu ponto de vista*”.

Por conta disso, acreditamos que essa tendência no *corpus*, ou seja, uma grande ocorrência de disfluências na condição 2 (parcialmente planejada), deve-se ao fato de o falante estar antecipando problemas. Em outras palavras, ele está em uma situação em que sua fala não está totalmente planejada, o falante se depara com a produção de uma oração relativa, sentença comprovadamente mais complexa, e, se antecipando a problemas, produz disfluências para que sua fala seja construída perfeitamente para seu objetivo comunicacional.

Importante ressaltar, também, que nossa hipótese, pelo menos em AOR, foi confirmada: realmente houve mais casos de disfluências na condição 2 do que na condição 1. Embora haja essa primeira confirmação, muito ainda há para verificar nas outras localizações de disfluências. É o que será feito na próxima seção em relação à localização DPOR.

5.2 Caso DPOR: mais disfluências na condição planejada

Nossa hipótese previa que a condição 2 teria mais casos de disfluências do que a condição 1. No entanto, em DPOR ocorreu o oposto, algo que prontamente explicaremos nesta seção. Para começar essa explicação, seguiremos as mesmas diretrizes em AOR. Primeiramente, apresentamos a tabela geral de ocorrência das disfluências na localização DPOR, para após tecermos comentários sobre os dados:

TABELA 16: ocorrência das disfluências em DPOR.

Disfluências	Total
Pausa preenchida	3 (27,28%)
Repetição	1 (9,09%)
Substituição	5 (45,45%)
Substituição + substituição	1 (9,09%)
Truncamento	1 (9,09%)
Total	11 (100%)

A tabela apresenta ocorrência de 11 disfluências na localização DPOR. Importante lembrar que as disfluências foram assim classificadas toda vez que surgiram após a construção substantivo + pronome relativo + substantivo, como, por exemplo, “A *arrumadeira fechou no banheiro o menino cujo tio serviu e... ao exército*”. Neste exemplo, a disfluência “*e... ao*” (neste caso, uma substituição) ocorre após a construção alvo (pelo exemplo, “*menino cujo tio*” – substantivo + pronome relativo + substantivo).

Ao separarmos as disfluências em duas condições em DPOR, deparamo-nos, então, com um número maior de ocorrências de disfluências na condição 1 do que na condição 2, como exposto nos próximos dados:

TABELA 17: Ocorrência das disfluências na condição 1 em DPOR.

Disfluências	Condição 1
Pausa preenchida	1 (14,29%)
Repetição	1 (14,29%)
Substituição	3 (42,84%)
Substituição + substituição	1 (14,29%)
Truncamento	1 (14,29%)
Total	7 (100%)

TABELA 18: Ocorrência das disfluências na condição 2 em DPOR.

Disfluências	Condição 2
Pausa preenchida	2 (50%)
Repetição	0 (0%)
Substituição	2 (50%)
Substituição + substituição	0 (0%)
Truncamento	0 (0%)
Total	4 (100%)

De acordo com os dados apresentados, a condição 1 apresentou 7 ocorrências de disfluências (com destaque para as substituições, que isoladas ou juntas a outras disfluências, ocorrem 4 vezes no *corpus*, aproximadamente 57 % dos casos). Em contrapartida, a condição 2 apresentou 4 ocorrências de disfluências. Esses dados podem ser melhor visualizados no quadro comparativo a seguir:

TABELA 19: Comparação entre as disfluências na condição 1 e 2 em DPOR

Disfluências	Condição 1	Condição 2	Total
Pausa preenchida	1 (33,34%)	2 (66,66%)	3 (100%)
Repetição	1 (100%)	0 (0%)	1 (100%)
Substituição	3 (60%)	2 (40%)	5 (100%)
Substituição + substituição	1 (100%)	0 (0%)	1 (100%)
Truncamento	1 (100%)	0 (0%)	1 (100%)
Total	7 (63,63%)	4 (36,37%)	11 (100%)

Excetuando-se as pausas preenchidas, todas as disfluências da condição 1 empataram ou superaram às da condição 2. Mais uma vez, usamos como exemplo o caso das substituições. Na condição 1, há 3 ocorrências de substituição (60%), como, por exemplo, a sentença “*a universitária ajudou na tarefa a menina que saiu de férias sem, estava sem ajuda do professor*”; enquanto que, na condição 2, há 2 ocorrências (40%), como, por exemplo, “*o motorista esqueceu na escola,..... a menina, ...que havia viajado com o monitor.... e não tinha, che, chegado na hora*”.

