



**UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS HUMANAS, LETRAS E ARTES
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA
DOUTORADO EM LINGUÍSTICA**

ALTHIERE FRANK VALADARES CABRAL

**A INFLUÊNCIA DA ANIMACIDADE NO PROCESSAMENTO DE RELATIVAS
DE SUJEITO E DE OBJETO NO PORTUGUÊS BRASILEIRO E EUROPEU**

**JOÃO PESSOA
2016**

UFPB
BIBLIOTECA CENTRAL

ALTHIERE FRANK VALADARES CABRAL

**A INFLUÊNCIA DA ANIMACIDADE NO PROCESSAMENTO DE RELATIVAS
DE SUJEITO E DE OBJETO NO PORTUGUÊS BRASILEIRO E EUROPEU**

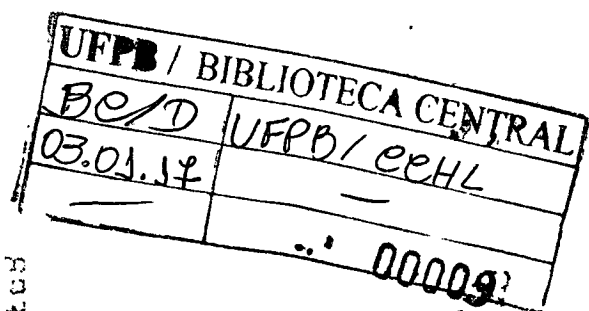
Tese apresentada ao Programa de Pós-Graduação em
Linguística da Universidade Federal da Paraíba,
UFPB, como requisito parcial para obtenção do título
de Doutor em Linguística, sob a orientação do
Professor Doutor **Márcio Martins Leitão**, da UFPB,
e da co-orientação do Professor Doutor **Eduardo
Kenedy** da Universidade Federal Fluminense, UFF.

COESP
801 (043)
C447i

JOÃO PESSOA
2016

ALTHIERE FRANK VALADARES CABRAL

UFPB
BIBLIOTECA CENTRAL



C117i Cabral, Althiere Frank Valadares.
A influência da animacidade no processamento de
relativas de sujeito e de objeto no português brasileiro e
europeu / Althiere Frank Valadares Cabral.- João Pessoa,
2016.
149f.
Orientador: Márcio Martins Leitão
Coorientador: Eduardo Kenedy
Tese (Doutorado) - UFPB/CCHL
1. Linguística. 2. Cláusulas relativas. 3. Sujeito e objeto.
4. Animacidade. 5. Semântica. 6. *Parser*. 7. Frequência.

UFPB/BC

CDU: 801(043)

ALTHIERE FRANK VALADARES CABRAL

**A INFLUÊNCIA DA ANIMACIDADE NO PROCESSAMENTO DE RELATIVAS
DE SUJEITO E DE OBJETO NO PORTUGUÊS BRASILEIRO E EUROPEU**

Aprovado em ____ / ____ / ____

BANCA EXAMINADORA

____ Nota ____
Prof. Dr. Márcio Martins Leitão
(Orientador)

____ Nota ____
Prof. Dr. Eduardo Kenedy
(Co-orientador)

____ Nota ____
Prof^a. Dr^a. Mahayana Cristina Godoy
(Examinadora Externa)

____ Nota ____
Prof^a. Dr^a. Elisângela Nogueira Teixeira
(Examinadora Externa)

____ Nota ____
Prof. Dr. José Ferrari Neto
(Examinador Interno)

____ Nota ____
Prof^a. Dr^a. Rosana Oliveira
(Examinadora Interna)

Média ____

A Diadorin, Sofia e Tales: meus três motivos.

AGRADECIMENTOS

Ao professor *Márcio Leitão*, que na condição de orientador, superou minhas expectativas, que já eram das melhores. Apesar de toda sabedoria e competência, sua orientação para mim fica marcada pela humildade e simplicidade com que conduziu tudo que me ensinou, sempre respeitando meu tempo de trabalho. Rigoroso na metodologia, preciso na visão teórica, e antes de tudo, cuidadoso na condução deste trabalho. Pude contar com ele em todos os momentos da pesquisa, suas contribuições sempre foram as mais férteis possíveis. Contudo, para além da figura do professor, creio ter construído, no decorrer deste trabalho, uma amizade bastante sólida, que aumenta ainda mais a admiração que tenho pelo grande profissional que ele é. E por meio dessa amizade, conheci e Arthur e Alice, seus filhos, de quem levo as mais lindas lembranças, lembranças que carregarei na memória, e na “*carcunda*”, palavra cheia de significados para meu grande amigo Arthur. Agradeço também à sua esposa Miriam. Conviver com sua família em Portugal foi uma das melhores surpresas deste doutorado.

Como co-orientador, tive o privilégio de ter *Eduardo Kenedy*. A competência, o conhecimento, a perspicácia para os fatos da língua fazem desse professor um profissional sem igual, pude aprender com ele muito mais do que poderia supor. Mas nossa relação também não se limitou à pesquisa, e hoje o tenho em alta conta na lista de meus amigos mais queridos. Sua contribuição para este trabalho também é inestimável.

Mais uma pessoa se apresentou como uma grande surpresa no decorrer da construção desta tese. Em Portugal, a proposta era tão somente de fazermos alguns experimentos na Universidade do Minho, mas quem me acolheu naquela instituição foi a professora *Maria do Carmo Lourenço-Gomes*, a doce Cacao. Nunca terei como agradecer o carinho com que me acolheu e com que me orientou. Foram poucas semanas de convívio, porém, dessas semanas nasceu uma parceria muito fértil e feliz.

A *Conceição e José Cabral*, meus pais, a quem devo a formação primeira que me fez entender o que era a vida.

A meu irmão *Zé Brígido*, e a minha cunhada *Mavi*, que talvez sejam, neste mundo, as pessoas que mais torcem por mim, e por quem serei eternamente grato.

Ao meu outro irmão posticho, *Iradilson*. A ele tenho que agradecer por ter sido o amigo que mais me ajudou nesta tese, e ele sequer é linguista. Mas ele traduziu muitos textos, revisou tantos outros, me emprestou seus alunos para que eu pudesse realizar parte dos experimentos, e, sobretudo, me apoiou quando mais precisei. Construímos nossa amizade no decorrer do processo de doutorado e ele acabou sendo adotado por minha família inteira. Sei que tenho um irmão com quem posso contar para o resto da vida.

A *Gitanna Bezerra*, uma colega cuja competência se destaca de maneira espetacular. Recorri a ela algumas vezes no decorrer da pesquisa, e ela sempre foi incrivelmente atenciosa, mostrou-me caminhos novos e interessantes, leu meus textos, corrigiu e enriqueceu muito este trabalho.

A *Patrícia Tondinelli*, minha amiga-irmã, perto ou longe, estamos conectados.

Ao professor *José Ferrari*, que tantas vezes me socorreu na minha luta com a estatística, e à professora *Rosana Oliveira* pelo apoio que sempre me deu.

A três amigos que foram fundamentais para que os experimentos acontecessem: *Maria Tânia*, *Fábio Alexandre* e *Dayvyd Lavaniery*. Todos professores do IFRN, cederam-me seus alunos para que eu pudesse realizar os experimentos com eles.

Ao Nuno e à Sofia, que me acolheram em Portugal.

A todos os *colegas do LAPROL*, pela parceria em vários momentos.

A todos os voluntários que participaram dos experimentos realizados no Instituto Federal do Rio Grande do Norte e na Universidade do Minho, em Portugal.

Ao *IFRN*, instituição na qual trabalho e que fomentou meu doutorado.

Aos meus alunos.

E por fim, a todos os meus amigos, que mesmo não estando diretamente relacionados à escrita desta tese, motivaram-me, torceram por mim, acreditaram no meu trabalho.

Poema marginal

*Havia um rio, de margens estreitas, de modo que de um lado se podia ver o outro
lado
No meio do nada, aquele pedaço do Pandeiros, filho do velho Chico
Se ia
Pelo caminho, meu pai me puxava o cavalo
Por lá, barulhos de bichos
E aquela conversa bonita sobre os pássaros
Causos
e canções
Fecha a porta, mariquinha
E aquelas não-me-toque se fechavam
Havia um rio
e que jamais então houve de novo
Jamais o mesmo rio que banhava a menina
Jamais a mesma menina que se banhava no rio
se suas margens eram estreitas
de modo que de um lado se podia ver o que vinha vindo do outro lado
aquele infinito de águas eu não ousei provocar
nunca atravessei
foi ele quem me atravessou*

Diadorin Pietà Nunes Cabral

RESUMO

Nesta tese, investigamos a influência do traço de animacidade sobre o processamento de cláusulas relativas de extração de sujeito e de objeto, levando em consideração dados do Português Brasileiro e Europeu. Para tanto, realizamos um estudo de *corpus* e cinco estudos experimentais utilizando a técnica experimental da leitura automonitorada. Muitos estudos sobre processamento de sentenças afirmam que há uma assimetria entre cláusulas relativas de sujeito e de objeto. A maioria desses estudos mostra que sujeitos são mais rápidos para processar do que objetos em cláusulas relativas. As razões para esta assimetria, em geral, são ocasionadas pela complexidade sintática dessas cláusulas. Tais estudos normalmente não controlam o traço de animacidade e os estímulos usados nestes experimentos geralmente possuem antecedentes animados para a extração de sujeito e de objeto. Alguns trabalhos, entretanto, indicaram que o traço de animacidade é um fator relevante não apenas durante o processamento de sentenças, mas também na distribuição de padrões de cláusulas relativas em línguas naturais. A hipótese que assumimos aqui é que se o traço de animacidade realmente influencia a distribuição de cláusulas relativas, então o processamento *on-line* dessas estruturas seria afetado por esse traço. O estudo de corpus foi realizado em Português Brasileiro e Europeu. Os resultados mostram que cláusulas relativas de sujeito e objeto animados não representam a estrutura mais usual no Português. O objeto animado é especialmente raro na língua. Baseado nesta informação, estudos experimentais foram realizados. Nos dois primeiros experimentos, ambos idênticos, um em Português Brasileiro e o outro em Português Europeu, pretendíamos mensurar o tempo de processamento em cláusulas relativas de sujeito e objeto, todas elas com um traço animado. A partir destes experimentos, poderíamos confirmar e corroborar com o que é apresentado pela literatura no que diz respeito a este assunto: cláusulas relativas de sujeito foram realmente mais rápidas para processar. Nos experimentos 3 e 4, respectivamente em Português Brasileiro e Europeu, quatro condições foram testadas, sujeito animado e inanimado, assim como objeto animado e inanimado em cláusulas relativas. Resultados significantes foram obtidos e eles indicam um efeito principal de traço de animacidade. Em várias medidas, os objetos inanimados foram lidos ainda mais rapidamente do que os sujeitos. No quinto experimento, aplicamos a técnica de leitura automonitorada, contudo, com palavras isoladas. Testamos separadamente substantivos animados e inanimados. Os resultados deste último experimento não foram significativos. Isto mostra que o traço de animacidade provavelmente não atua quando apenas palavras isoladas são lidas, mas em vez disso este traço atua quando a distribuição da estrutura frasal é apresentada. Visando analisar nossos resultados, aplicados à DLT, a Teoria de Dependência de Localidade, de Gibson (2000), como um modelo teórico básico. Segundo essa teoria, o *parser* está apto para acessar, entre outras fontes de informação, a frequência de distribuição de padrões animados na língua e faz isso nos estágios iniciais do processamento. Em outras palavras, a frequência das estruturas com itens lexicais animados e inanimados foi determinante no processamento de cláusulas relativas de sujeito e objeto. A assimetria entre estas estruturas foi confirmada, mas foi refutada quando o traço de animacidade foi apropriadamente controlado.

Palavras-chave: Relativas, sujeito, objeto, animacidade, semântica, *parser*, frequência.

ABSTRACT

In this work, we investigate the influence of the animacy feature over the processing of subject and object extraction in relative clauses, taking into consideration Brazilian and European Portuguese data. In order to achieve this goal, a corpus study was performed as well as five experimental studies through a self-paced experimental technique. Many studies on sentence processing claim there is an asymmetry between object and subject relative clauses. The majority of these studies show that subjects are faster to process than objects in relatives clauses. The reasons for this asymmetry, in general, are due to syntactic complexity of these clauses. Such studies normally do not control the animacy feature and the stimuli used in those experiments generally have animate antecedents for both subject and object extraction. Some works, however, have indicated that animacy feature is a relevant factor not only during sentence processing, but also in the distribution of relative clause patterns in natural languages. The hypothesis we assume here is that if the animacy feature really influences the distribution of relative clauses, then the on-line processing of these structures would be affected by this feature. A corpus study was performed in Brazilian and European Portuguese. The results show that animated subject and object in relative clauses do not represent the most usual structure in Portuguese. The animated object is especially rare in the language. Based on this information, experimental studies were made. In the two initial experiments, both identical, one in Brazilian Portuguese and the other in European Portuguese, we intended to measure the processing time in both subject and object of relative clause, all of them with an animate feature. From these experiments, we could confirm and corroborate what is presented by the literature concerning this topic: subject relative clauses were indeed processed faster. In the experiments 3 and 4, respectively in Brazilian and European Portuguese, four conditions were tested, animated and inanimated subject as well as animated and inanimated object in relative clauses. Significant results were achieved and they indicate an animacy feature main effect. In several measurements, the inanimated object was read even faster than subjects. In the fifth experiment, we applied the self-paced reading technique, however, with isolated words. We tested separately animate and inanimate nouns. The results of this last experiment were not significant. This shows that the animacy feature probably does not act when only isolated words are read, but instead this feature takes over when a phrasal structure distribution is presented. In order to analyze our results, we applied the DLT, Dependency Locality Theory, from Gibson (2000), as a basic theoretical model. According to this theory, the parser is able to access, among other sources of information, the frequency of distribution of animacy patterns in the language and it does so in the early stages of the processing. In other words, the frequency of structures with animated and inanimated lexical items was determinant in processing subject and object relative clauses. An asymmetry between these structures was confirmed, but this asymmetry was undone when the animacy feature was properly controled.

Key-words: Relatives, subject, object, animacy, semantics, parser, frequency.

LISTA DE FIGURAS

Figura 1	Figuras retiradas do trabalho de Mangas (2011, p. 60).	31
Figura 2	Figuras retiradas do trabalho de Mangas (2011, p. 60).	32
Figura 3	Gráfico representativo do <i>corpus</i> escrito do PB	79
Figura 4	Gráfico representativo do <i>corpus</i> falado do PB	80
Figura 5	Gráfico representativo do <i>corpus</i> escrito e falado do PB	81
Figura 6	Gráfico representativo do <i>corpus</i> escrito do PE	82
Figura 7	Gráfico representativo do <i>corpus</i> falado do PE	83
Figura 8	Gráfico representativo do <i>corpus</i> escrito e falado do PE	84
Figura 9	Participantes voluntários da pesquisa durante a realização do experimento, alunos do Instituto Federal do Rio Grande do Norte, em Natal, RN, Brasil.	94
Figura 10	Participantes voluntários da pesquisa durante a realização do experimento, alunos da Universidade do Minho, em Braga, Portugal.	94
Figura 11	Resultado do tempo de resposta - PB	98
Figura 12	Resultado do tempo de processamento do segmento crítico - PB e PE	102
Figura 13	Gráfico comparativo – Tempo de Leitura – Segmento Crítico – PB e PE	122
Figura 14	<i>Corpora</i> – escrito e falado, em PB e PE - Comparativo	122
Figura 15	Relativas de sujeito e de objeto	127

LISTA DE TABELAS

Tabela 1	<i>Corpus</i> do Holandês e do Alemão (MAK, VONK E SCHRIEFERS, 2002)	36
Tabela 2	Condições Experimentais do 1º Experimento de Mak, Vonk E Schriefers (2002)	38
Tabela 3	Principais modelos de processamento de frase (KENEDY, 2015, p. 153)	45
Tabela 4	Resultados de Frazier e Clifton, (1997).	57
Tabela 5	Exemplos do frases experimentais de Gilboy et al. (1995).	57
Tabela 6	Resultados de Schafer, Clifton e Frazier, (1996).	58
Tabela 7	<i>Corpora</i> – escrito e falado, em PB e PE.	85
Tabela 8	<i>Corpora</i> do Holandês, do Alemão e do Português	86
Tabela 9	Proporção de erros e acertos para as relativas de sujeito e objeto em PB.	97
Tabela 10	Resultados do tempo de processamento do segmento crítico do PB.	99
Tabela 11	Proporção de erros e acertos para as relativas de sujeito e objeto em PE.	100
Tabela 12	Resultados do tempo de processamento do tempo de resposta em PE	101
Tabela 13	Resultados do tempo de processamento do segmento crítico do PE	101
Tabela 14	Quadrado latino dos experimentos 3 e 4	108
Tabela 15	Proporção de erros e acertos para as relativas de sujeito e objeto em PB.	110
Tabela 16	Resultados do tempo de resposta – PB	111
Tabela 17	Resultados do tempo de leitura do segmento crítico – S2 – em PB.	112