Embora ocorra essa peculiaridade, há uma explicação plausível para tal. Todas as orações relativas da condição 1 tendem à produção de oração relativa padrão. Esses dados foram verificados na pesquisa de Corrêa et al (2008), ao apresentarem maior tendência de produção de orações relativas padrão na condição 1, em face à maior produção de orações relativas cortadoras e resumptivas na condição 2. Por conta disso, os falantes, no momento da produção de sentenças na condição 1, voltaram sua atenção para a produção da relativa padrão, fazendo com que houvesse mais gastos de processamento para produção de tal oração. Devido a esse gasto maior de processamento, houve, então, um momento de sobrecarga e, por consequência, mais disfluências. É como se o falante “relaxasse” após a produção complexa (oração relativa padrão) e, por conta disso, produzisse mais disfluências.

Até então, focamos o nosso olhar para antes e depois da oração relativa, basta agora verificarmos a ocorrência das disfluências na própria oração relativa (NAOR). É nessa

localização que esperamos dar uma contribuição maior ao que está sendo pesquisado neste trabalho. Por isso, a próxima seção está voltada à ocorrência das disfluências em NAOR.

5.3 Análise minuciosa de NAOR

Localização em que a complexidade é maior e que se espera maior ocorrência de disfluências, NAOR prova ser uma localização interessante a ser analisada. Tudo por conta de um número de ocorrência de disfluências e de dados que mostram que essa localização deve ser vista com muito cuidado. Por partes, seguiremos uma rotina igual a que fizemos nas outras localizações, ou seja, a seguir, apresentamos um quadro geral de ocorrência das disfluências em NAOR:

TABELA 20: Ocorrência das disfluências em NAOR.

Disfluências	Total
Alongamento	1 (2,17%)
Pausa preenchida	11 (23,92%)
Pausa preenchida + alongamento	1 (2,17%)
Pausa preenchida + substituição	6 (13,05%)
Repetição	3 (6,53%)
Substituição	19 (41,31%)
Substituição + repetição + pausa preenchida	2 (4,34%)
Substituição + substituição	2 (4,34%)
Truncamento	1 (2,17%)
Total	46 (100%)

Os dados apresentam 46 ocorrências de disfluências na localização NAOR, que nunca custa esclarecer, é a localização em que ocorre disfluências na construção substantivo + pronome relativo + substantivo. Para exemplificar sentenças com disfluências em NAOR,

apresentamos a seguinte sentença disfluente “o mestre promoveu para a faixa azul este menino ééé com o qual o gerente brigou”. Esta sentença é classificada como NAOR, pois a disfluência “ééé” (nomeada como pausa preenchida), ocorre em meio à construção alvo sugerida por nós (neste caso, “menino ééé com o qual o gerente” – substantivo + pronome relativo + substantivo).

As 46 disfluências apresentaram uma possibilidade maior de análise quando separadas pelas condições 1 e 2. Nos quadros seguintes, apresentamos as disfluências em NAOR, separadas por condições:

TABELA 21: Ocorrência das disfluências na condição 1 em NAOR.

Disfluências	Condição 1
Alongamento	1 (5,88%)
Pausa preenchida	2 (11,76%)
Pausa preenchida + alongamento	0 (0%)
Pausa preenchida + substituição	0 (0%)
Repetição	2 (11,76%)
Substituição	11 (64,72%)
Substituição + repetição + pausa preenchida	0 (0%)
Substituição + substituição	0 (0%)
Truncamento	1 (5,88%)
Total	17 (100%)

TABELA 22: Ocorrência das disfluências na condição 2 em NAOR.

Disfluências	Condição 2
Alongamento	0 (0%)
Pausa preenchida	9 (31,04%)
Pausa preenchida + alongamento	1 (3,45%)
Pausa preenchida + substituição	6 (20,68%)
Repetição	1 (3,45%)
Substituição	8 (27,58%)
Substituição + repetição + pausa preenchida	2 (6,90%)
Substituição + substituição	2 (6,90%)
Truncamento	0 (0%)
Total	29 (100%)

Na condição 1, há ocorrência de 17 disfluências, sendo as substituições responsáveis por 64,72% do *corpus* (11 ocorrências). Em relação à condição 2, há ocorrência de 29 disfluências em NAOR. Destaque, para esses dados, é a grande ocorrência da disfluência pausa preenchida, não somente de forma isolada, como também de forma combinada com outras disfluências, o que corresponde a 62,07% do *corpus*, somando formas isolada e combinada. Quando há a comparação direta entre as duas condições em disfluências na localização NAOR, algumas observações importantes podem ser feitas:

TABELA 23: Comparação entre as disfluências na condição 1 e 2 em NAOR

Disfluências	Condição 1	Condição 2	Total
Alongamento	1 (100%)	0 (0%)	1 (100%)
Pausa preenchida	2 (18,19%)	9 (81,81%)	11 (100%)
Pausa preenchida + alongamento	0 (0%)	1 (100%)	1 (100%)
Pausa preenchida + substituição	0 (0%)	6 (100%)	6 (100%)
Repetição	2 (66,66%)	1 (33,34%)	3 (100%)
Substituição	11 (55%)	8 (45%)	20 (100%)
Substituição + repetição + pausa preenchida	0 (0%)	2 (100%)	2 (100%)
Substituição + substituição	0 (0%)	2 (100%)	2 (100%)
Truncamento	1 (100%)	0 (0%)	1 (100%)
Total	17 (36,96 %)	29 (63,04%)	46 (100%)

Apesar das categorias alongamento, repetição e truncamento terem maior ocorrência na condição 1 do que na condição 2, elas juntas não são o suficiente para tornarem as ocorrências das disfluências na condição 1 (17 ocorrências, 36,96% dos casos) mais frequentes do que na condição 2 (29 ocorrências, 63,04% dos casos). Para exemplificar as duas condições, usamos duas sentenças em que ocorrem pausas preenchidas: na condição 1, temos “*o mágico escondeu dentro da caixa a menina ééé de cabelos pretos da qual o turista falou*”, enquanto que na condição 2 temos a seguinte sentença exemplo “*o vendedor atraiu para a loja a menina....ããã cujo pai passou mal*”.

Como era esperado, a condição 2 apresentou mais disfluências do que a condição 1. Entretanto, dessa vez, há uma diferença de disfluências frequentes em NAOR, se comparada à AOR e DPOR. Em AOR e DPOR, a disfluência mais frequente é a substituição, tanto na condição 1, como na condição 2. Em NAOR, a condição 1 apresenta mais casos de substituição e a condição 2 mais casos de pausas preenchidas (isoladas ou juntas a outras

disfluências), como já exposto anteriormente. É exatamente esse dado novo que nos permite outro tipo de análise.

Por conta da grande ocorrência de pausas preenchidas (combinadas ou não) na condição 2, acreditamos, então, que essa disfluência seja responsável pela diferenciação das duas condições na localização NAOR. O que acontece em NAOR é algo completamente diferente do que ocorre nas outras duas localizações, o que, provavelmente, é decorrência das características das orações relativas em si.

Eklund & Shriberg (1998 in Moniz, 2006) afirmam que a localização das pausas preenchidas em pontos específicos da sentença são indicadores de estratégias feitas pelos falantes no momento da produção da fala. Quando as pausas preenchidas ocorrem mais frequentemente no início da sentença, o falante estaria fazendo um planejamento global do enunciado. Por outro lado, quando as pausas preenchidas ocorrem em meio à sentença, o falante, então, estaria fazendo uma procura lexical. Portanto, de acordo com os achados de Eklund & Shriberg (1998), acreditamos que a maior ocorrência das pausas preenchidas em meio à sentença, em nosso caso, em NAOR, em nosso *corpus* estaria obedecendo a uma procura lexical por conta dos falantes. Entretanto, como as orações relativas são espécies de novas sentenças, acreditamos também que as pausas preenchidas têm, de certa forma, função de planejamento discursivo.

Em pesquisas como a de Shriberg (1994), Candea (2000) e Eklund (2004), as pausas preenchidas são as disfluências que têm maior frequência. Em Moniz (2006), repetições e alongamentos foram os mais frequentes. Em nosso *corpus*, houve tendência maior às substituições, em primeiro lugar, e às pausas preenchidas, em segundo lugar, como assim fica exposto na próxima tabela:

TABELA 24: Ocorrência das disfluências em todo *corpus*

Disfluências	Total
Alongamento	2 (2,71%)
Pausa preenchida	15 (20,27%)
Pausa preenchida + alongamento	1 (1,36%)
Pausa preenchida + substituição	7 (9,45%)
Repetição	6 (8,10%)
Substituição	34 (45,94%)
Substituição + repetição + pausa preenchida	2 (2,71%)
Substituição + substituição	4 (5,40%)
Truncamento	3 (4,06%)
Total	74 (100%)

Combinados ou não, as substituições aparecem em 63,5% dos casos, enquanto as pausas preenchidas, combinadas ou não, são 33,79% dos casos. Na pesquisa de Moniz (2006), alongamentos e repetições foram preferidas pelos falantes, em face às pausas preenchidas, para preparação de unidades subsequentes, tendência essa que não se confirmou em nossa pesquisa. As pausas preenchidas parecem ser as estratégias disfluentes adotadas, na maioria dos casos, em uma busca lexical e preparação do que virá pela frente.

Por fim, ao analisarmos as vocalizações que ocorrem nas pausas preenchidas, algumas observações podem ser feitas. Em Moniz (2006), as pausas que mais ocorrem são a forma *aa* (78,47%), *aam* (17,15%) e *mm* (4,38%). Tendência não verificada em nosso *corpus*, como pode ser visualizada no quadro a seguir:

TABELA 25: Formas de pausas preenchidas em PB.