Tabela 18	Proporção de erros e acertos para relativas de sujeito e objeto em PE.	113
Tabela 19	Resultados do tempo de resposta – PE	114
Tabela 20	Resultados do tempo de leitura do segmento crítico – S2 – em PE.	115
Tabela 21	Tempo de leitura da primeira palavra	119
Tabela 22	Tempo de resposta	120

LISTA DE SIGLAS

- AFS** – Estratégia do antecedente Ativo
- CRO** – Cláusula Relativa de Objeto
- CROa** – Cláusula Relativa de Objeto Animado
- CROi** – Cláusula Relativa de Objeto Inanimado
- CRS** – Cláusula Relativa de Sujeito
- CRSa** – Cláusula Relativa de Sujeito Animado
- CRSi** – Cláusula Relativa de Sujeito Inanimado
- DLT** – Teoria da Dependência Local
- FF** – Forma Fonética
- FL** – Forma Lógica
- ms** – Milésimos de segundo
- PB** – Português Brasileiro
- PE** – Português Europeu
- SC** – Sintagma Complementizador
- SN** – Sintagma Nominal
- SP** – Sintagma Preposicional
- SPLT** – Teoria da Predição Sintática Local
- SV** – Sintagma Verbal
- TGP** – Teoria do *Garden-Path*

SUMÁRIO

1.	CAPÍTULO 1 – Introdução	16
2.	CAPÍTULO 2 – Estado da arte	22
2.1	Estrutura da Relativa	22
2.2	Estudos sobre assimetria entre CRS e CRO	29
2.3	A influência do traço da animacidade no processamento de sentenças relativas	35
2.4	Modelos teóricos de processador sintático	45
2.5	A animacidade: uma breve definição	71
3.	CAPÍTULO 3 – Estudo de <i>corpus</i>	77
3.1	Sobre os <i>corpora</i>	77
3.2	Português Brasileiro	79
3.3	Português Europeu	83
3.4	Comparação PB/PE	85
4.	CAPÍTULO 4 – Estudos experimentais	88
4.1	Experimentos 1 e 2 – Cláusulas relativas de sujeito e objeto com antecedentes animados em PB e PE	90
4.2	Experimentos 3 e 4 – CRS e CRO com antecedentes animados e inanimados em PB e PE	104
4.3	Experimentos 5 – Leitura Automonitorada com palavras isoladas	117
4.4	Discussão geral	122
4.5	Nossa análise e os modelos teóricos de processamento	128
5.	CAPÍTULO 5 – Considerações finais	134
5.1	<i>“The cost of the unlikely structure is balanced by the benefits of having a simple structure”</i>	134
	Referências	139
	Anexos	145

*Feliz de quem vive ainda no mundo dos substantivos:
O resto é literatura...*
– Mario Quintana, Baú dos espantos.

CAPÍTULO 1

Introdução

O processamento de frases tem ocupado parte importante das investigações da Psicolinguística Experimental nos últimos anos. Com o intuito de investigar o que acontece quando compreendemos e produzimos frases, pesquisadores vêm desenvolvendo métodos e aplicando técnicas que descrevem e explicam fenômenos linguísticos de maneira cada vez mais refinada.

Neste trabalho, debruçamo-nos sobre o processamento de frases, mais especificamente, sobre o processamento de cláusulas relativas de sujeito e de objeto. Nosso intuito é compreender como o *Parser*, o processador sintático, procede, quando do processamento de relativas, considerando-se um possível efeito do traço da animacidade¹ ao longo desse processo.

Este trabalho teve início no Brasil, na cidade de Natal, no Rio Grande do Norte. A princípio, intencionávamos pesquisar a influência da animacidade no processamento de relativas no Português Brasileiro. No decorrer do processo, porém, compreendemos que a ampliação da abrangência da pesquisa para outra variedade da língua poderia ser bastante produtiva no sentido de compreendermos se os fenômenos investigados no Português Brasileiro encontram evidências também no Português Europeu ou se são fatos observáveis apenas em uma dessas variedades.

Diante disso, parte dos nossos experimentos, bem como do estudo de *corpus* a que nos referiremos mais adiante, foi feita na cidade de Braga, em Portugal. Dito isso, estamos assumindo que há nesta pesquisa um caráter comparativo entre as variedades Brasileira e Europeia da Língua Portuguesa.

O mote desta pesquisa foi, a princípio, a assimetria no processamento dessas cláusulas relativas de sujeito e de objeto. Muitos trabalhos em psicolinguística (cf. FORD, 1983; KING & JUST, 1991; KING & KUTAS, 1995;

¹ Animacidade – termo que se refere, grosso modo, à propriedade de um elemento possuir ou não possuir vida. Na seção 2.5 do segundo capítulo, definiremos o termo para os propósitos desta tese.

FRAUENFELDER, SEGUI & MEHLER, 1980; HOLMES & O'REGAN, 1981; MECKLINGER, et al., 1995; SCHRIEFERS, FRIENDERICI & KÜHN, 1995; FRAZIER, 1987; OLIVEIRA, 2013; COSTA, et al., 2009; MANGAS, 2011, MAK, VONK & SCHRIEFERS, 2002 & 2006) vêm apontando para uma diferença entre o processamento dessas estruturas, a maior parte deles demonstrando as cláusulas relativas de sujeito, doravante CRS, como estruturas mais fáceis de serem processadas do que as cláusulas relativas de objeto, doravante CRO, uma vez que são lidas em menos tempo, exibindo, pois, um menor custo de processamento. As estruturas de que tratamos estão representadas nas frases abaixo:

1. O rapaz [que ☐ convidou você] estava furioso.
2. O rapaz [que você convidou ☐] estava furioso.

Trabalhos há, contudo, que relativizam essa assimetria. Mak, Vonk e Schiefers (2002, 2006), por exemplo, trazem resultados interessantes a partir de uma variável normalmente desconsiderada na maioria dos estudos que testaram a assimetria entre CRS e CRO, o traço da animacidade. Com efeito, perceba-se que, nos exemplos acima, tanto o sujeito quanto o objeto no domínio das relativas possuem o traço “mais animado” (e “mais humano, inclusive). A partir de um estudo de *corpus* e da aplicação das Técnicas Experimentais de Leitura Automonitorada e de Rastreamento Ocular, os autores controlaram a animacidade, e os resultados normalmente encontrados em outros trabalhos foram relativizados. Ou seja, a partir da manipulação e do controle de um traço de natureza semântica, parece haver uma mudança no processamento dessas sentenças. Para ilustrar esta discussão, apresentamos abaixo exemplos das estruturas a respeito das quais estamos discorrendo.

3. **(CRSa) Relativa de sujeito animado:**
A mulher [que ☐ matou o rapaz] é muito bonita.
4. **(CRSi) Relativa de sujeito inanimado:**
A pedra [que ☐ matou o rapaz] não era muito grande.
5. **(CROa) Relativa de objeto animado:**
A mulher [que o rapaz matou ☐] é muito bonita.
6. **(CROi) Relativa de objeto inanimado:**
A pedra [que o rapaz jogou ☐] não era muito grande.

Sem, por ora, atermo-nos a questões de cunho teórico em sentido *stricto*, vamos agora apresentar, ainda que sucintamente, alguns dos modelos de processador sintático com os quais a psicolinguística trabalha atualmente. Tais modelos se diferem em alguns aspectos. A este trabalho interessa sobretudo o fato de estar em discussão o tipo de informação a que o *parser* tem acesso nos primeiros estágios da compreensão da frase.

Modelos mais modulares compreendem o *parser* como modular, serial e incrementacional, como é o caso da Teoria do *Garden-Path*, de Frazier (1979). Visto assim, trabalhando de maneira simples, o *parser* dá conta do processamento sintático de uma dada estrutura em tempo ótimo. Por outro lado, modelos mais conexionistas compreendem que o falante tem acesso a informações de vários tipos, ou seja, informações sintáticas são acessadas simultaneamente com informações, por exemplo, semânticas e pragmáticas.

Posto de outra forma, o questionamento que se faz é o seguinte: o *parser*, no momento inicial do processamento, analisaria apenas informações sintáticas, o que é conhecido como encapsulamento sintático, ou ele teria acesso a informações de natureza lexical, semântica, pragmática ou mesmo discursiva?

Para tanto, faz-se necessário considerar que, para a teoria do *Garden-Path*, que será mais detalhada no próximo capítulo, há estágios distintos no processamento linguístico. A priori, no primeiro estágio do processamento, a que nos referimos como momento *on-line*, o *parser* opera de maneira a atribuir representações estruturais ao *input* linguístico. Num segundo momento, num estágio mais tardio, as informações de natureza semântica, pragmática e discursiva entram em cena, ao que chamamos de momento *off-line*. O primeiro estágio é, portanto, reflexo, automático, distinto do segundo que é mais reflexivo. A separação entre esses estágios não é fácil ou clara, haja vista serem processos que ocorrem em frações de segundo.

Interessa-nos, particularmente, dois modelos de processamento, cujas perspectivas teóricas se distinguem. Para uma delas o *parser* é definido como um mecanismo modular, que analisa a estrutura exclusivamente a partir de informações sintáticas no primeiro estágio da compreensão, para a outra perspectiva, o processamento é mais interativo, em que as informações

estruturais e não estruturais são processadas simultaneamente, de maneira que o *parser* poderia avaliar diversas análises ao mesmo tempo, e no caso de estruturas ambíguas, construir diversas estruturas sintáticas em paralelo, deixando para escolher aquela que se apresentar como melhor no decurso do processamento linguístico, ao fim da sentença.

Para além dessas duas concepções, há modelos interativos, que pressupõem, por um lado, um *parser* incremental e serial, porém, sem o chamado encapsulamento sintático dos modelos modularistas. Tais modelos são de especial importância para as análises e discussões que esta tese pretende empreender. A DLT² – The Dependency Locality Theory – (GIBSON, 2000), um modelo que se encaixa nos pressupostos dos modelos interativos, será devidamente discutida por seus pressupostos na seção 2.4.4, do próximo capítulo.

Mais adiante neste trabalho, discutiremos com a devida atenção alguns modelos de processamento que descrevem esses estágios. Nesta introdução, vale-nos afirmar que o estudo que ora se instaura pretende lançar a discussão sobre a natureza das informações acessadas pelo *parser* no primeiro momento da compreensão da frase, ou mais detidamente, da compreensão das cláusulas relativas de sujeito e de objeto no tocante à influência da frequência do traço da animacidade sobre essas estruturas.

Para esta empreitada, para além da revisão bibliográfica concernente ao tema, valer-nos-emos de um estudo de *corpus*, e no tocante à parte experimental, realizamos cinco experimentos por meio da Técnica Experimental de Leitura Automonitorada, sendo que o último foi feito com palavras isoladas.

A organização formal desta tese dar-se-á em 5 capítulos. O primeiro é composto por esta introdução, que de maneira breve descreve o fenômeno que pretendemos abordar.

No segundo capítulo, pretendemos descrever o estado da arte, quando elencaremos alguns trabalhos que abordam o tema de nossa pesquisa com uma breve discussão em torno de seus resultados. O capítulo será composto por 5 seções, na primeira (2.1), descreveremos a estrutura da relativa do ponto

² Teoria da Dependência Local – Manteremos a sigla tal qual a teoria é conhecida, no Inglês – DLT.

de vista sintático. Na segunda (2.2), abordaremos alguns estudos sobre assimetria entre CRS e CRO. Na terceira (2.3), discutiremos alguns resultados experimentais que testaram e discutiram a influência da animacidade em estruturas relativas. Na quarta seção (2.4), traremos à baila os modelos de processamento que nos interessam à discussão ora abordada. Na subseção (2.4.4), discutiremos a DLT, de Gibson (2000), que assumiremos nesta pesquisa como modelo que abarca o fenômeno aqui investigado. Na quinta seção (2.5), discutiremos o traço da animacidade para delineação de nossos propósitos nesta tese.

O terceiro capítulo trará os resultados de dois estudos de *corpus* que orientam nossas análises no que tange à frequência do fenômeno investigado, o primeiro em Português Brasileiro – PB – e o segundo em Português Europeu – PE. Com esses *corpora* analisados, pretendemos discutir a distribuição de CRS e CRO no que tange à animacidade do termo antecedente da relativa.

No quarto capítulo, os cinco experimentos por nós realizados serão descritos e discutidos em quatro seções distintas. Na primeira seção, (4.1), descreveremos e discutiremos os resultados dos experimentos 1 e 2, em que testamos, por meio de leitura automonitorada, CRS e CRO com termos animados, respectivamente em Português Brasileiro e Europeu. O objetivo é averiguar se, controlando sistematicamente a animacidade, trazendo todos os antecedentes com o traço animado, as relativas de sujeito serão mais fáceis de processar. Nossa hipótese é de que serão, conforme aponta literatura em psicolinguística. Na segunda seção, (4.2), discutiremos os experimentos 3 e 4, em que controlamos o traço da animacidade de CRS e CRO. O experimento 3 foi feito em PB e o experimento 4 foi feito em PE. Partimos da hipótese de que a manipulação da animacidade dos termos antecedentes mudará o tempo de processamento das relativas e as CRS não serão, necessariamente, as estruturas mais rápidas de serem lidas. Na terceira seção (4.3), traremos os resultados do experimento 5, feito com palavras isoladas animadas e inanimadas do PB, partimos da hipótese de que a animacidade só será acessada na estrutura frasal, e palavras isoladas não serão lidas com tempos diferentes apenas pela manipulação do traço. Na quarta seção (4.4), faremos uma discussão geral de nossos resultados com base nos pressupostos teóricos

assumidos na pesquisa e na seção (4.5) discutiremos nossos resultados diante dos modelos teóricos de processador sintático aqui discutidos.

Por fim, no quinto capítulo faremos nossas considerações finais, quando discutiremos os resultados de todos os experimentos realizados, bem como as análises feitas a partir do estudo de *corpus* para, diante de nossa discussão teórica, termos uma descrição adequada dos fenômenos aqui empreendidos com o intuito de explicá-los.

*Na noite de estreia, depois que todos se foram e as luzes se apagaram, a bailarina
deitou-se no fio e adormeceu*

– Marina Colasanti, Um espinho de marfim.

CAPÍTULO 2

Estado da Arte

A complexidade de uma cláusula relativa bem como seu papel central na recursividade fazem dessa estrutura um ponto de discussão fundamental à descrição e à explicação do aparato humano de produção e compreensão da linguagem.

Segundo Costa et al. (2009, p. 211), as relativas “são um indicador válido por se reconhecer que são uma área em que diferentes populações encontram dificuldades, não havendo, contudo, consenso relativamente à natureza da dificuldade”. Cláusulas relativas são estruturas linguísticas encaixadas em outras estruturas por meio de adjunção. Sintagmas complexos que, recursivamente, podem modificar um sintagma nominal, SN, adjungindo-lhe tantas propriedades quanto o discurso corrente e a memória de trabalho do falante/ouvinte permitirem.

Disto isso, neste capítulo, revisaremos alguns trabalhos que discutem as relativas de sujeito e de objeto para compreender melhor como essas estruturas são processadas na mente humana do ponto de vista da psicolinguística.

2.1 Estrutura da Relativa

Para uma descrição das relativas e para uma melhor elucidação do fenômeno que abordaremos neste trabalho, trazemos os exemplos que se seguem:

7. A menina foi à festa.
8. A menina que conheci foi à festa.
9. A menina que conheci que usava os óculos foi à festa.
10. A menina que conheci que usava os óculos que foram roubados foi à festa.

As sentenças acima relacionadas ilustram o encaixamento de estruturas relativas. Em 8, “*que conheci*”, uma estrutura sentencial se encaixou à outra estrutura por meio de um pronome relativo. Em 9, duas estruturas, “*que conheci*” e “*que usava os óculos*” foram encaixadas. Em 10, foram três estruturas, “*que conheci*”, “*que usava os óculos*” e “*que foram roubados*”. A princípio, poderíamos encaixar várias outras estruturas, uma vez que a relativa é uma estrutura adjungida. Contudo, outros fatores apresentam-se como limitadores quando aumentamos a complexidade de encaixe das relativas.

Por ora, é relevante considerar que, em uma língua natural, temos estruturas que funcionam como predicadores, como complementos, ou como adjuntos. As relativas, em específico, são adjuntos oracionais, a respeito dos quais discutiremos mais adiante.