Tipo de Pausa Preenchida	Total
<i>ããã</i>	14 (53,85%)
<i>ééé</i>	8 (30,76%)
<i>ooo</i>	4 (15,39%)
Total	26 (100%)

Com exceção de “ããã”, todos os outros dados são distintos. Isso significa dizer que, mesmo sendo falantes de uma mesma língua, os falantes de PB e de PE usam estratégias disfluêntes diferentes. Tal detalhe pode levantar uma discussão entre os mais entusiastas sobre a existência de uma língua brasileira e de uma língua portuguesa, uma vez que até as disfluências de pausas preenchidas, nas duas línguas, são distintas. Todavia, não queremos entrar no mérito dessa discussão, apenas trazer mais uma informação sobre esse tema.

As disfluências, em todas as suas nuances, apresentam diversas peculiaridades. Entretanto, resolvemos focar as dificuldades computacionais do formulador sintático, principalmente ao que concerne às disfluências nas orações relativas. Durante este texto, comprovamos que as disfluências seriam em maior número na condição 2 do que na condição 1. Não só isso, conseguimos verificar que tal hipótese tinha cabimento, além de ainda podermos fazer comentários bastante pertinentes sobre o formulador sintático. Caminho percorrido, hora de encerrar nossa jornada.

6 CONSIDERAÇÕES FINAIS

Fácil. Não. Trabalhoso. Sim. Essa é a melhor definição para o caminho que percorremos nesta dissertação. Tratar de fenômenos da fala relacionados com a mente não é uma tarefa simples de ser feita. Percebemos essa situação na pele, logo no primeiro momento, quando buscávamos uma teoria que desse sustentação ao que estávamos estudando. Encontramos terreno firme no Gerativismo, uma teoria linguística que, entre outros suportes teóricos, dá conta de estruturas, nossa necessidade, pois tratávamos de fala espontânea. Não somente isso, vimos que precisávamos de entendimento das mais modernas pesquisas gerativas, em outras palavras, o Programa Minimalista. Isso se deve porque entender estruturas complexas, árvores sintáticas, nós sintáticos, camadas CP, entre outros, sem esse aparato teórico, seria impossível.

Como citamos fala espontânea, tratamos diretamente de produção da fala. Como o tema da dissertação está voltado para o fenômeno das disfluências, logo mencionamos as dificuldades que o formulador sintático enfrenta na produção de uma fala, principalmente em contextos complexos. Quando citamos complexidade, incluem-se, nessa ocasião, as camadas sintáticas, principalmente, a camada CP. Dificuldades nessa área, por exemplo, influenciam estratégias diferenciadas ao falante, estratégias como as disfluências, por exemplo. Com isso, vimos que precisávamos ter conhecimento sobre a dificuldade do formulador sintático, explicações favoráveis que encontramos no terreno da Psicolinguística.

Só após o entendimento psicolinguístico, pudemos, então, adentrar às questões das disfluências. Tivemos de recorrer a autores que já haviam tratado, de certa forma, sobre esse tema. Vimos que as disfluências já foram vistas como autocorreções (ver Levelt, 1989), receberam uma terminologia específica para uma melhor análise (ver Shriberg, 1994) e que foram analisadas em falantes de língua portuguesa europeia (ver Moniz, 2006). Como nosso foco esteve voltado às disfluências em camadas complexas, também tivemos de beber em fontes que tratassem desse tema. Encontramos em Horne et al (2005), Jaeger (2005), Clark e Wasow (1998), estudos que trataram sobre a complexidade das estruturas e a localização sintática das disfluências.

Com tudo isso, bastava, agora, transferir todos esses conhecimentos para uma pesquisa em língua portuguesa brasileira. Para isso, precisávamos de uma boa base de dados,

principalmente relacionados à produção de orações relativas. Encontramos em Corrêa et al (2008) a grande base de dados de que precisávamos. Primeiro, buscamos compreender o trabalho do grupo científico, para após adaptarmos alguns detalhes para continuação de nossas pesquisas.

Observamos que a pesquisa realizada por Corrêa et al (2008) possuía duas condições distintas para argumentar, entre outros fatores, sobre as demandas do processamento interno da OR na produção da sentença. Essas duas condições consistiam em uma que era inteiramente planejada e outra que era parcialmente planejada. A diferença entre essas duas condições está no modo de recolha dos dados. Quando as informações do teste são apresentadas de forma adiantada, admite-se que a produção pode ser inteiramente planejada, resultando em baixa demanda de processamento. Pelo contrário, quando as informações sobre a sentença são apresentadas de forma on-line, admite-se que a produção pode ser parcialmente planejada, resultando em alta demanda de processamento. Diante dessas especificidades, os participantes, divididos pelas duas condições, produziram sentenças com orações relativas de todos os tipos.

Feito isso, coube a nós ouvir todos os arquivos de áudio e anotar as disfluências que ali ocorriam. Fizemos isso, pois tínhamos um objetivo: demonstrar que, entre as duas condições estabelecidas por Corrêa et al (2008), a condição parcialmente planejada apresentaria mais disfluências que a condição inteiramente planejada. Para verificar tal hipótese, analisamos as disfluências em três localizações distintas: antes da oração relativa (AOR), depois da oração relativa (DPOR) e na oração relativa (NAOR). Para definirmos essas localizações, estipulamos que as disfluências deveriam ocorrer antes, durante ou depois da estrutura substantivo + pronome relativo + substantivo, como por exemplo, “*a redado, a redatora* entrevistou para o jornal,... o menino,... com o qual... o advogado concordou”. Nesse caso, a disfluência “*a redado, a redatora*” ocorre antes da estrutura alvo, pelo exemplo apresentado, o trecho que corresponde à “o menino, com o qual o advogado”. Ocorrência de disfluências NAOR e DPOR pode ser visualizada nos seguintes exemplos, respectivamente, “o mágico escondeu dentro da caixa a menina *ééé* de cabelos pretos da qual o turista falou” e “o motorista esqueceu na escola,..... a menina, ...que havia viajado com o monitor.... e não tinha, *che, chegado* na hora”.

Feitas as análises, confirmamos nossa hipótese: realmente há maior ocorrência de disfluências na condição parcialmente planejada do que na condição inteiramente planejada,

com exceção à localização DPOR. Isso se deve porque, de acordo com os dados de Corrêa et al (2008), a condição inteiramente planejada tende à maior produção de orações relativas padrão, em face à condição parcialmente planejada, que tende à maior produção de orações relativas cortadora e resumptiva. Por conta dessa característica, os falantes, na condição inteiramente planejada, preocupavam-se com a produção da relativa padrão – fator que traz uma sobrecarga maior de processamento –, sendo que, após a produção da relativa, os falantes “relaxavam”, ocasionando, então, mais casos de disfluências.

Em AOR, há uma grande ocorrência da disfluência substituição nas duas condições estipuladas por Corrêa et al (2008). Diferentemente de DPOR, a condição parcialmente planejada apresentou mais casos de disfluências. Esse dado nos trouxe a seguinte reflexão: o falante, antecipando a produção de uma sentença complexa, ou seja, a oração relativa, produz disfluências com intuito de produzir a melhor sentença para dar conta de seu objetivo na comunicação, que é a de passar uma mensagem inteligível.

Algumas reflexões também foram feitas na ocorrência das disfluências em NAOR. Na condição inteiramente planejada, essa localização apresentou maior ocorrência da disfluência substituição – fator parecido com todas as outras localizações –, porém na condição parcialmente planejada, curiosamente, a disfluência que se destaca são as pausas preenchidas. Essa tendência é consequência da própria característica das orações relativas, sentenças altamente complexas. Além disso, as pausas preenchidas revelaram estratégias utilizadas pelos falantes: quando a pausa preenchida ocorre em meio a sentença, o falante estaria em um momento de procura lexical. Embora essa afirmação seja verdade, não se pode descuidar das características da oração relativa. Ela é como uma nova sentença sendo gerada dentro de uma outra sentença. Por isso, também entendemos que as pausas preenchidas estariam no início da sentença, nesse caso, no início da oração relativa. Quando as pausas preenchidas ocorrem em início de sentença, afirma-se, então, que o falante estaria em um momento de planejamento discursivo.

Assim chegamos ao término desse texto. Admitimos que mais análises poderiam ser feitas, como por exemplo, a ocorrência das disfluências em sentenças que não tiveram orações relativas, afinal, se não houve gastos computacionais para produzir uma sentença complexa, por que então o falante teve momentos disfluentes?

Outras abordagens também foram pensadas por nós, como, por exemplo, a gravação de dois dias de um congresso e a análise desses palestrantes. Chegamos a ter esse material,

mas como já tínhamos em mente a análise das disfluências em orações relativas, verificamos que não houve muitas produções dessas sentenças durante todo o congresso.

Também pensamos em um teste perceptivo, nos moldes do que foi feito em Moniz (2006), essa sim, uma tarefa interessante para trabalhos futuros. Outra boa possibilidade, também, é a análise de sentenças de um único tipo, como aqui, que foi feito apenas com orações relativas. Nossa sugestão é a análise das disfluências em sentenças passivas, construção que pode ser tão complexa como as orações relativas.

Enfim, encerramos esse texto, não com a certeza de que concluímos reflexões. Pelo contrário, apenas engatinhamos em um terreno tão desconhecido e fascinante, como é o terreno das disfluências.