Os predicadores projetam seus argumentos na frase. Noutras palavras, são termos que suscitam outros termos para que possam se realizar no discurso. Um exemplo de predicadores são os verbos. Um verbo se realiza em sua forma finita quando da projeção de seus argumentos.

Assim, o verbo “amar”, por exemplo, só terá sua grade argumental saturada quando da inserção de um sujeito e de um objeto. Ou seja, no discurso, esse verbo não se realiza como “amar Maria” ou “João amar”, mas como, por exemplo, “João ama Maria”. Sendo o verbo “amar” um verbo transitivo direto, ele predica ao projetar seus dois argumentos necessários à construção do discurso, mesmo que esses argumentos, eventualmente, não sejam foneticamente realizados.

Em contrapartida, o SN “Maria” foi complemento do verbo, elemento sem o qual não teríamos a completude dessa estrutura. Dizemos com isso que o “objeto direto” funciona como complemento, elemento sem o qual não podemos ter uma sentença bem formada, pelo menos nos casos em que eles são necessários.

Há termos do discurso, porém, que não predicam e tampouco são necessários à completude de uma estrutura argumental, são os termos adjungidos, tradicionalmente chamados de adjuntos. Logo, “João ama Maria” é constituído por um predicador e dois argumentos, um interno (objeto direto) e um externo (sujeito). Contudo o discurso possui ainda estruturas que não predicam e que não são exigidas por predicadores, assim temos:

11. João ama Maria incondicionalmente.

Perceba-se que “incondicionalmente” modifica a estrutura sentencial, mesmo não tendo sido uma exigência necessária à completude da frase. As cláusulas relativas são adjuntos cuja estrutura nos é fundamental neste trabalho.

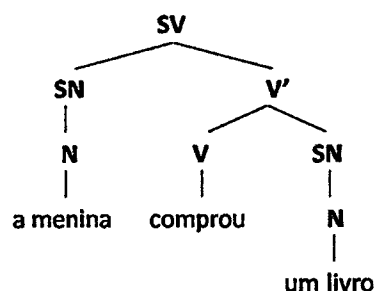
Para maior clareza, descrevemos abaixo as árvores sintáticas que nos são importantes à compreensão do fenômeno linguístico que pretendemos investigar neste trabalho. Para tanto, tomaremos a representação arbórea da teoria X-Barra, uma vez que, conforme Kenedy (2013, p. 196), a teoria X-Barra

é interessante e útil justamente porque ela nos oferece um modelo de representação arbórea capaz de dar conta de todos os tipos de relação sintática, seja a de um núcleo e complemento, ou a de especificador e núcleo ou ainda a de adjunção entre sintagmas.

Como pretendemos compreender a estrutura de uma cláusula relativa, interessa-nos, sobremaneira, representar a constituição das estruturas adjungidas. A seguir, representamos a relação de um núcleo com o seu complemento (argumento interno) e com o seu especificador (argumento externo).

12.

A menina comprou um livro.



Em 12, temos a estrutura de uma sentença com verbo transitivo direto. Note-se: “comprar” se apresenta saturado com dois argumentos, um interno, “um livro”, e outro externo “a menina”.

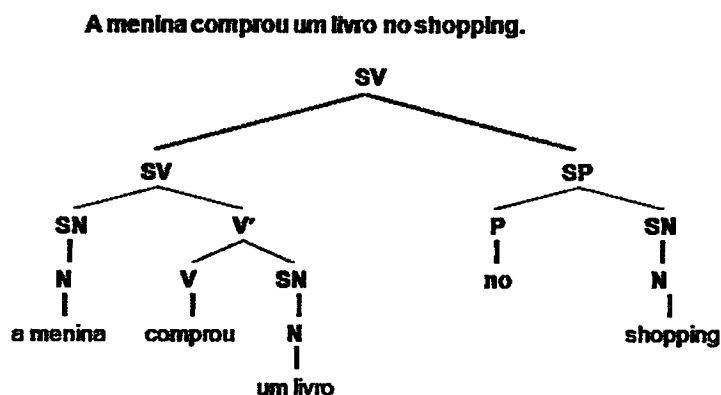
Cabe ser posto, porém, que a concatenação entre o verbo e seus argumentos não se dá ao mesmo tempo. Primeiro ocorre a concatenação, ou seja, a operação *merge*, entre o verbo e seu argumento interno “um livro”.

Posteriormente, “*comprou um livro*” sofre o segundo *merge*, agora com o argumento externo, “*a menina*”.

Essa ordem nos é importante, uma vez que contraporemos em nossa investigação as diferenças de processamento entre estruturas de sujeito e de objeto.

Na representação acima, V' , (lê-se v-barra) é a projeção intermediária de V com seu argumento interno, “um livro”. SV é a projeção máxima de V , indicando a relação entre V e seu especificador, o argumento externo. Na sequência, apresentamos a representação arbórea de uma estrutura com um termo adjungido.

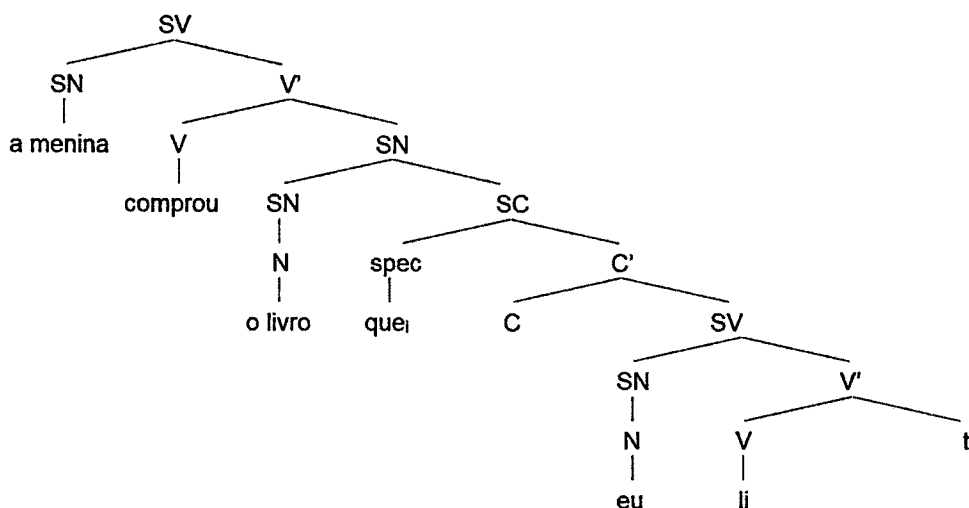
13.



Em 13, temos um adjunto preposicionado, um SP. Como se vê, a adjunção acontece pela combinação de dois sintagmas (SV/SP), ou nas palavras de Kenedy (2013, p.199), “isso significa que o sintagma que recebe a adjunção e, dessa forma, projeta-se novamente na estrutura sintática é o sintagma verbal. Por sua vez, o SP é o sintagma que se adjunge ao SV.” Ou seja, as relações de adjunção se dão entre as projeções máximas dos sintagmas.

Por fim, temos os sintagmas complementizadores, SC, uma categoria funcional que “estabelece a relação entre orações numa dada sentença” (Kenedy, 2013, p.200). As cláusulas relativas são estruturas adjungidas a outras cláusulas, e a relação que se dá entre elas está abaixo representada. A representação arbórea é simplificada, omitindo categorias funcionais, mas atende aos objetivos desta tese.

14. A menina comprou o livro que eu li.



Como se pode perceber, o complementizador, em sua projeção máxima, estabelece uma relação de adjunção com a projeção máxima do sintagma nominal. Mais que isso, o pronome relativo, quando introduz uma relativa, se move para a periferia esquerda da oração relativa, por meio da operação denominada *Move*. Mais adiante descreveremos com mais rigor essa operação, por ora, basta-nos afirmar a existência da operação que é constitutiva de uma relativa, e que esse movimento, denominado movimento A' (A-Barra) é fundamental à gramaticalidade da sentença.

O sintagma verbal, por sua vez, encontra-se completo com seus argumentos, interno e externo, devidamente saturados. A derivação de frases se dá por meio da aplicação de algumas operações como *Merge* e *Move*, a que nos referimos há pouco.

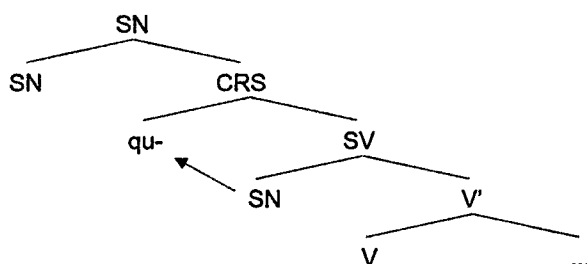
Chomsky (1995) defende que a derivação ocorre por meio de fases. Assim, a memória de trabalho de um falante, mesmo por sua limitação, opera com fases aplicando tais operações e enviando para as interfaces, a que denominamos Forma Fonética (FF) e Forma Lógica (FL) o resultado dessas fases.

Ou seja, "a menina comprou um livro" tem seu ciclo fechado quando da concatenação de seus argumentos, como "que eu li" terá seu ciclo completo quando da saturação dos sintagmas que constituem a cláusulas como ainda do movimento A' necessário à formação de uma relativa.

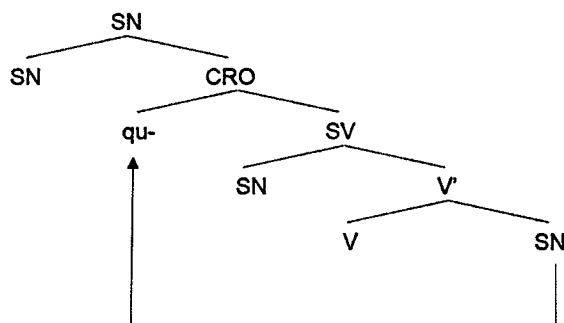
O sujeito de uma sentença, conforme exposto até aqui, possui uma relação com o verbo distinta da relação que há entre o verbo e o objeto. Em um verbo transitivo direto, o objeto é o complemento verbal, o argumento interno, item necessário à saturação do verbo quanto à sua grade argumental. O sujeito também é necessário a essa mesma saturação, mas como argumento externo, estabelecendo com o verbo uma relação menos direta.

Diante disso, podemos afirmar que, sintaticamente, sujeito e objeto são elementos distintos entre si. Para além dessa consideração, a psicolinguística vem demonstrando, nos últimos anos, que as cláusulas relativas de sujeito são diferentes das relativas de objeto, e para esse fato, atribuem-se explicações, sobretudo, de ordem sintática. Posto assim, sentenças como 15 e 16 são processadas de formas distintas:

15. CRS: A menina que_i [] conhece o escritor comprou o livro.



16. CRO: A menina que_j o escritor conhece [] comprou o livro.



Trata-se de duas representações bastante simplificadas, mas que atendem aos objetivos desta tese porque demonstram que a cadeia formada no Move de um CRS é local, mas de um CRO não é. Algo que será discutido mais adiante.

Em 15, “a menina” é o termo relativizado, na cláusula relativa esse termo assume a posição de sujeito por meio da correferência com o pronome relativo – a menina conhece o escritor. Em 16, novamente sendo o termo relativizado,

esse termo funciona como objeto – o escritor conhece **a menina**. Como se vê nas representações arbóreas, tais diferenças têm implicações diretas no movimento A' a que nos referimos há pouco. Os índices, “i” e “j” representam o movimento realizado pelo constituinte.

Segundo Braga, Kato e Mioto (2009, p. 242), “o processo de relativização se realiza por meio de um pronome-Q”. Ainda segundo esses autores, as construções acontecem a partir de um movimento (operação *Move*), em que o pronome relativo (pronome-Q), é movido para a periferia esquerda da oração, uma vez que, se permanecessem onde foram gerados, as sentenças seriam agramaticais.

O pronome movido, por sua vez, deixa, no lugar em que foi gerado, uma categoria vazia, conforme demonstrado acima, movendo-se para a posição de especificador do complementizador. Na versão minimalista da gramática gerativa, segundo Sedrins (2013, p. 202), “a operação *Mover* é na realidade produto da atuação de três mecanismos: Copiar, Concatenar, Apagar”.

Ainda segundo esse autor, “de uma maneira bem simplificada, podemos entender movimento como a propriedade, verificada nas línguas naturais, de um constituinte ser interpretado em uma posição distinta daquela em que é foneticamente realizado”.

Como se pode notar, a formação de uma relativa de sujeito ou de objeto se dá de maneira semelhante, mas o lugar em que o termo relativizado é gerado varia, quando tratamos de uma CRS ou de uma CRO. Gibson (cf. 2000) afirma que a complexidade de integração estrutural depende da distância ou localidade entre dois elementos a serem integrados. Ainda segundo esse autor, um fenômeno de complexidade bem estabelecido a ser explicado por uma teoria de recurso computacional é a complexidade superior de uma CRO quando comparada a uma CRS. Em outras palavras, entre o termo relativizado e o seu lugar de aposição, há mais material linguístico numa CRO do que numa CRS, o que facilita a leitura das relativas de sujeito. Porém, a DLT (GIBSON, 2000) considerará outros fatores quando da análise das diferenças de processamento que há entre essas estruturas. Ao que mais adiante discutiremos.

Nesta seção, portanto, descrevemos a estrutura de uma cláusula relativa, uma estrutura adjungida, encaixada a outro constituinte sentencial por

meio de um pronome relativo, que se move para a periferia esquerda do sintagma complementizador. Sendo um adjunto, a relação entre a estrutura relativa, ou do sintagma complementizador, se dá na projeção máxima de SN.

2.2 Estudos sobre assimetria entre CRS e CRO

Conforme apontamos na seção anterior, CRS e CRO são estruturas diferentes entre si. Do ponto de vista sintático, as relativas de sujeito possuem um argumento externo relativizado, ao passo que as relativas de objeto relativizam um argumento interno.

Alguns estudos (cf. FORD, 1983; KING & JUST, 1991; KING & KUTAS, 1995; FRAUENFELDER, SEGUI & MEHLER, 1980; HOLMES & O'REGAN, 1981; MECKLINGER, et al., 1995; SCHRIEFERS, FRIENDERICI & KÜHN, 1995; FRAZIER, 1987; OLIVEIRA, 2013; COSTA, et al., 2009; MANGAS, 2011, MAK, VONK & SCHRIEFERS, 2002& 2006) no âmbito da psicolinguística experimental vêm apontando uma diferença de processamento das CRS em relação às CRO. Ou seja, os usuários das línguas naturais processam essas estruturas de maneira diversa, o que acarreta um maior ou menor tempo para a leitura de sentenças que contêm CRS ou CRO. Para Costa et al. (2009, p. 212), trabalhos diversos demonstram “haver uma assimetria entre tipos de relativas na aquisição destas estruturas, o que indicará que eventuais déficits relacionados com as relativas não se prendem globalmente com o tipo de estrutura, mas apenas com alguns sub-aspectos”. Posto desta maneira, debruçamo-nos sobre a assimetria que há no processamento de CRS e CRO com o fim de compreender melhor como o *parser* as processa nos primeiros estágios da compreensão e sobre a influência do traço da animacidade nas referidas estruturas.

Essa assimetria foi verificada no Inglês nos trabalhos de Ford (1983), de King & Just (1991), e de King & Kutas (1995), em que a cláusula perde a ambiguidade na palavra que segue o pronome relativo. Cabe ser posto que a língua Inglesa possui uma estrutura semelhante à do Português, no que se refere às relativas, que também desambigua a estrutura relativa logo após o pronome, sendo assim, quando temos uma relativa de sujeito, a palavra que se

segue ao pronome relativo é um verbo; se for uma relativa de objeto, teremos um SN.

A mesma assimetria foi encontrada no Francês, que não possui ambiguidade no relativo, já que ao pronome relativo está marcado como o sujeito (*qui*) ou como o objeto (*que*) da cláusula relativa (FRAUENFELDER, SEGUI & MEHLER, 1980; HOLMES & O'REGAN, 1981). Resultados correspondentes também foram obtidos para o Alemão (MECKLINGER et al., 1995; SCHRIEFERS, FRIEDERICI, & KÜHN, 1995).

Usando várias medidas diferentes, esses estudos mostram que cláusulas relativas de objeto estão associadas a mais dificuldade de processamento que as cláusulas relativas de sujeito.