REFERÊNCIAS

BRENNAN, S. ; M. SCHOBBER. How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language*, 44, 2001.

_____. Processes that shape conversation and their implications for computational linguistics. ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. 28., Hong Kong, 2000. *Proceeding...* Hong Kong: [s.n], 2000.

CANDEA, M. *Contribution a l'Etude des Pauses Silencieuses et des Phenomenes Dits «d'Hesitation» en Français Oral Spontane – Etude sur un corpus de récit en classe de Français* 2000. Tese (Doutorado) - Universite Paris III – Sorbonne Nouvelle, Paris,2000.

_____. *Les euh et les allongements dits « d'hesitation » : deux phenomenes soumis a certaines contraintes en francais oral non lu*. In: Actes des 23è JEP, [S.l.: s.n.], 2000.

CHOMSKY, Noam. *Aspects of the Theory of Syntax*. Cambridge ;, Massachusetts: MIT Press, 1965.

_____. *Lectures on government and binding*. Dordrecht: Foris, 1981.

CLARK, H. H. ; TREE, J. E. Fox. Using uh and um in spontaneous speech. *Cognition*,, 2002.

_____; WASOW, T. Repeting words in spontaneous speech. *Cognitive Psychology*, 37, 1998.

CORRÊA, Letícia M. et al. Mercedes. *Avoiding processing cost: differential strategies in the production of restrictive relative clauses*. Annual CONFERENCE ON ARCHITECTURES AND MECHANISMS FOR LANGUAGE PROCESSING, 14., Cambridge, 2008. [*Proceeding...*].Cambridge: [s.n.], 2008.

EKLUND, R. ; E. SHRIBERG. *Crosslinguistic disfluency modeling: a comparative analysis of Swedish and American English human-human and human-machine dialogs*. In: Proceedings of the International Conference on Spoken Language Processing, 6, Sydney, Australia, 1998.

HEIKE, A. E. A content-processing view of hesitation phenomena. *Language and Speech*, 1981.

HORNE, M., FRID, J., ROLL, M. *Hesitation disfluencies after the clause marker att 'that' in Swedish*. Working Papers 51 Lund University, Dept. of Linguistics, 2005, 85–104.

JAEGER, F. *Optional that indicates production difficulty: Evidence from disfluências*. *Proceedings of DiSS'05*, Disfluency in Spontaneous Speech Workshop, Aix-en-Provence, France, 10–12 September 2005.

KENEDY, E. *Gerativismo*. In: MARTELOTTA, Mário Eduardo Toscano. (Org.). In: *Manual de lingüística*. São Paulo: Contexto, 2008. v. 1, p. 127-140.

LEITÃO, M. *Psicolinguística Experimental: Focalizando o processamento da linguagem*. In: Martelotta, M. (org.) *Manual de Linguística*. São Paulo: Contexto, 2008.

LEVELT, Willem. *Speaking: From Intention to Articulation*. Cambridge, Ma, MIT Press. Linguagem. In: MARTELOTTA, Mário Eduardo Toscano (org.) *Manual de Linguística*. São Paulo: Contexto, 1989.

_____. *Monitoring and self-repair in speech*. In *Cognition*, 1983

MATA, A. I. *Apresentação preliminar do CPE Faces: um ‘corpus de Português Europeu falado por adolescentes em contexto escolar’ para o estudo da prosódia dos estilos de fala*. In: Actas do XI Encontro Nacional da APL, vol. I – Corpora. Lisboa, APL/Colibri, 1995.

_____. *Para o Estudo da Entoação em Fala Espontânea e Preparada no Português Europeu: metodologia, Resultados e Implicações Didáticas*, 1999. Tese (Doutorado, F.L.) - Universidade de Lisboa, Lisboa, 1999.

MIOTO, Regina et al. *O estudo da gramática*. In: *Manual de sintaxe*. Insular: Florianópolis, 1999. p. 24-39.

MIRANDA, Fernanda. *O custo de processamento de orações relativas: um estudo experimental sobre relativas com pronome resumptivo no Português Brasileiro*. 2008. Dissertação (Mestrado em Letras) - Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2008.

MODELS OF BILINGUALISM ETC. Modelo de processamento de Levelt (1989).

Disponível em:

<<http://homepage.ntlworld.com/vivian.c/SLA/Bilingualism%20Models%20and%20Memory.htm>>. Acesso em: 17 jul. 2011.

MONIZ, Helena Gorete Silva. *Contributo para a caracterização dos mecanismos de (dis)fluência no português europeu*. Dissertação (Mestrado) em Letras: Lisboa, Universidade de Lisboa, 2006.

NAKATANI, C. H. ; HIRSCHBERG, J. A corpus-based study of repair cues in spontaneous speech. *Journal of the Acoustical Society of America*, 95, 1994.