Como já foi posto, em Inglês, a cláusula perde a ambiguidade pela palavra seguinte ao pronome relativo (um verbo ou um SN). Cláusulas relativas em Holandês e Alemão, contudo, são finalizadas em verbo. Portanto, em Holandês e Alemão as cláusulas relativas não perdem a ambiguidade pela categoria sintática da palavra que segue o pronome relativo. Em Alemão, em muitos casos a cláusula relativa perde a ambiguidade pela marca de caso do pronome relativo ou do SN que segue o pronome relativo. Em Holandês, não há nenhuma marca de caso nos pronomes relativos e SNs completos, e, portanto, as cláusulas relativas perdem a ambiguidade apenas no auxiliar. A marca de número no auxiliar tira a ambiguidade da sentença com relação a uma cláusula relativa de sujeito. Posto de outra forma, o Alemão e o Holandês possuem diferenças estruturais nas sentenças relativas quando comparadas ao Português e ao Inglês que precisam ser consideradas neste trabalho.

Em Holandês e Alemão foi investigado se a designação da função do sujeito é influenciada pelo ajuste temático entre os termos na cláusula relativa e o verbo principal (MECKLINGER et al., 1995; SCHRIEFERS et al., 1995). Nos experimentos, foram usados participípios passados que foram enviesados para tomar um dos termos na cláusula relativa como o sujeito e o outro como o objeto. Foram apresentadas sentenças, em que um dos protagonistas seria mais provável de ser o sujeito do verbo principal que o outro. Os resultados mostraram que o ajuste temático não foi suficiente para que o tempo de processamento do sujeito fosse menor do que do objeto. Diante disso, os

autores deduziram que o *parser* não tem acesso a esses traços semânticos no primeiro momento do processamento.

Trueswell, Tanenhaus, e Garnsey (1994), por sua vez, usaram a animacidade para manipular o ajuste temático de um SN inicial de sentença com o verbo que o seguia em Língua Inglesa. Esta manipulação do ajuste temático afetou os tempos de leitura para cláusulas relativas reduzidas e completas. Geralmente, cláusulas relativas reduzidas como em 17, abaixo, levam a tempos de leitura mais longos no *by-phrase*, que cláusulas relativas completas, pois os leitores preferem uma leitura de verbo principal do particípio passado que tem que ser reanalisado na leitura do *by-phrase*. O *by-phrase* é “o elemento de caráter agentivo em passivas” (JUNIOR e CORRÊA, 2015, p. 93). Os autores mostraram que a diferença entre cláusulas relativas reduzidas e cláusulas relativas completas desapareceu quando o substantivo antes do verbo foi inanimado, como em 18.

17. O réu examinado pelo advogado³...

18. A evidência examinada pelo advogado⁴...

Esse trabalho é relevante à nossa discussão haja vista termos a manipulação do traço da animacidade como fator que modifica o tempo de leitura de cláusulas relativas, do que se pode inferir que o acesso a esse traço pode acontecer já nos estágios iniciais do processamento, ao que mais adiante nos deteremos mais detalhadamente. Sobre essas estruturas, Traxler (2012) faz considerações sobre o traço da animacidade e de como a frequência atua para um processamento mais ou menos custoso dessas estruturas. Sobre isso, discorreremos mais detidamente na seção 2.4.4.

Em Português, também alguns trabalhos já demonstraram a assimetria das CRS em relação às CRO, como na maioria dos casos, também trazendo as relativas de sujeito como mais rápidas de serem lidas do que as relativas de objeto. Oliveira (2013) realizou três experimentos com a Técnica Experimental de Leitura Automonitorada com estruturas com CRS e CRO. Seu foco eram as

³The defendant examined by the lawyer was...

⁴The evidence examined...

diferenças entre as relativas com encaixe central e à direita. Em seu trabalho, o autor comprovou uma assimetria entre tais estruturas, sendo as CRS significativamente mais rápidas de serem lidas do que as CRO.

Oliveira (2013) tinha seu foco no encaixe das cláusulas relativas, e no tocante às diferenças no tempo de processamento o autor explicou a partir das diferenças sintáticas que há entre elas, como a distância e a complexidade dos termos relativizados. Segundo ele, as CRS são mais facilmente processadas porque “esse tipo de extração descreve uma cadeia linearmente e estruturalmente menor” (OLIVEIRA, 2013, p.77). A seguir, transcrevemos 4 estruturas retiradas do trabalho de Oliveira (2013, p.1).

19. O banqueiro que encontrou o repórter atacou o senador.
20. O banqueiro que o repórter encontrou atacou o senador.
21. O senador atacou o repórter que encontrou o banqueiro.
22. O senador atacou o repórter que o banqueiro encontrou.

Conforme se pode observar, todos os substantivos relativizados (banqueiro, repórter, senador) são animados, notadamente humanos.

Assim como Oliveira (2013), Mangas (2011) também analisou estruturas relativas e encontrou uma assimetria entre elas. A autora desenvolveu seu experimento com 6 crianças com deficiência auditiva e 6 crianças sem deficiência. A partir do reconhecimento de imagens foi elaborado um teste com sentenças contendo CRS e CRO, sendo elaborado ainda um teste de preferência. No teste de identificação de imagens, as crianças eram expostas a duas imagens, como as que se seguem, e, na sequência, o experimentador questionava a criança:

a- Experimentador: mostra-me o soldado que o médico está a pintar.
(MANGAS, 2011, p.61)

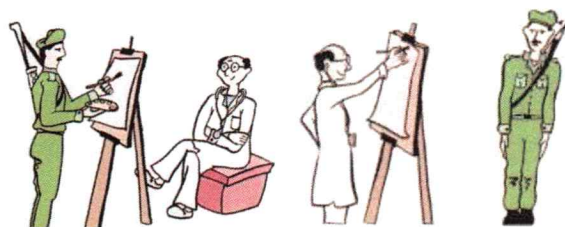


Figura 1: Figuras retiradas do trabalho de Mangas (2011, p. 60).

Em seguida, eram postas outras duas figuras, como as que se seguem, então a criança era questionada.

b- Experimentador: mostra-me o menino que está a pentear o rei. (MANGAS, 2011, p.62).

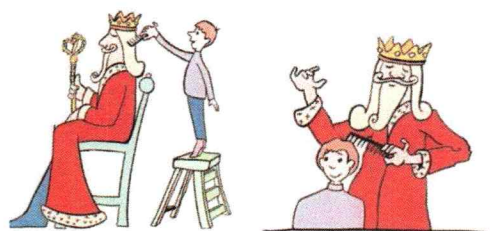


Figura 2: Figuras retiradas do trabalho de Mangas (2011, p. 60).

Cabe salientar que todas as figuras eram animadas (soldado, médico, rei e menino), o mesmo se repetiu com todas as figuras do teste.

Nas tarefas de preferência, as crianças tinham que responder a questões como as que se seguem:

c- Experimentador: há dois meninos a passear na rua: um menino encontra o professor e o outro menino encontra um amigo. Que menino gostavas mais de ser?

d- Experimentador: há dois meninos a passear no jardim. O avô abraça um menino e a mãe abraça o outro menino. Que menino é que gostavas mais de ser?

Neste segundo teste, havia estruturas com termos inanimados, mas segundo a própria autora, não houve um equilíbrio rigoroso do traço da animacidade, das vinte estruturas testadas, 6 eram inanimadas, como a que se segue:

23. Gostava mais de ser o menino que o rádio desperta.

O foco da autora eram as diferenças de processamento dessas sentenças entre crianças com deficiência auditiva e crianças sem deficiência auditiva. Em ambos os testes, tanto nas CRS quando nas CRO, as crianças

com deficiência auditiva foram mais lentas nas tarefas, e em ambos os grupos de sujeitos, as CRS foram lidas significativamente mais rápido que as CRO. Ou nas palavras da autora: “Todas as crianças apresentaram melhores resultados na identificação das relativas de sujeito do que na identificação das relativas de objecto” (MANGAS, 2011, p.80).

Posto desta forma, Mangas (2011) também encontra uma assimetria entre CRS e CRO, sendo as relativas de sujeito mais fáceis de processar do que as relativas de objeto.

O trabalho de Mangas, por sua vez, tem suas bases no trabalho de Costa et al., (2009), que também testaram relativas de sujeito e de objeto com foco na assimetria de processamento dessas estruturas. Já no início do trabalho dos referidos pesquisadores, eles fazem algumas considerações acerca de fatores estruturais que distinguem as CRS das CRO.

as relativas de sujeito e objecto são estruturas complexas que envolvem dependências A-barra (com movimento), distinguindo-se pelo facto de apenas as de objecto envolverem uma dependência em que há intervenção de um argumento/papel temático, isto é, em que o papel temático do objecto atravessa o papel temático do sujeito (COSTA et al., 2009, p. 212).

Para ilustrar o que afirmam acima, os autores trazem os exemplos que se seguem:

- a. **Gostava de ser o rapaz que visitava o avô.**
- b. **Gostava de ser o rapaz que o avô visitou.**

O trabalho de Oliveira (2013), mencionado anteriormente, explica a assimetria no processamento dessas estruturas com base em argumentos da mesma ordem que Costa et al. (2009), ou seja, com base em fatos de natureza sintática.

Costa et al. (2009) realizaram dois experimentos, sendo um com produção de relativas e outro com compreensão, a partir do reconhecimento de imagens. Os testes foram feitos com crianças sem déficit de linguagem, crianças com déficit e um grupo controle, formado por adultos. Os

pesquisadores encontraram a mesma assimetria relatada na maior parte da literatura que trata do tema. As CRS foram significativamente mais fáceis de processar, tanto em tarefas de produção quando de compreensão, do que as CRO. Costa et al. usaram imagens de pessoas e animais, todos animados, nas tarefas de reconhecimento de imagem.

Por ora, o que está posto é que, nos três trabalhos supracitados que analisaram as relativas na língua Portuguesa (OLIVEIRA, 2013; MANGAS, 2011 E COSTA et al., 2009), as CRS e CRO são assimétricas entre si, necessariamente sendo as CRS mais fáceis de processar do que as CRO.

2.3 A influência do traço da animacidade no processamento de sentenças relativas

Alguns trabalhos vêm demonstrando que fatores de ordem semântica também podem interferir no processamento de CRS e CRO. Conforme apontamos na seção anterior, Trueswell, Tanenhaus, e Garnsey (1994), Mecklinger et al., (1995) e Schriefers et al., (1995), manipularam traços semânticos, incluindo-se a animacidade, e a partir de diferentes medidas apontaram para influências mais ou menos significativas desses traços no processamento de relativas.

Dois trabalhos, contudo, trataram mais diretamente do traço da animacidade em CRS e CRO, e por isso discutiremos seus resultados nesta seção. Ambos os trabalhos são de Mak, Vonk e Schriefers (2002, 2006).

No trabalho de 2002, os autores partem do argumento de que, embora a literatura que aborda a assimetria entre CRS e CRO venha demonstrando por medidas diversas que as relativas de sujeito são mais rápidas, portanto, mais fáceis de processar do que as de objeto, esses trabalhos desconsideram estruturas com termos inanimados, ou pelo menos não controlam esse traço nos experimentos.

A partir dessa consideração, os autores citam o trabalho de Zubin (1979), em que o autor estudou o efeito da animacidade na distribuição de cláusulas relativas de sujeito e de objeto. Para Zubin, regularidades podem ser conduzidas por saliência e por animacidade.

Ou seja, Segundo Zubin (1979), um item linguístico que é mais saliente tende a ser sujeito da sentença. No argumento do autor, a disponibilidade do item é fator importante para se estabelecer a saliência.

Logo, quando tratamos de uma relativa, temos um termo antecedente, retomado pelo pronome relativo, e um termo dentro da subordinada encaixada. Visto assim, o termo mais disponível é o primeiro a aparecer, ou seja, o termo antecedente, que vem a ser dessa forma o item mais saliente, portanto, primeiro candidato a ser o sujeito da sentença.

Zubin (1979) estabelece ainda que um item animado será preferido para ser sujeito da sentença. Ou seja, quando temos um termo antecedente animado, há duas forças que atuam sobre ele para que ele seja o sujeito da sentença. As evidências para essas considerações Zubin encontrou num estudo de *corpus* com cláusulas relativas do Alemão com itens animados e inanimados.

A partir do estudo feito por Zubin e das considerações no tocante à realização das relativas nas línguas humanas, Mak, Vonk e Schriefers (2002) realizaram dois estudos de *corpus*, um em Alemão e outro em Holandês. Para o estudo de *corpus* do Holandês foram usadas duas edições do jornal diário *Trouw* (16 e 17 de setembro de 1993).

Estes textos contêm 120.000 palavras. Para o estudo de *corpus* Alemão foram usadas quatro edições da versão digital do jornal diário *Die Welt* (15, 19 e 20 de Novembro de 1995 e 2 de Dezembro de 1995). Estes textos contêm 75.000 palavras.

Só interessavam as estruturas com sujeito e objeto, e foram encontradas nessas condições 286 relativas em Holandês e 168 no Alemão. A distribuição das relativas se deu conforme a Tabela abaixo. Cabe ser posto, porém, que não discutiremos a representatividade desses *corpora*. Nossas discussões se darão, sobretudo, a partir de dados experimentais, informações oriundas de *corpus* tão somente nos instrumentalizam frente às nossas análises.

Tabela 1: *Corpus* do Holandês e do Alemão (MAK, VONK E SCHRIEFERS, 2002)

Relativa	Holandês	Alemão	Total
Sujeito Animado	140(49%)	82(49%)	222(49%)
Sujeito Inanimado	66(23%)	62(37%)	128(28%)
Objeto Animado	2 (1%)	2 (1%)	4(1%)
Objeto Inanimado	78 (27%)	22(13%)	100(22%)
Relativas em números absolutos dos <i>Corpora</i> – Porcentagem entre parêntesis			

Como se pode perceber, as construções mais comuns nas duas línguas foram as de sujeito animado; contudo, as construções de sujeito inanimado e objeto animado também foram muito frequentes. As estruturas de objeto animado, por outro lado, foram muito raras em ambas as línguas. Cabe ressaltar que, conforme temos apontado neste capítulo, a maior parte dos estudos que constitui a literatura atual que aborda a assimetria entre CRS e CRO parte da análise de CRS e CRO animadas. Ou seja, algo que parece não representar o que de fato ocorre nas línguas naturais, ao menos se consideramos os excertos das duas línguas estudadas como representativas de um universo maior.

Outro parêntese, brevemente apontado anteriormente, tem que ser aberto nesta nossa análise. O Holandês e o Alemão possuem especificidades no que se refere às relativas que inviabilizam uma comparação direta com o Português. Diferentemente do Português e do Inglês que, desambiguizam a relativa pela categoria gramatical da palavra que se segue ao pronome relativo, em Alemão essa desambiguação ocorre por uma marca de caso que acontece na maioria das relativas. O Holandês, por sua vez, não possui essa marca de caso, mas possui um auxiliar que tira essa ambiguidade. Diante disso, o trabalho de Mak, Vonk e Schriefers (2002) traz discussões em torno dessa marca de caso do Alemão e do auxiliar do Holandês que não serão feitas em nosso trabalho.

Os autores analisam seus dados e consideram que o que encontram é compatível com os pressupostos de Zubin (1979), uma vez que há mais relativas de sujeito, e que as relativas de objeto são, em sua maioria, inanimadas. Posto isso, eles afirmam que a animacidade influencia a distribuição de relativas no Alemão e no Holandês.

Diante das considerações inferidas do estudo de *corpus*, Mak, Vonk e Schriefers (2002) fizeram dois estudos experimentais em Holandês. O primeiro com a Técnica Experimental de Leitura Automonitorada e o segundo a partir da Técnica de Rastreamento Ocular.

Os experimentos desse estudo de Mak, Vonk e Schriefers (2002) foram fundamentais para a estruturação dos experimentos desta tese, mas cabe ser posto que nosso trabalho não é uma réplica do trabalho desses autores. Nossas opções metodológicas foram distintas justamente porque as diferenças entre o Português e o Holandês não nos permitiriam uma comparação direta entre essas línguas. Por exemplo, o tempo de processamento do auxiliar desambiguador foi fundamental à análise do Holandês. O Português, por sua vez, não possui esse auxiliar. Assim como no Inglês, a ordem das palavras de uma CRS e de uma CRO é diferente em Português.

Mais que isso, houve uma preocupação no trabalho de Mak, Vonk e Schriefers (2002), quanto à relação entre a animacidade do termo antecedente e do termo da cláusula encaixada. No nosso trabalho, porém, concentramos nossas análises no termo antecedente, tanto em nosso estudo de *corpus* quanto em nossos experimentos.

Na sequência seguem as condições experimentais do primeiro experimento reportado por Mak, Vonk e Schriefers (2002).

Tabela 2: Condições Experimentais do 1º Experimento de Mak, Vonk E Schriefers (2002)

(A) Cláusula relativa de sujeito com objeto animado.