O’ SHAUGHNESSY, D. *Analysis of false starts in spontaneous speech*. In Proceedings of the International Conference on Spoken Language Processing. Canada, 1992.

PINKER, Steven. *O instinto da linguagem: como a mente cria a linguagem*. São Paulo: Martins Fontes, 2002.

QUINTANILHA, Thiago, OLIVEIRA, Antônio Marcos ; DIAS, Sammy Cardoso. *As disfluências da fala em foco: descrevendo um caso, desvendando mecanismos*, 2010. No prelo.

RAPOZZO, E. Da teoria de P&E ao PM: algumas ideias-chave. In: CHOMSKY, Naom. *O Programa Minimalista*. E. RAPOZZO (trad). Lisboa: Caminho 1999.

SHRIBERG, E. *Preliminaries to a Theory of Speech Disfluency*. 1994 Tese (Doutorado) - Universidade da Califórnia, Califórnia, 1994.

ZHAO, Yuan ; JURAFSKY, Dan. *A preliminary study of Mandarin filled pauses*. In: DISS' 05, Aix-en-Provence, France, 2005.

ANEXO A – Sentenças AOR

Sentenças com disfluências AOR Condição 1

1. a universi... a universitá... a universitária ajudou na tarefa a menina cujo professor saiu de férias.
2. A redado, a redatora entrevistou para o jornal,... o menino,... com o qual... o advogado concordou.
3. O cantor incluiu na bandaáá .. o menino que o apresentador gosta.
4. O a taxista largou na escola o irmão da menina loirinha que estava com ele.
5. O salva-vidas arrastou para a praia o o técnico deste menino que perdeu a hora.

Sentenças com disfluências AOR Condição 2

1. A juíza convocou para audiên, para a audiência.... a menina da qual o estudante precisou....para atestar seu ponto de vista.
2. A arquiteta trouxe pra, para a reunião, a meninada qual o faxineiro havia reclamado.
3. O mágico escondeu dentro da caixa,..... o turista de, O mágico escondeu dentro da caixa, a menina de quem o turista falou. A menina da esquerda...
4. A redatora entrevistou para o jornal...ééé... o menino com quem o advogado concordou.
5. A reda, a redatora entrevistou para o jornal o menino ... ããã de quem o advogado discordou.
6. A juíza convoco convocou para a audiência,..... a menina,... que o estudante precisou.
7. O pala o palhaço entreteve na festa... do menino... que o repórter discordou.
8. A arquiteta trouxe para, para a reunião a menina que o faxineiro reclamou.
9. Atle atleta escolheu para o jogo, uma menina do qual o empresário torceu para ela.
10. O atacante agarrou pela, pelas pernas.... o menino..... que não tinha mais instrutor, pois esse havia pedido demissão
11. O mestre promoveu pra faixa, pra, para a faixa azul,.... o menino que...havia brigado com o gerente e ganhou a luta.
12. A ateleta escolheu para o jogo... a menina cujo empresário tinha torcido.

ANEXO B – Sentenças NAOR

Sentenças com disfluências NAOR Condição 1

1. O mágico escondeu dentro da caixa a menina ééé de cabelos pretos da qual o turista falou.
2. A arquiteta trouxe para a reunião a menina a qual, a menina que o faxineiro reclamou, da qual o faxineiro reclamou.
3. O mestre promoveu para a faixa azul este menino ééé com o qual o gerente brigou.
4. O síndico abordou no corredor um dos meninos lo... morenos de cabelos curtos para quem o escritor telefonou.
5. A redatora entrevistou para o jornal o menino ... que o advogado, com o qual o advogado concordou.
6. A psicóloga ouviu com atenção a menina com qual, com quem o pipoqueiro implicou.
7. O palhaço entreteve na festa o menino que o pa, que o repórter discordou.
8. A juíza convocou para a audiência o meni... o estudante que precisou desta menina.
9. A enfermeira carre, carregou para a maca o menino queee o fotógrafo riu.
10. O síndico abordou no corredor o menino que o, que o escritor telefonou.
11. O motorista esqueceu na escola o, a menina que o monitor viajou.
12. A juíza convocou para a audiência a menina q, que o estudante precisou dela.
13. O vendedor atraiu para a loja o pai da, desta menina o qual passou mal.
14. O mestre promoveu para a faixa azul, um menino de cane, um menino de cabelos negros com quem o gerente havia brigado.
15. A redatora entrevistou para o jornal... um menino ruivo e qu e com sardas que sempre conseguia a concordância do advogado que havia entrevistado...
16. O síndico abordou no corredor o menino para cujo escritor, para cujo escritor telefonou.
17. A redatora entrevistou para o jornal o advoga... o menino cujo advogado concordou.