Vanwege het onderzoek moeten de inbrekers, die de bewoner beroofd hebben, nog een tijdje op het politiebureau blijven.

Devido às investigações, os ladrões, que roubaram o inquilino, tiveram que ficar na delegacia por algum tempo.

(B) Cláusula relativa de objeto com objeto animado.

Vanwege het onderzoek moeten de inbrekers, die de bewoner beroofd hebben, nog een tijdje op het politiebureau blijven.

Devido à investigação, o inquilino, que os assaltantes roubaram, teve que ficar na delegacia por algum tempo.

(C) Cláusula relativa de sujeito com objeto inanimado.

Vanwege het onderzoek moeten de inbrekers, die de computer gestolen hebben, nog een tijdje op het politiebureau blijven.

Devido à investigação, os assaltantes, que roubaram o computador, tiveram que ficar na delegacia por algum tempo.

(D) Cláusula relativa de objeto com objeto inanimado.

Vanwege het onderzoek moet de computer, die de inbrekers gestolen hebben, nog een tijdje op het politiebureau blijven.

Devido à investigação, o computador, que os assaltantes roubaram, teve que permanecer na delegacia por algum tempo.

Foi analisado o tempo de processamento da palavra imediatamente após do auxiliar desambiguador. Nas estruturas com objeto animado, os tempos de leitura do segmento foram mais rápidos para as estruturas de sujeito do que as de objeto. Contudo, quando o objeto era inanimado, essa diferença desapareceu. Segundo os autores, por haver, como descreve Zubin (1979), duas forças que impelem o substantivo animado a ser o sujeito da sentença, quando o termo é inanimado. Na argumentação deles, o contrário acaba acontecendo também. Ou seja, quando temos um substantivo inanimado, o *parser* já é guiado no sentido de construir uma estrutura de objeto. Essa discussão será ampliada frente aos dados que reportaremos na análise dos experimentos 3 e 4 constitutivos desta tese, no capítulo 4. Queremos pontuar, porém, que essas forças de que trata Zubin (1979), referidas por Mak, Vonk e Schriefers (2002), serão discutidas sob o olhar da DLT, que considera fatores de diversas ordens nas análises ora reportadas na próxima seção deste capítulo, bem como na análise de nossos dados nos capítulos 3 e 4 desta tese.

A Técnica de Leitura Automonitorada, porém, na opinião dos autores do estudo, poderia não captar se a desambiguação ocorreu no auxiliar ou se houve um efeito de encapsulamento tardio. Os autores quiseram aumentar a

visão da natureza dos efeitos e para tanto fizeram outro experimento, a partir de condições semelhantes, porém, com a Técnica Experimental de Rastreamento Ocular. Para muitos pesquisadores, o rastreamento ocular pode captar com maior precisão alguns fatos linguísticos. Não discordamos dessa perspectiva, inclusive admitimos a inclusão dessa técnica para futuros trabalhos que podem expandir os resultados desta tese. Contudo, conforme exporemos mais adiante, os resultados que encontramos foram suficientemente robustos com a Técnica de Leitura Automonitorada, suprimindo-nos de dados significativos para as inferências que pretendíamos empreitar.

Para além desse nosso posicionamento, Mitchell (2004) argumenta que, muito embora existam críticas em torno da Técnica de Leitura Automonitorada, a maior parte delas relacionadas ao fato de a fragmentação da frase diminuir o tempo de leitura, ou tornar a leitura menos natural, muitos estudos foram feitos no sentido de refutar os resultados experimentais da Leitura Automonitorada, mas os resultados a partir de outras técnicas, como o EEG e o Rastreamento Ocular acabam por, sistematicamente, confirmar dados encontrados na Leitura Automonitorada. Tais fatos fazem com que Mitchell (2004) considere essa técnica como uma ferramenta de grande valia à psicolinguística, pois atrelada à confiabilidade de seus resultados, ainda pesa a questão de baixo custo, quando relacionada a outras técnicas experimentais. Concordamos com o autor e, embora, consideremos lícito que nossos resultados sejam ampliados por meio, por exemplo, da replicação por meio da Técnica de Rastreamento Ocular, estamos seguros de que os resultados aqui trazidos possuem a precisão necessária à nossa discussão.

Os resultados do segundo experimento com rastreamento ocular feitos por Mak, Vonk e Schriefers (2002) confirmaram o que foi encontrado no primeiro.

As duas medidas analisadas foram o tempo da primeira passagem⁵ e o tempo de regressão⁶, assim como na leitura automonitorada, a diferença foi encontrada apenas nas relativas com termos animados, em que o processamento do sujeito é mais rápido. Essa diferença desaparece quando o

⁵First-pass, é a primeira passagem do olhar sobre um determinado ponto aferido no processamento.

⁶O tempo de regressão é a contagem do tempo que o olhar retoma um lugar específico dentro da estrutura experimental em questão.

objeto relativizado é inanimado. Segundo os autores, a diminuição do tempo do objeto inanimado, não computada em experimentos que não o incluem, faz com que o tempo de processamento dos objetos, animados e inanimados, seja reduzido e isso faz com que não haja diferença significativa entre o objeto e o sujeito.

No experimento de Leitura Automonitorada, uma interação entre a animacidade de objeto e tipo de cláusula foi encontrada na palavra que segue o auxiliar. A questão era se esta interação foi um efeito de *spillover* do auxiliar, ou se estava presente apenas no encapsulamento da cláusula. Os dados do experimento de Rastreamento Ocular mostram que o efeito não é atrasado, pois a interação foi encontrada no auxiliar por si só, nos tempos de leitura de primeira passagem. Uma vez que o auxiliar não foi o final da cláusula neste experimento, o efeito não pode ocorrer devido ao encapsulamento da cláusula.

Além da interação da animacidade de objeto e tipo de cláusula nos tempos de leitura de primeira passagem no auxiliar, houve também uma interação destes fatores na duração do tempo de regressão no substantivo final da cláusula. Tempos de leitura nas cláusulas relativas de objeto estavam atrasados apenas quando o objeto da cláusula relativa era animado. A interação nesta posição pode ser compreendida como um efeito de encapsulamento da cláusula. Portanto, os dados apresentaram efeitos imediatos e de encapsulamento.

Posto dessa forma, os dois experimentos confirmam a premissa de seus autores de que o traço da animacidade é acessado pelo *parser* já nos primeiros momentos de processamento da frase em estruturas relativas. Mais que isso, independente da localização precisa do efeito da animacidade, esse traço afeta a análise de cláusulas relativas.

Em 2006, Mak, Vonk e Schriefers retomam a discussão em outro artigo. Nesse novo trabalho os autores procuraram explorar os fatores que podem contribuir para a modulação da preferência usual por cláusulas relativas de sujeito. Para tanto, realizaram 3 experimentos.

Os autores partem do pressuposto de que, ao se manipular o conteúdo semântico de verbos em frases mais ou menos plausíveis na condição de sujeito e de objeto, o tempo de leitura das frases menos plausíveis se mantém.

Numa tradução livre, as frases, testadas em Holandês, equivaliam aos exemplos que se seguem:

24. Amanhã o professor que ensinou os alunos entregará os diplomas⁷.

25. Amanhã os estudantes que o professor ensinou receberão os diplomas⁸.

O verbo “ensinar” torna o SN “*estudantes*” altamente implausível como sujeito da sentença relativa, ainda assim, alguns estudos experimentais citados por Mak, Vonk e Schriefers (2006) (MAK et al., 2002; MECKLIGER et al., 1995; SHRIEFERS et al., 1995) demonstram que essa informação não interfere no tempo de processamento das relativas, e estruturas como 24 continuam sendo mais rápidas de serem lidas do que 25. Ainda segundo esses autores, a maioria desses estudos usou cláusulas relativas em que tanto o antecedente como o SN interno à cláusula relativa eram termos animados.

Diante disso, foram realizados três novos experimentos. No experimento 1, foram testadas estruturas em que o SN antecedente e o SN interno eram inanimados. A intenção era investigar se a animacidade *per se* influencia a decisão do *parser*.

Nos experimentos 2 e 3, foram usadas sentenças em que um dos SNs é animado e o outro inanimado, com o intuito de se aferir o efeito da mudança do animacidade desses SNs. O primeiro e o segundo experimento foram a partir da Técnica Experimental de Leitura Automonitorada e o terceiro foi feito a partir de Rastreamento Ocular.

As frases testadas no Holandês no experimento 1 do trabalho de Mak, Vonk e Schriefers (2006), traduzidas livremente, equivaliam às que se seguem.

26. Segundo o panfleto, o gel que remedia os vazamentos deve funcionar imediatamente⁹.

27. Segundo o panfleto, os vazamentos que o gel remedia devem desaparecer imediatamente¹⁰.

⁷ Morgen zal de professor, die de studenten ontmoet heeft, de diploma's uitreiken.

⁸ Morgen zal de professor, die de studenten ontmoet hebben, de diploma's uitreiken.

⁹ Volgens de folder moet de gel, die de lekkages verhelpt, in één keer werken.

¹⁰ Volgens de folder moeten de lekkages, die de gel verhelpt, in één keer verdwenen zijn.

Partindo dos resultados que encontraram em 2002, os autores esperavam encontrar tempos de leitura mais rápidos para estruturas como 27 do que estruturas como 26, contudo os resultados foram contrários a isso, e as estruturas de sujeito foram mais rápidas do que as de objeto. Ao que parece, quando os dois SNs são inanimados, a animacidade guia a interpretação inicial da cláusula relativa, mas isso ocorre na combinação entre o termo antecedente e o SN interno à relativa.

Segundo os autores, além da animacidade, há outro fator que deve ser levado em consideração quando se estuda processamento de cláusulas relativas, já que a cláusula relativa é uma declaração sobre o antecedente, o antecedente é o tópico da sentença. Visto que há uma tendência para o tópico ser o sujeito de uma cláusula, eles estabeleceram a hipótese de que o antecedente será escolhido como sujeito, caso as outras influências na escolha do sujeito da cláusula relativa sejam as mesmas. Isto explicaria a preferência por cláusulas de sujeito em sentenças em que o SN antecedente e o SN interno à cláusula relativa são iguais em animacidade.

Nos experimentos 2 e 3, os autores então investigaram a influência da animacidade e da topicalização. Partiu-se da prerrogativa de que há dois fatores potenciais que podem afetar a atribuição da função sintática do sujeito e do objeto em cláusulas relativas com dois SNs completos: a preferência para o SN antecedente como sujeito da cláusula relativa e a preferência de uma entidade animada como sujeito da sentença. Os autores consideraram SNs completos aqueles constituídos por um substantivo, ou, melhor dizendo, por um sintagma cujo núcleo fosse um substantivo, eles preferiram descartar os SNs constituídos por pronomes. No Experimento 2, os autores investigaram a influência desses dois fatores no processamento de cláusulas relativas. Usaram para tanto cláusulas relativas com uma entidade animada e uma inanimada.

O primeiro fator no experimento foi a animacidade do sujeito – animado ou inanimado – e o segundo o tipo de cláusula – de sujeito ou de objeto. Os autores encontraram efeito de interação entre a animacidade e a topicalização, ou seja, uma evidência de que o conteúdo semântico pode interferir no *parsing* de sentenças relativas. Contudo, tais resultados não encontram correspondentes diretos no Português, pois o tempo mais relevante aferido

pelos autores foi do auxiliar desambiguador, e o Português não possui tal termo. Os autores demonstraram que não há diferença no tempo de processamento de relativas de sujeito e de objeto quando o antecedente é animado. Os resultados do terceiro experimento, feito a partir da Técnica de Rastreamento Ocular, confirmaram o que já tinha sido encontrado na Leitura Automonitorada, no segundo experimento.

Os resultados de Mak, Vonk e Schriefers (2006) confirmam, segundo eles, a hipótese da topicalização, que provê uma explicação para o padrão de resultados encontrados por eles. Segundo essa hipótese, o leitor opta por um dos SNs na cláusula relativa como o sujeito tendo como base a validade do tópico, que subordina, por sua vez, a animacidade e a topicalização do antecedente.

A interação desses dois fatores conduz a um padrão de resultados. A hipótese da topicalização também provê predições adicionais que podem ser testadas empiricamente. A dificuldade de cláusulas relativas de objeto deveria ser afetada pelo SN interno à cláusula relativa sendo o SN completo ou um pronome.

Além disso, mesmo quando o SN interno à cláusula relativa é um SN completo, o *parsing* de cláusulas relativas de objeto deveria ser mais fácil quando o referente do SN completo é tópico do discurso comparado quando este é uma nova entidade.

Em suma, nesta seção, procuramos elencar alguns estudos que abordaram a influência da animacidade no processamento de CRS e CRO. A animacidade do antecedente, *per se*, parece não conduzir a decisão do *parser*, ao menos no que se refere ao Holandês, língua cuja estrutura da relativa se difere em muito do Português, já que ela possui um auxiliar desambiguador que define se a cláusula é de sujeito ou de objeto, além de possuir um verbo no final da sentença e de a ordem das palavras não mudarem quando se trata de uma CRS ou de uma CRO.

A esta tese, interessa ser posto que a assimetria verificada em diversas línguas por meios experimentais distintos é relativizada quando o traço semântico da animacidade é controlado.

Em nosso trabalho, tomaremos a língua Portuguesa, em suas variedades Brasileira e Europeia, para ampliarmos a visão que se tem do

fenômeno com base em experimento envolvendo CRS e CRO. Para tanto, precisamos traçar nossa discussão a partir de um modelo de processamento que abarque o *parsing* dessas sentenças e que explique os resultados por nós encontrados, empreitada que empreendemos na próxima seção.

2.4 Modelos teóricos de processador sintático

A psicolinguística experimental procura explicar fenômenos linguísticos a partir de evidências comportamentais mensuráveis por técnicas e metodologias diversas. Para além de tais técnicas, modelos teóricos constituídos a partir de diferentes perspectivas são criados e a explicação dos fenômenos estudados depende fortemente das evidências experimentais para sua validação e desenvolvimento.

O *parser* é o analisador sintático, e “se refere aos procedimentos mentais, que determinam a estrutura de uma sentença, integrando os mecanismos de produção e compreensão da linguagem” (GRAVINA, 2009). Um dos pontos fulcrais de uma teoria que se propõe a compreender o funcionamento linguístico na mente humana é saber como o *parser* opera, quais são os elementos que são acessados no momento do processamento da linguagem, durante o *parsing* da sentença. Esse processador atribui estrutura às sentenças e fornece representações estruturais, que por sua vez recebem análises interpretativas, quando entram em cena informações de natureza semântica, pragmática e discursiva. Nem todas as teorias que tratam de processamento nomeiam o processador de *Parser*. Nesta tese, contudo, optamos por fazê-lo, mesmo na descrição de tais modelos.

Conforme discutimos na seção 2.1 deste capítulo, o léxico é constituído por vocábulos que funcionam como predicadores, complementos ou adjuntos, na estrutura frasal. O vocábulo na frase é um item sintático, que se for um predicador, projeta sua estrutura argumental quando da realização da sentença. Ou seja, um verbo, por exemplo, ao se realizar na sentença, projeta a quantidade de argumentos que preencherá essa estrutura, esses argumentos são itens sintáticos que vão saturar a seleção do verbo para que a sentença esteja completa.

Todavia, nesse ínterim em que as concatenações acontecem, informações de naturezas distintas entre si são acessadas para que a interpretação da frase aconteça. Um dos objetivos de uma teoria de processamento de frases é descrever e explicar como e quando isso acontece. Abaixo, transcrevemos do trabalho de Kenedy (2015, p.153) uma Tabela que resume os três tipos principais de modelo de processamento, a respeito dos quais discorreremos mais adiante neste texto.

Tabela 3: Principais modelos de processamento de frase (KENEDY, 2015, p. 153)

TGP	Conexionismo	DLT
Serial e modular (Ex. FRAZIER & FODOR, 1978; FRAZIER, 1979; FRAZIER E RAYNER, 1982).	Paralelo e não modular (Ex. MACDONALD et al., 1994; TANENHAUS et al., 1995).	Serial e não modular (Ex. GIBSON, 2000).

2.4.1 Modelos Seriais e modulares –Teoria do *Garden-Path* (TGP).

Alguns modelos de processamento são caracterizados por serem modulares, neste trabalho interessa-nos, especialmente, a teoria do *Garden-Path* (FRAZIER, 1979). Para esses modelos, traços de natureza semântica não são lidos pelo *parser* no primeiro estágio da compreensão da sentença.

Ou seja, o processamento da linguagem humana, numa visão estritamente modularista, primeiro tem acesso aos itens sintáticos sem acessar informações semânticas, que só entrariam em cena posteriormente. “O conceito de modularidade da mente se refere à ideia de que existem subcomponentes relativamente independentes e de função específica no processamento da linguagem” (MAIA, 2015, p.15). O *parser* seria um desses módulos mentais, o módulo sintático. As concatenações que ocorrem entre os itens lexicais são realizadas por esse módulo sintático, algo que acontece muito rapidamente, de modo reflexo, e está condicionado, inclusive, às limitações da memória de trabalho. É justamente o fato de esse módulo se dedicar exclusivamente a essa tarefa faz com que o processamento seja tão rápido.

Kenedy (2015, p. 152) faz uma importante distinção entre o conceito de modularidade formulado por Frazier (1979), inicialmente tratado por Frazier e Janet Fodor (1978) e o conceito de modularidade de Jerry Fodor (1983)

Deve-se ter em conta que a noção de modularidade no acesso às informações usadas na construção de sintagmas e frases foi explicitamente formulada por Frazier (no fim da década de 70) e tal noção é sensivelmente diferente do modelo de modularidade difundido pela argumentação de Jerry Fodor (1983). De uma maneira geral, a noção de modularidade de Fodor diz respeito à existência de sistemas de conhecimento específicos para tratar tipos particulares de informação cognitiva (por exemplo, atenção, memória, linguagem – ou dentro da linguagem, submódulos ou micromódulos como sintaxe, semântica, fonologia etc.). Já a modularidade sustentada inicialmente por Frazier e Janet Fodor (1978), e depois de uma maneira generalizada nos estudos da teoria do *Garden-Path* (TGP), diz respeito à hierarquia de acesso à informação linguística durante o funcionamento do *parser*, no desempenho linguístico. (KENEDY, 2015, p.152)

As teorias consideradas estruturais, para além da modularidade, assumindo a modularidade nos termos de Frazier (1979), consideram ainda que o *parsing* da sentença seja serial, incrementacional e encapsulado.

No que se refere à serialidade, o *parsing* de uma sentença acontece “em uma série linear de operações mentais, em um processo em que a saída da operação anterior fornece a entrada da próxima operação” (MAIA, 2015, p.16). Além disso, o termo serial refere-se ao fato que, diante de uma ambiguidade, o *parser* computa imediatamente uma, e apenas uma análise sintática. Diferentemente do que sustentam os modelos de processamento paralelos, que admitem múltiplas análises simultâneas, diante das quais, depois da desambiguação, opta por uma delas.

A noção de incrementacionalidade diz respeito ao fato de que “a estrutura vai sendo construída à medida que cada item vai sendo acessado, sem atraso, sem esperar o fim da frase” (MAIA, 2015, p.16). Por fim, e de fundamental importância para nossa discussão, sendo encapsulado, o *parser* tem acesso apenas a informações de natureza sintática nesse momento inicial, visto assim, para os modelos ditos estruturais, as informações de natureza semântica, por exemplo, não seriam acessadas imediatamente, e o *parser* não se comprometeria com elas no momento inicial do processamento. Contudo,

evidências experimentais *on-line* vêm mostrando que alguns traços semânticos são de fato acessados, já no primeiro momento da leitura.

O *parser* incremental, ou seja, a ideia de que o *parser* constrói a representação sintática à medida que as palavras vão sendo encontradas, as quais são imediatamente incorporadas à estrutura, condiz com a limitação de memória de trabalho do falante/ouvinte. Mais que isso, faz com que o tempo de processamento seja otimizado.

Entre os modelos modulares temos a teoria do *Garden-Path* (FRAZIER, 1979), comumente traduzida para o Português como Teoria do Labirinto, que propõe um processamento modular, serial, incremental, encapsulado. Ou seja, considera-se apenas uma análise estrutural diante de ambiguidades sintáticas. Diante da ambiguidade, o *parser* se compromete apenas com uma estrutura, uma interpretação particular que atende a restrições sintáticas impostas por princípios estruturais, como a Aposição Local – *Late Closure* – e a Aposição Mínima – *Minimal Attachment*– que pressupõem que a análise escolhida será aquela que estiver primeiro disponível para o *parser*, que por sua vez implicam menor complexidade estrutural.

Basicamente, por ter se comprometido com uma, e apenas uma, estrutura possível no *parsing* de uma frase ambígua, o *parser* pode ter que reanalisar a estrutura quando a primeira opção feita por ele não foi a correta, nesse momento, acontece o chamado efeito *Garden-Path*, o *parser* cai no labirinto e isso demanda custo operacional.

Teorias que concebem o processamento como paralelo, ou seja, teorias que predizem que o *parser* constrói diversas estruturas sintáticas simultâneas, processando-as paralelamente, postulam que quando da desambiguação frase, ele escolherá a melhor opção entre as várias processadas.

A priori, quando da concepção desta teoria (TGP), pretendia-se postular uma universalidade dos seus princípios, não se restringindo apenas ao Inglês, língua de onde se originaram as postulações iniciais da teoria, mas para todas as línguas humanas, no que se refere a estratégias cognitivas. O princípio da aposição local foi formulado da seguinte maneira: “Quando possível, aponha os itens lexicais que vão sendo encontrados à oração ou sintagma correntemente sendo processado” (FRAZIER, 1979, apud MAIA, 2015, p.14). Já o princípio da aposição mínima está assim formulado: “Aponha um item à estrutura

sentencial, usando o mínimo possível de nós sintáticos” (FRAZIER, 1979, apud MAIA, 2015, p.14)

No tocante à Aposição Mínima, essa universalidade vem sendo comprovada nos testes já realizados em diversas línguas por métodos distintos, contudo, o mesmo não se observa quanto ao Princípio da Aposição Local. Frases ambíguas como 28, testadas em algumas línguas vêm mostrando resultados diferentes daqueles inicialmente propostos pela Teoria do Labirinto.

28. Alguém atirou no empregado da atriz que estava na varanda.

Se a relativa “que estava na varanda” estiver aposta ao SN “a atriz”, a aposição será local, contudo, se estiver aposta ao SN “o empregado”, a aposição será não local.

Estudos experimentais demonstraram que no Inglês a preferência é pela aposição local, contudo, Cuetos e Mitchell (1988) em trabalho seminal, compararam a língua inglesa com a Espanhola e obtiveram resultados divergentes. No Espanhol a preferência foi pela aposição não local, o que veio a relativizar esse princípio. No Brasil, os trabalhos iniciados por Ribeiro (1998), trazem resultados, para o Português, semelhantes aos apresentados no espanhol, com uma preferência não local. A partir de então, o princípio vem sendo testado em outras línguas e os leitores apresentam preferências distintas de acordo as línguas particulares. Diante disso, os outros pressupostos da Teoria do *Garden-Path*, TGP, foram questionados por muitos pesquisadores.

No que se refere ao acesso do *parser* a um traço como a animacidade, de natureza semântica, a teoria do labirinto de Frazier (1979) não prediz que esse acesso ocorra. O chamado encapsulamento sintático, constitutivo dessa teoria, assume que há uma atribuição de uma estrutura sintática, de maneira que o leitor, no momento do processamento, entra, às cegas, num labirinto estrutural, comprometendo-se com uma única estrutura possível. Mais adiante, tendo confirmada a sua previsão sintática, o processamento prossegue sem custos adicionais. Se essa estrutura com a qual o *parser* se comprometeu não se confirmar, o leitor reanalisa a frase, o que gera um custo maior de

processamento (o efeito *Garden-Path*). Visto assim, um traço semântico não é acessado pelo *parser* no primeiro instante da leitura da frase. Ou seja, a teoria do *Garden-Path* assume um processamento serial, o que pressupõe uma imediatividade na análise do *input*, construindo um esqueleto constituído apenas por informações sintáticas.

Dessa forma, podemos afirmar que o modelo proposto por Frazier (1979) apreende bem o modo como a mente processa frases, contudo, como as informações são acessadas e em que momento isso ocorre nem sempre é tão claro. Quando as informações de natureza não sintática são acessadas, e em que medida elas interferem no processamento, é uma questão que precisa ser mais elucidada por evidências empíricas.

Na formulação inicial da teoria, Lyn Frazier e Janet Fodor (1978) propõem um modelo de processamento em dois estágios. O modelo proposto, denominado *Sausage Machine* (Máquina de Salsicha) prediz que o processamento ocorra num primeiro estágio, que consiste na atuação do *Preliminary Phrase Packager*, uma espécie de empacotador preliminar de sintagmas, que atribui nós lexicais e sintagmáticos para os *inputs* lexicais, formando cadeias ou pacotes estruturados, que são enviados para o segundo estágio, denominado *Sentence Structure Supervisor*, que combina os sintagmas formando um marcador frasal completo.

No primeiro momento, no estágio inicial, há uma visão restrita, em que o falante/ouvinte lida apenas com cerca de seis palavras. No segundo momento há uma visão mais ampla da sentença completa. Essa divisão de tarefas é fundamental para que a memória de trabalho não fique sobrecarregada, possibilitando o *parsing* em um tempo ótimo. Nas palavras de Maia, (2015, p. 16-17):

a TGP é uma teoria em que se propõem a existência de dois estágios fundamentais no processo de compreensão de frases: um processador modular, serial e de domínio específico produz uma análise sintática inicial da frase, baseado em métricas de simplicidade e localidade. Essa estrutura seria, então, interpretada, em um segundo estágio, menos reflexo do processo de compreensão.

Dito isso, esse modelo assume que o *parser* trabalha com decisões sem informações que estejam além do *input* imediato. Ou seja, o *parser* ignora tanto

as palavras que já foram computadas anteriormente, como ainda as que estão por vir, mas construindo hipóteses sintáticas sobre o que virá na sequência frasal.

A TGP se configura como um modelo de extrema importância para a psicolinguística como um todo. Sua longevidade atesta a pertinência de seus princípios. Formulada em 1979, desde quando dados empíricos vêm ora confirmando e ora refutando o que foi estabelecido quando de sua criação. Mesmo considerando a importância desse modelo, os fatos que motivam a pesquisa aqui instaurada são de uma natureza diferente dos pressupostos desse modelo. Procuramos atestar a influência de informações semânticas no processamento de sentenças nos momentos iniciais, a TGP nega tal possibilidade. Posto assim, o modelo não se adéqua aos nossos propósitos neste trabalho.

2.4.2 Modelos paralelos e não modulares: satisfação de condições

Noutra perspectiva teórica, MacDonald, Pearmultter e Seidenberg (1994) propõem um modelo que se enquadra nas Teorias de Satisfação de Condições, são os chamados modelos conexionistas. Nesse âmbito, a discussão sobre as decisões frente a uma ambiguidade estrutural é analisada sob outro viés. No que tange à TGP, as teorias de satisfação de restrições são bem distintas.

Para MacDonald, Pearmultter e Seidenberg (1994), a informação lexical tem a tarefa de direcionar as preferências de análise no caso de ambiguidades estruturais. Nessa perspectiva, há uma similaridade entre a resolução de uma ambiguidade estrutural com a resolução de ambiguidade lexical. Há um acesso a múltiplas análises, simultaneamente. A frequência de uso é determinante quanto aos níveis de ativação das análises alternativas, e o contexto é fundamental em todo processo. Posto dessa maneira, estruturas sintáticas não seriam construídas em tempo real, mas acessadas no léxico. Por isso mesmo, as preferências de uso em caso de ambiguidade estrutural seriam em parte determinadas pela frequência de uso de um dado item.

Dessa forma, por exemplo, um verbo usado predominantemente na língua como transitivo, assim será visto primeiro pelo *parser*. Ou seja, se um

traço semântico, como a animacidade, for de fato acessado pelo *parser*, qual seria a implicação disso para o *parsing* de uma sentença com termos animados ou inanimados?

Conforme apontamos na seção 2.3, o traço da animacidade afeta o processamento de relativas de sujeito e de objeto segundo apontam alguns estudos experimentais que manipularam esse traço. Estudos de *corpus* também demonstram que a distribuição de CRS e CRO nas línguas naturais é afetada pela animacidade. Um ponto que há de se ressaltar nessa discussão é o fato de os estudos experimentais que aferiram o processamento de relativas de sujeito e de objeto nas últimas décadas, em sua maioria, praticamente desconsiderarem esse traço na elaboração dos experimentos, e as estruturas animadas sempre foram as mais utilizadas experimentalmente.

Conforme já discutimos anteriormente nesta tese, discussão essa em que pretendemos nos deter com mais afinco no próximo capítulo, no qual faremos um estudo de *corpus* em torno da questão, a frequência de ocorrência de CRS e CRO com termos animados e inanimados não é equivalente nas línguas naturais. Mais que isso, estruturas relativas com antecedentes inanimados são bastante comuns, e normalmente não são testadas nos experimentos que abarcam o tema. Dito isso, questiona-se: se a frequência de uso é determinante na decisão do *parser* na análise de algumas estruturas na perspectiva dos modelos de satisfação de restrições, e se a animacidade afeta a distribuição de sentenças relativas na língua, essa frequência não conduziria o *parser* na análise de relativas de sujeito e de objeto?

No que tange especificamente à frequência, Ellis (2002) argumenta que há efeitos de frequência no processamento da fonologia, do léxico, da morfologia e da sintaxe. Segundo esse autor, teorias baseadas no uso sustentam que a aquisição da língua parte dos exemplos a que somos expostos no decorrer da vida. Assim, o aprendizado é fragmentado em milhares de construções e determinantes de produtividade de padrão incluem a lei poderosa da prática, numa disputa de traços e satisfação de restrição, regida por efeitos de tipologia e de frequência.

Ainda para Ellis (2002), o aprendizado de uma língua é a aprendizagem associativa de representações que refletem as probabilidades de ocorrência de mapeamentos forma-função. A frequência é, assim, uma chave determinante

da aquisição, pois "regras" da língua, em todos os níveis de análise (da fonologia, através da sintaxe, até o discurso), são regularidades estruturais que emergem da análise do tempo de vida do aprendiz das características distribucionais do *input* da língua.

Nessa perspectiva, em níveis mais altos, considera-se que a compreensão da língua é determinada pela vasta quantidade de informação estatística do ouvinte sobre o comportamento de itens léxicos em suas línguas.

Tais considerações de Ellis (2002) estão em conformidade com o modelo de MacDonald, Pearmultter e Seidenberg (1994), que leva em consideração a frequência de uso quando do processamento linguístico.

Os modelos conexionistas, baseados em restrição, têm como princípio fundamental o processamento paralelo, segundo o qual, há um número indefinido de estruturas sintáticas, no caso de ambiguidades, que são construídas em paralelo, com a dissolução da ambiguidade, uma delas prevalece. A DLT, contrariamente a isso, pressupõe a construção de uma e apenas uma estrutura sintática, porém, diferentemente da TGP, a construção dessa estrutura se dá a partir do uso de várias informações de naturezas diversas.

Levy e Gibson (2013) consideram os avanços que as teorias de satisfação de condições proporcionaram aos estudos psicolinguísticos. Contudo, são cautelosos ao precisarem o lugar em que ocorre a maior ou menor dificuldade de processamento da frase nesse modelo teórico. Essa consideração nos faz questionar se o processamento paralelo é condizente com o processamento da linguagem.

À parte o conceito de processamento paralelo, não admitido na DLT, no mais, o modelo descreve o processamento de estruturas linguísticas a partir de mecanismos semelhantes aos constitutivos dos modelos de satisfação de restrições.

2.4.3 A hipótese de *Construal*

Frazier (1995) argumenta que modelos de satisfação de restrições atribuem, à frequência, um papel relevante, sem, porém, mostrar de maneira clara como isso afetaria o processamento. Além do mais, segundo ela, a

frequência não comportaria análises em línguas distintas, uma vez que a frequência muda de acordo com as línguas particulares e, por isso, teorias com base na satisfação de restrições oferecem, segundo Frazier, apenas uma teoria das representações que são pré-armazenadas no léxico.

Frazier e Clifton (cf.1996, 1997), justamente no intuito de responder a essas questões não respondidas pela teoria do *Garden-Path*, formulam a hipótese de *Construal*, que procura explicar o processamento de estruturas não primárias.

As orações relativas sempre foram objeto de estudo em muitas línguas, estudos que acontecem sob diferentes perspectivas teórico-metodológicas. O princípio da Aposição Local, por exemplo, não é uniforme em todas as línguas no tocante às relativas, o que de alguma forma fragiliza a Teoria do *Garden-Path*. Os resultados de tais estudos, muitas vezes bem distintos entre si quando comparados entre línguas diferentes, motivaram Frazier e Clifton (1996) a formularem e apresentarem a hipótese de *Construal*, uma proposta que, segundo eles, admite uma subespecificação, que se refere a estruturas não primárias, no marcador frasal sintático construído durante a análise de sentenças *on-line* em quantidade e tipo muito limitados.