Sentenças com disfluências NAOR Condição 2

1. A taxista largou na escola a menina que, cujo irmão está com gripe.

2. A arrumadeira fechou no banheiro o menino....ããã cujo pai serviu o exército.
3. O vendedor atraiu para a loja a menina....ããã cujo pai passou mal.
4. A universitária ajudou na tarefa a menina ...ããã cujo pai saiu de férias.
5. O mágico escondeu dentro da caixa ...ããã menina de quem o turista falou.
6. A enfermeira carregou para a maca o menino..... que o fotógrafo riu. Não, de quem o fotógrafo riu.
7. A enfermeira carregou para a maca, o fotógrafo do qual, o o o menino do qual o fotógrafo tinha rido
8. A juíza convocou para a audiência a menina.....ããã que, de quem o estudante precisou.
9. O mestre promoveu para a faixa azul o menino...ééé... com quem alguém brigou.
10. O motorista esqueceu na escola ...a menina ééé com quem o via... ééé.. com quem o monitor viajou.
11. O motorista esqueceu na escolaa menina ééé... com quem o monitor viajou.
12. O mestre promoveu para a faixa azul o menino...que brigou o ge... O mestre promoveu para a faixa azul o menino.....o menino com o qual o gerente brigou.
13. O mestre promoveu para a faixa azul,... o ge, o menino com quem o gerente brigou.
14. A redatora entrevistou para o jornal, o garoto de quem, com quem o advogado concordou. A redatora entrevistou para o jornal o garoto com quem o advogado concordou.
15. A atleta escolheu para o jogo a menina.....ããã que o, de quem, que o empresário torceu. Para quem o empresário torceu.
16. A juíza convocou para a adiên, para a audiência... a menina.... cuj, que o estudante, que a estudante precisou.
17. O síndico abordou um, no corredor um menino.....éééum menino..... que o escritor telefonou.
18. O mestre promoveu para a faixa azul o menino...ããã em que o gerente/ que o gerente brigou.
19. O síndico abordou no corredor o meninoããã para quem o instrutor.... não, com quem o instrutor telefonou..... O síndico abordou no corredor o menino...que que o instrutor telefonou.
20. A redatora entrevistou para o jornal....um adv, o menino cujo o qual o advogado tinha concordado com ele.
21. O motorista esqueceu na escola.... ããã ...menina que, que o monitor tinha viajado com ela.

22. O atacante agarrou pelas pernas...o ins... o menino.. de quem o instrutor pediu demissão.
23. O atacante agarrou pelas pernas o menino.... ããã.... O menino que o instrutor... Não sei.)
24. A universitária ajudou na tarefa.....ééé.... a meninããã do qual o professor saiu de férias.
25. O atacante agarrou pelas pernas...os o menino que o instrutor tinha pedido demissão.
26. A universitária ajudou na tarefa, a meni, uma das meninas, a que havia saído de férias.uma
27. A taxista largou na, na escola a menina ...ããã ããã de quem o ir, que o irmão pegou gripe.
28. O salva-vidas apresentou para a praiaa menina....ããã de que o técnico perdeu a hora.
29. O atacante agarrou pelas pernas.... o menino.....ããã que o instrutor pediu demissão.

ANEXO C – Sentenças DPOR

Sentenças com disfluências DPOR Condição 1

1. A arrumadeira fechou no banheiro o menino cujo tio serviu e... ao exército.
2. O salva-vidas arrastou para a praia... o menino cujo técnico perdeu a, a hora.
3. A arquiteta trouxe para a reunião a menina que ooo faxineiro reclamou.
4. A arrumadeira fechou no banheiro o menino ruivo que era sobri, sobrinho de um coronel do exército.
5. A universitária ajudou na tarefa a menina que saiu de férias sem, estava sem ajuda do professor.
6. O palhaço entreteve na festa, um menino... louro... que nunca havi..., não conseguiu concordância do repórter.
7. O motorista esqueceu na escola... uma menina loura de cabelos encaracolados...cujo monitor havia viajado... O motorista esqueceu na escola uma menina loura de cabelos encara, encaracolados que havia viajado com o monitor.

Sentenças com disfluências DPOR Condição 2

1. A arquiteta trouxe para a reunião... a menina de quem ooo faxineiro reclamou.
2. A atleta escolheu para o jogo a menina que o o empresário torceu.
3. O motorista esqueceu na escola,..... a menina, ...que havia viajado com o monitor.... e não tinha, che, chegado na hora.
4. A psicóloga ouviu com atenção, a menina, cujo pipoqueiro criou uma certa...A psicóloga ouviu com atenção uma das meninas, aquela com a qual o pipoqueiro criou uma certa implicância. A psicóloga ouviu com atenção uma das meninas, porque ela reclamava... que o pipoqueiro vivia implicando com ela.