Para tanto, seus autores concebem que há, nas línguas naturais, estruturas primárias e não primárias, segundo Bezerra (2013),

diferentemente dos sintagmas primários, cujo processamento é regido pelos princípios estruturais, de modo que são apostos de imediato e exibem uma preferência para uma análise estruturalmente definida, os sintagmas secundários recebem uma análise subespecificada ou indeterminada, pois não são apostos de imediato à estrutura, mas associados a ela.

Para a hipótese de *Construal*, quando se afirma que há subespecificação no *parsing* de algumas estruturas, está se afirmando que o *parser* não constrói de imediato uma estrutura sintática determinada, mas, indeterminada e que pode, inclusive, ser guiado por informações não estruturais. Assim, o acesso a informações de natureza distinta da sintática pode acontecer nas construções não primárias, é preciso, contudo, delimitar em que tempo isso ocorre no *parsing*, algo que o modelo não deixa totalmente claro. Posto assim, o *parser*, para a hipótese de *Construal*, teria acesso, ainda

que em quantidade muito limitada e para construtos muito específicos, a informações de natureza semântica, pragmática e discursiva. Contudo, a maneira que tais informações são acessadas e, principalmente o momento em que isso acontece, ainda são questões a serem respondidas.

Segundo Frazier e Clifton (1997), diferentes teorias fornecem respostas distintas a esta questão da subespecificação. Na teoria *Garden-Path* (FRAZIER, 1979, 1987), a subespecificação não era possível. Para a hipótese de *Construal*, isso continua sendo verdade, se estivermos tratando de estruturas primárias, mas o mesmo não ocorre em estruturas não primárias.

Para Frazier e Clifton (cf. 1996, 1997), como já foi posto, o modelo de *Construal* é uma hipótese sobre a existência de subespecificações muito limitadas. No modelo teórico, uma análise imediata plenamente determinada é construída para "relações primárias", conforme definido abaixo em (a) e (b), as quais são analisadas como na teoria *Garden-Path* (FRAZIER, 1987; FRAZIER & CLIFTON, 1996).

- a. O sujeito e o predicado principal de qualquer cláusula finita;
- b. Complementos e constituintes obrigatórios de frases primárias.

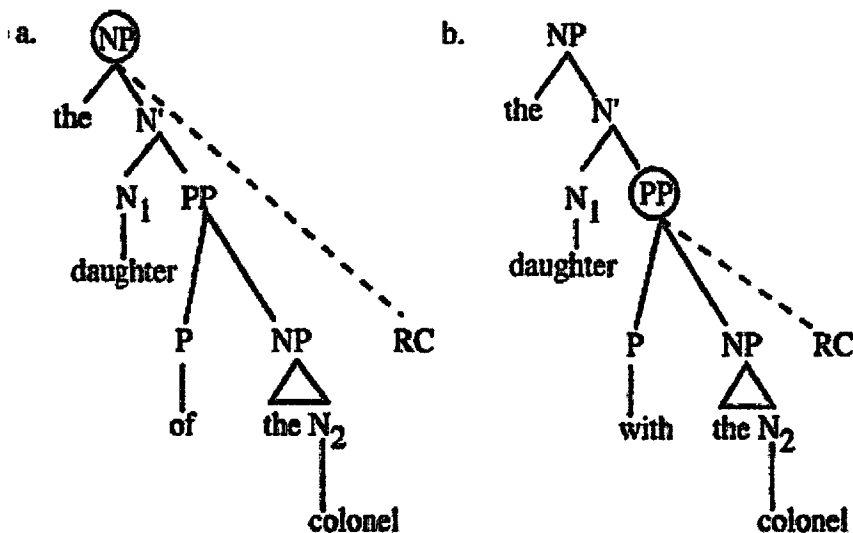
A hipótese do *Construal* assume uma análise subespecificada que é atribuída apenas a relações não primárias, conforme estabelecido em (c) e (d).

- a. Analise um *input*, X, como instanciando uma frase primária se possível;
- b. Caso contrário, associe X ao domínio de processamento temático corrente (a projeção estendida do último atribuidor *theta* real).

Para Frazier e Clifton (1997), um item *input* X é preferencialmente tomado para ser uma relação primária, e tais princípios enquanto aposição mínima governam sua análise. Se X não pode ser analisada como uma relação primária, este é associado ao domínio de processamento temático corrente.

Tomemos por base o processamento de cláusulas relativas complexas. Segundo a hipótese de *Construal*, o que será importante é a identidade do

último atribuidor *theta*, pois a cláusula relativa (CR) será associada à projeção estendida deste item, conforme ilustrado nas árvores abaixo, também transcritas de Frazier e Clifton (1997):



Como se vê nas árvores acima representadas, em (a), 'filha' (*daughter*) é o último atribuidor *theta* efetivo quando a cláusula relativa é encontrada. Desta forma, a cláusula relativa será associada à sua projeção estendida. Duas cabeças potenciais estão, assim, disponíveis: o SN mais longo encabeçado por N₁ 'filha' (*daughter*) e o SN mais curto encabeçado por N₂ 'coronel' (*colonel*). Diferentemente, em (b) o último atribuidor *theta* é a preposição 'com' (*with*). Assim, o domínio corrente é o PP encabeçado por 'com' (*with*) e apenas SN₂ 'o coronel' (*the colonel*) está disponível dentro deste domínio para funcionar como cabeça da cláusula relativa. (cf. FRAZIER E CLIFTON, 1997).

Para Frazier e Clifton (1997), uma variedade de fatores semânticos e pragmáticos influencia na seleção do processador. Modificadores (incluindo cláusulas relativas) preferencialmente modificam frases que são "referenciais". Postos os interesses da presente pesquisa, questionamos se o traço da animacidade poderia guiar o *parser* no *parsing* de cláusulas relativas de sujeito e de objeto. Contudo, a hipótese de *Construal* não esclarece o momento exato em que essas informações não sintáticas entram em cena. Em outras palavras, se algumas informações semânticas são acessadas durante *parsing* de estruturas não primárias, na formulação de Frazier e Clifton (cf. 1996, 1997),

mas o modelo não é preciso quanto ao modo e ao momento em que isso acontece.

Para os autores, relações não primárias, que são opcionais e não servem como complementos de itens léxicos, não influenciam na determinação da análise de um item lexical. Esta seria a razão de relações sintáticas especificadas amplamente serem imediatamente atribuídas apenas no caso de relações primárias potenciais. A sintaxe das relações primárias desambigua a descrição léxica de itens de *input* em qualquer ocasião particular de uso, porém, a precisão desse momento é relevante para uma análise da natureza que pretendemos aqui.

Frazier e Clifton (cf. 1996) revisaram evidências em várias línguas, como no Inglês, no Espanhol e no Alemão, que sugerem a existência de tal distinção básica e apresentam evidências para a afirmação do que a hipótese *Construal* faz acerca do papel do domínio de processamento temático corrente.

Os autores ainda apresentam evidências de como sintagmas não primários, associadas a um domínio em que elas poderiam receber múltiplas análises possíveis e serem interpretadas, e argumentaram que este processo de interpretação não parece envolver um processo de resolução através de satisfação competitiva de restrições alternativas.

Em um estudo de questionário escrito, Gilboy, Sopena, Clifton e Frazier (1995) investigaram a interpretação atribuída a sentenças com cláusulas relativas seguindo cabeças complexas contendo uma preposição não atribuidora de theta, como em (4a), na Tabela que se segue, ou a preposição atribuidora de theta 'com' (*with*) (em Espanhol, *con*), como em (4b).

Nos exemplos em Inglês e nas suas respectivas versões em Espanhol, falantes nativos (de Inglês ou Espanhol, respectivamente) atribuíram mais interpretações NP₂ a exemplos com preposições atribuidoras de *theta*, conforme indicado na Tabela 4, como esperado para a hipótese *Construal*.

Tabela 4: Resultados de Frazier e Clifton, (1997).

Frases	Tipo de preposição	% NP ₂	
		Inglês	Espanhol
a. Os Turistas admiraram o museu da cidade que eles visitaram em Agosto ¹¹ .	Não atribuidora de <i>theta</i>	53	49
b. O conde pediu o bife com o molho que o <i>chef</i> preparou muito bem ¹² .	Atribuidora de <i>theta</i>	69	83

Quando mais de um lugar de aposição permissível está disponível dentro do domínio de processamento temático corrente, como o processador escolhe entre eles? Para os autores, uma variedade de fatores semânticos e pragmáticos influencia na seleção do processador. Modificadores (incluindo cláusulas relativas) preferencialmente modificam frases que são "referenciais" no sentido de que eles são introduzidos por um determinante. Exemplos como estes da Tabela 5 que se segue foram incluídos no estudo de Gilboy et al. (1995).

Tabela 5: Exemplos do frases experimentais de Gilboy et al. (1995).

Frases	Tipo de NP	% de escolha de NP ₂
a. Ontem me entregaram o suéter de algodão que foi importado ilegalmente ¹³ .	Não referencial	26
b. Ontem me entregaram o suéter do algodão que foi importado ilegalmente ¹⁴ .	Referencial	55

Mais respostas NP₂ foram dadas quando NP₂ foi introduzido pelo determinante (5b) do que quando não foi (5a). Esses resultados, que foram reportados por Gilboy et al. (1995) tem sido replicados de maneira bem sucedida em Espanhol por Igoa (1995) e em Alemão por Hemforth, Konieczny & Scheepers, (1996).

¹¹ The tourists admired the museum of the city that they visited in August.

¹² The count ordered the steak with the sauce that the chef prepared especially well.

¹³ Yesterday they gave me sweater of cotton that was illegally imported.

¹⁴ Yesterday they gave me the sweater of the cotton that was illegally imported.

Modificadores são também atraídos a uma cabeça realçada (*focused*). Schafer, Clifton e Frazier (1996) manipularam as posições do foco em um estudo de compreensão auditiva, usando sentenças como as exemplificadas abaixo:

Tabela 6: Resultados de Schafer, Clifton e Frazier, (1996).

Frases	% escolha de NP ₂
a. O detetive examinou a ENTRADA da casa que mostrou claros sinais de danos ¹⁵ .	34,9
b. O detetive examinou a entrada da CASA que mostrou claros sinais de danos.	47,4

Conforme esperado, NP₂ foi escolhido mais frequentemente quando N₂ 'casa' foi destacado (6b) do que quando N₁ 'entrada' foi (6a).¹⁶ Desta forma, cláusulas relativas preferem cabeças que são referenciais e focalizadas.

Compreendemos, pois, que diante de uma análise feita pelo marcador frasal em estruturas relativas, pode ocorrer uma subespecificação. Contudo, os exemplos supracitados de trabalhos experimentais não respondem com clareza os questionamentos empreendidos nesta tese.

Mais que isso, há evidências experimentais, explicitadas na próxima sub-seção, que atestam o acesso a traços semânticos, bem como a interferência do contexto nos primeiros momentos do processamento, mesmo em estruturas primárias. Isso nos leva a considerar a arquitetura de modelos de processamento mais interativos, que podem em alguns aspectos, responder melhor aos nossos anseios.

2.4.4 Modelos interativos: DLT

A DLT, (The Dependency Locality Theory), de Gibson (2000), é uma teoria sobre complexidade linguística. Segundo sua proposta, esse modelo associa custo de integração estrutural crescente com a distância de aposição,

¹⁵ The detective eyed the entrance of the house that showed clear signs of damage.

¹⁶ Resultados adicionais reportados por Schafer et al. (1996) mostram que o efeito do foco não deve ser tomado para refletir um processo de casamento do estado informacional da cabeça e da cláusula relativa (dado x novo), mas pode em vez disso simplesmente refletir a saliência adicional da cabeça realçada (FRAZIER & CLIFTON, 1997, p. 283).

e custo de armazenamento com categorias sintáticas preditas. Esta teoria, em Português: “Teoria da Dependência Local”, provê uma unificação de um arranjo amplo de fenômenos de processamento díspares, incluindo o tempo de leitura *on-line* de cláusulas relativas com extração de sujeito e de objeto.

Conforme já afirmamos, há “modelos que assumem um *parser* interativo e um processamento paralelo, isto é, um processador que não tem acesso restrito apenas à informação sintática no momento *on-line* da compreensão e está apto para postular múltiplas análises possíveis perante uma sentença ambígua” (BEZERRA, p. 20, 2013). Existem modelos não modularistas, a exemplo do modelo conexionista (de Satisfação de Condições), de MacDonald, Pearlmutter, Seidenberg, (1994), e o interativo, de Gibson (2000), Gibson e Fedorenko (2011). Esses dois modelos são muito diferentes, pois o conexionista compreende o processamento como paralelo, enquanto o interativo descreve o processamento como incrementacional e serial, mas ambos compreendem o *parser* como sensível à informação semântica, não sendo, portanto, modulares.

Perceba-se: a TGP se caracteriza por ser serial e modular, o conexionismo por ser paralelo e não modular, já os modelos interativos são seriais e não modulares.

A Teoria da Dependência Local, DLT (GIBSON, 2000), portanto, enquadra-se no *hall* dos modelos interativos, modelos segundo os quais, qualquer tipo de informação (semântica, sintática, discursiva, frequência) que estiver disponível no curso da computação da sentença pode ser acessado e processado imediatamente. Dessa maneira, embora o processamento seja incrementacional e serial, não há um tipo de informação que seja privilegiada. Visto assim, informações semânticas podem auxiliar o *parsing* sem com isso tornar o processo mais custoso.

Para maior elucidação dessas questões, é preciso estar atento para o fato de que a compreensão linguística possui momentos iniciais e tardios. Para os modelos que partem de pressupostos modularistas, o momento inicial é quando o *parser* opera atribuindo ao material linguístico uma representação estrutural. No momento mais tardio, ou final, é que são acessadas e integradas informações de outros níveis linguísticos.

Sendo assim, o primeiro estágio é mais reflexo e o segundo, o mais reflexivo do processamento. Nas palavras de Kenedy, (2015, p.152), “entende-se que os modelos possam assumir ou um processamento serial e incremental ou, por oposição, um processamento paralelo e distribuído”. Por sua vez, “a DLT identifica-se como um modelo interativo porque, tal como a TGP, assume a serialidade e a incrementalidade na construção de representações linguísticas, mas diferentemente da TGP, sustenta que certos tipos de informações não estruturais podem ser vistas pelo *parser*” (KENEDY, 2015, p.153).

Ainda para Kenedy (2015, p. 153), “talvez a primeira e mais relevante informação não estrutural a considerar como atuante no processamento de frases seja a frequência”. O autor diz ainda que por frequência se entende a familiaridade que temos com determinados itens linguísticos. Mesmo que a experiência individual com a língua possa apresentar diferenças consideráveis de falante para falante, é inegável que haja construções mais ou menos comuns em uma língua particular.

O trabalho de Ellis (2002) citado na seção anterior, também referido por Kenedy (2015), demonstra que “palavras que apresentaram alto índice de ocorrência em muitos *corpora* do Inglês” foram lidas mais rapidamente em tarefas experimentais do que palavras com baixa frequência. Visto assim, o próprio efeito *Garden-Path* poderia ser provocado não apenas por questões de ordem estrutural, mas pela frequência com que uma palavra ou estrutura é usada na língua.

Kenedy (2015, p. 154) argumenta ainda que, se considerarmos o efeito de frequência na formulação de modelos psicolinguísticos *on-line*, o contexto deve ser acessado pelo *parser*. Contudo, os estudos até agora empreendidos nesse sentido ainda não são conclusivos do ponto de vista empírico.

Na proposição da Teoria da Dependência Local (DLT), Gibson (2000) evidencia como são processadas algumas estruturas frasais a partir do custo de processamento relacionado à distância da aposição. Segundo ele, durante o *parsing* da sentença, há dois componentes a serem considerados: a integração de novos referentes do discurso e o custo de armazenamento desses referentes.

Em seu trabalho, o autor evidencia suas considerações a partir de dados empíricos e inclui, entre seus exemplos, a cláusulas relativas encaixadas. Grosso modo, ele considera a complexidade dessas estruturas numa escala crescente. Há aquelas mais ou menos complexas pelo número de encaixes, mas também pelo tipo de encaixe. Há um aumento da complexidade, segundo esse autor, quando da inclusão de termos com dependência de caso distintas entre si (acusativo ou nominativo, por exemplo).

Uma extensão da hipótese da dependência incompleta prediz que a complexidade é indexada pela quantidade máxima de dependências incompletas do mesmo tipo, em que duas dependências sintáticas são iguais se o mesmo Caso¹⁷ (como o Caso nominativo para o sujeito de um verbo, e o Caso acusativo para o objeto de um verbo, e assim por diante) é atribuído na relação. De acordo com esta hipótese, tipos diferentes de dependências sintáticas incompletas não interferem umas nas outras, mas dependências sintáticas incompletas similares simultaneamente presentes num estado de processamento são difíceis para que o processador mantenha seus traços. (cf. GIBSON, 2000).

A complexidade da cláusula relativa está nela mesma, independente de haver, por exemplo, ambiguidade na estrutura frasal. Visto assim, Gibson parte da análise de relativas para descrição dos fatos da teoria que propõe. E para tanto, ele parte de trabalhos que, conforme estamos discutindo desde o início desta tese, desconsideram o traço da animacidade, ou pelo menos não o controlam sistematicamente.

Gibson (2000) afirma que a complexidade da extração do objeto é capturada através de várias medidas, incluindo o monitoramento de fonema, decisão léxica *on-line*, tempos de leitura e precisão de resposta a questões de teste, para tanto, o autor cita Holmes (1973); Hakes, Evans e Brannon (1976); Wanner e Maratsos (1978); Holmes e O'Regan (1981); Ford (1983); Waters, Caplan e Hildebrandt (1987); King e Just (1991). Ainda para Gibson, além da complexidade a que se referem esses autores, estudos em neurolinguística revelam que o volume de fluxo de sangue no cérebro é maior em áreas

¹⁷ Fenômeno linguístico que permite a discriminação de argumentos de um predador. (KENEDY, 2013, p. 239).

relacionadas ao processamento linguístico para extrações de objeto do que para extrações de sujeito, para tanto ele cita os trabalhos de Stromswold et al. (1996) e Just et al. (1996). Por fim, Gibson (2000) afirma que pacientes com histórico de infarto afásico não podem responder confiavelmente às questões de compreensão sobre as CRO, embora eles apresentem bom desempenho em CRS (CARAMAZZA E ZURIF 1976; CAPLAN E FUTTER 1986; GRODZINSKY 1989; HICKOK, ZURIF E CANSECO-GONZALEZ 1993).

Para comprovação de parte de seus argumentos, o autor faz um estudo experimental com relativas de sujeito e de objeto, as quais o autor exemplifica com as frases traduzidas abaixo:

a. Extração de sujeito

29. O repórter que enviou o fotógrafo para o editor esperava por uma boa história¹⁸.

b. Extração de objeto

30. O repórter que o fotógrafo enviou para o editor esperava por uma boa história¹⁹.

Nos experimentos realizados pelo autor, as CRS, como acontece na maioria dos estudos de assimetria entre CRS e CRO, foram lidas mais rapidamente também nos estudos de Gibson (2000).

Segundo Kenedy (2015, p 154), “por princípio, a DLT assume que se uma dada informação se tornar visível no *input*, então ela poderá influenciar as decisões do processador. Trata-se de uma questão eminentemente empírica a ser explorada”. Ora, a animacidade seria um fator “visível no *input*”? Ou mais que isso, a frequência de termos animados e inanimados interfere no processamento desses mesmos termos?

O modelo ora apresentado, proposto por Gibson (2000), atende às hipóteses de trabalho da presente tese, uma vez que o modelo considera fatores não sintáticos na configuração do processamento linguístico. Ao que mais adiante discutiremos com mais afinco.

Antes dessa formulação da DLT, Gibson (1998) propôs a Teoria da Predição Sintática Local (Sintactic Prediction Locality Theory), SPLT²⁰, teoria

¹⁸The reporter who sent the photographer to the editor hoped for a good story.

¹⁹The reporter who the photographer sent to the editor hoped for a good story.

²⁰ Teoria da Predição Sintática Local – Manteremos a sigla em Inglês – SPLT.

segundo a qual quando o leitor encontra a cláusula relativa, há uma diferença no número de elementos requeridos entre cláusulas relativas de sujeito e cláusulas relativas de objeto. Para completar a cláusula como uma cláusula relativa de sujeito, dois elementos são necessários: um traço de SN de sujeito e um verbo. Para completar a sentença como uma cláusula relativa de objeto, três elementos precisam ser mantidos na memória: um SN de sujeito, um traço de SN de objeto e um verbo. Como o *parser* prefere a interpretação com o mínimo de elementos requeridos, a cláusula é analisada como uma cláusula relativa de sujeito.

Perceba-se, a SPLT não prediz que a decisão possa ser influenciada pelo conteúdo semântico das entidades na cláusula relativa. Cabe-nos ainda ressaltar que Mak, Vonk e Schriefers (2002), apontam justamente para esse fato, segundo eles essa teoria então não abarcaria os fenômenos por eles abordados naquele trabalho, como também assumiremos aqui nesta tese.

Mak, Vonk e Schriefers (2002) também discutem a Active Filler Strategy²¹, a AFS, (FRAZIER, 1987; FRAZIER & FLORES D'ARCAIS, 1989). A AFS é baseada na diferença de estrutura sintática entre cláusulas relativas de sujeito e objeto. Nesse modelo, o pronome relativo na cláusula relativa de sujeito origina na posição do sujeito e é então movido para a posição inicial da cláusula. O pronome relativo na cláusula relativa de objeto é criado na posição de objeto e é também movido para a posição inicial da cláusula. A posição vazia que é deixada pelo pronome relativo na estrutura é chamada de *gap*, e o pronome relativo é o antecedente (*filler*) para este *gap*.

A AFS estabelece que quando um leitor encontra um *filler*, ele é endereçado à posição do *gap* anterior mais provável. Em cláusulas relativas, isto significa que o *filler* é relacionado à posição do sujeito, pois é a posição mais anterior possível. Esta estratégia é correta para cláusulas relativas de sujeito, mas é problemática em cláusulas relativas de objeto, pois a posição do sujeito já está ocupada. Assim, a AFS propõe que o leitor escolha a análise da cláusula relativa imediatamente no pronome relativo. A AFS (FRAZIER, 1987; FRAZIER & FLORES D'ARCAIS, 1989) é baseada na estrutura sintática da cláusula relativa. Informação semântica não pode influenciar a decisão de

²¹ Estratégia do Antecedente Ativo – Manteremos a sigla em Inglês – AFS.

parsing. Segundo Mak, Vonk e Schriefers (2002), para a AFS, a decisão de *parsing* é devida à diferença na estrutura sintática entre as cláusulas relativas de sujeito e objeto, e esta diferença é a mesma independente do objeto da cláusula relativa ser animado ou inanimado. Dito isso, assumimos, juntamente com Mak, Vonk e Schriefers (2002) que a AFS não constitui um modelo a partir do qual poderíamos analisar nossos dados.

A DLT, por outro lado, é um modelo de processamento que considera fatores estruturais quando do *parsing* de sentenças, sem, contudo, desconsiderar informações semânticas e pragmáticas, disponíveis no *input*, por exemplo. Para a DLT, a frequência também contribui para a facilidade/dificuldade de leitura de uma sentença. Noutras palavras, Gibson (cf.2000) assume que fatores como a frequência do item léxico, o custo de integração estrutural desse item, o seu custo de armazenamento, a plausibilidade da estrutura resultante e a complexidade do discurso são elementos a serem considerados durante o processamento *on-line*.

Para um olhar sobre essas questões, tomaremos algumas considerações de Traxler (2012) que discorre sobre os modelos de dois estágios, a exemplo da TGP, e modelos de satisfação de restrições.

Ao trazer à baila essas considerações, não estamos assumindo, para esta tese, os modelos de satisfação de restrições, a exemplo de MacDonald, Pearlmutter, Seidenberg, (1994), que concebem o processamento como conexcionista, em que múltiplas análises sintáticas são construídas em paralelo. Pretendemos, antes, esclarecer como trabalhos experimentais vêm desvendando o *parsing* de determinadas estruturas incluindo em suas análises a semântica e a frequência, a exemplo do que admite a DLT de Gibson (2000). Em tempo, cabe ser posto que a DLT, um modelo iterativo, que embora seja incrementacional, como a TGP, não prediz o encapsulamento sintático assumido por esta última. Ou seja, a DLT assume que outros fatores, além dos sintáticos, possam guiar o *parser* nos momentos iniciais do processamento.

Traxler (cf. 2012) traz discussões em torno da interferência do contexto no processamento de sentenças. Para a TGP, apenas informações sintáticas são acessadas pelo *parser*, posto assim, uma estrutura potencialmente ambígua não pode ser desambiguizada por informações fornecidas pelo

contexto. Ou seja, uma estrutura como 31 será tão custosa quanto a parte sublinhada da estrutura 32.

31. Encontrei a professora do meu amigo que mora em São Paulo.

32. Meu amigo tem uma professora que mora em Natal e outra que mora em São Paulo. Fui a um congresso esta semana e encontrei a professora do meu amigo que mora em São Paulo.

Embora a TGP desconsidere que o *parser* possa acessar o contexto trazido antes da sentença, Traxler (2012) relata que estudos experimentais vêm comprovando que o contexto facilita a leitura. Nesse caso, não havendo ambiguidade quanto à aposição de “*que mora em São Paulo*”, já que o contexto fornece essa informação, o leitor não entra no *Garden-Path* e processa 32 mais facilmente.

Essa consideração já coloca em xeque o conceito de encapsulamento sintático de que trata a TGP, pois pressupõe que informações contextuais são postas em cena durante o *parsing* dessas sentenças. A DLT (GIBSON, 2000) prevê isso, admitindo que outros fatores guiem o *parser* no momento inicial.

A frequência, segundo Traxler (2012) e Kenedy (2015), constitui um papel importante quando do processamento linguístico. A TGP prediz que o *parser* usa apenas informações da categoria da palavra para tomar decisões iniciais sobre quais estruturas sintáticas ele irá construir. Mas as palavras podem oferecer mais informações que isso. Traxler (2012, p. 153-154) exemplifica tal afirmação com os verbos “*tirar*” e “*colocar*”. Ambas as palavras pertencem à mesma categoria sintática, são verbos. Porém, esses verbos não são equivalentes. Para tanto, é preciso considerar a que tipo de estrutura sintática pertencem tais verbos. Eles são transitivos, mas temos outras considerações a fazer quando observamos os exemplos abaixo.

33. Tales tirou uma boa nota.

34. Tales colocou uma xícara. (?)

35. Tales colocou uma xícara na mesa.

Não podemos dizer que 34 seja uma sentença agramatical, porém, 35 parece melhor que 34. Isso acontece porque, embora ambos os verbos sejam

transitivos diretos, “*tirar*” é apropriado com apenas um objeto direto, mas “*colocar*” requer um objeto direto e um objetivo. Segundo Traxler (2012), são duas palavras que pertencem à mesma categoria sintática por serem verbos, todavia se diferem por possuírem subcategorias distintas entre si. Ainda segundo o autor, esses verbos exigem complementos bastante específicos, mas isso não se aplica a todos os verbos das línguas naturais, outros verbos podem ser mais flexíveis e exemplifica com o verbo “*ler*”, que, ainda segundo o autor, pode, por exemplo, ocorrer sem um complemento.

36. Sofia estava lendo.

Ele pode aparecer com um objeto direto, como em 36, em que o mesmo verbo é transitivo:

37. Sofia estava lendo uma história.

Ele pode aparecer ainda com um objeto direto e um objeto indireto, como em 37, sendo assim bitransitivo.

38. Sofia estava lendo uma história para Diadorin.

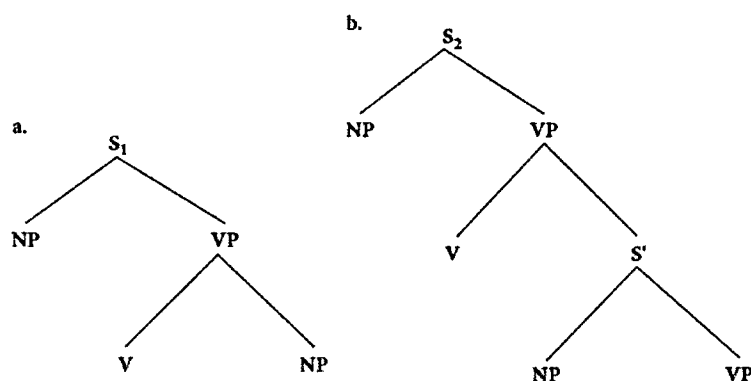
Ou seja, para Traxler (2012, p. 154) “*ler*” apresenta “algumas possibilidades de subcategoria, incluindo intransitivo, transitivo e bitransitivo; e cada uma destas possibilidades de subcategoria está associada a uma estrutura sintática diferente”. Para o autor, o *parser* baseado em restrição é diferente do *parser* modular, pois para as teorias baseadas em restrição a informação estrutural é “associada a palavras individuais no léxico e esta informação influencia quais hipóteses estruturais serão buscadas no decorrer do processamento das sentenças” (TRAXLER (2012, p. 154). Assim, o *parser* baseado em restrição se vale das informações de subcategorias quando há, por exemplo, mais de uma estrutura compatível com o input corrente.

Posto dessa maneira, 39 é vista como uma estrutura mais fácil de processar do que 40, tanto na TGP quanto em teorias de satisfação de restrições, mas por motivos distintos.

39. O estudante viu a resposta para a questão.

40. O estudante viu a resposta que estava no livro.

As árvores que se seguem, (a) e (b), representam, respectivamente, 39 e 40 e foram retiradas do trabalho de Traxler (2012, p. 156).



Como se vê, 39 é uma estrutura mais simples que 40, o que por si já satisfaz, para a TGP, a explicação de 40 ser mais custosa, uma vez que em 39 o leitor continuaria a trabalhar com o SV corrente. Há outros fatores, porém, que Traxler (2012) aponta para tal explicação com base nas teorias de satisfação de restrições. Segundo o autor, quando processamos uma sentença, informações relativas à probabilidade de uso de uma subcategoria podem ser acessadas. Ou seja, se 40, mesmo sendo mais simples que 39, possuir uma frequência muito maior na língua, o parser se vale dessa informação e 40 pode ser tão rápida de ser processada quando 39.

Ou seja, para Traxler (2012) um parser interativo se vale de informações constitutivas da experiência do falante e prediz a sua realização futura. Assim, a despeito da complexidade maior de uma dada estrutura, sua maior ou menor facilidade de processamento também está relacionada com o uso mais ou menos frequente dessa estrutura na língua.

Nos exemplos em questão, as predições da TGP foram as mesmas das teorias de satisfação de restrição, mas nem sempre isso acontece. A DLT, que admite a interveniência de fatores não estruturais, como a frequência, por exemplo, compreende o processamento de 41 e 42, a seguir, de forma distinta daquela compreendida pela TGP.

41. Dr. Phil realizou seus objetivos cedo²².

42. Dr. Phil percebeu que seus objetivos estavam fora de alcance²³.

Para Traxler (2002, p. 156), a teoria do *Garden-Path* tomará 42 como uma estrutura mais complexa e, portanto, mais custosa que 41, por razões semelhantes às aquelas elencadas anteriormente nesta seção. Contudo, as teorias baseadas em restrição argumentam que a frequência de uso muda esse quadro e ambas as estruturas possuem a mesma facilidade de processamento.

O que ocorre, nesse caso, é que o verbo “*realizar*”, frequentemente, possui um SN como complemento, em oposição, o verbo “*perceber*”, possui uma cláusula como complemento. Assim, embora 41 seja mais simples que 42, em 42, a frequência maior de complementos sentenciais para o verbo “*perceber*”, que, segundo Traxler (2012), é de 90% dos casos em inglês, faz com que haja um equilíbrio entre tais estruturas, e com que elas sejam umas tão fáceis de processar quanto as outras.

Ou seja, segundo Traxler (2012, p. 157) “o custo da estrutura improvável é equilibrado pelos benefícios de ter uma estrutura simples²⁴”. Para análise de nossos dados, no capítulo 4, tomaremos essa prerrogativa, perfeitamente aplicável aos resultados que apresentaremos.

A partir de estudos experimentais com Rastreamento Ocular e de Leitura Automonitorada Garnsey, et al. (1997) e Trueswell *et al.* (1993) mostraram que sentenças como 39 ocasionam uma menor dificuldade de leitura. Para Traxler (2012), a explicação disso está na assunção de que a probabilidade de um dado verbo se agrupar com determinadas estruturas é acessada pelo parser e isso influencia o seu processamento. Diante disso, o parser tenta antecipar a estrutura que está por vir, o que orienta as suas decisões.

Ou seja, o falante observa o quão frequente ele encontra estruturas sintáticas diferentes, de maneira a usar esta informação armazenada para ordenar as possibilidades diferentes. “Assim, uma sentença com um verbo

²² Dr. Phil realized his goals early on.

²³ Dr. Phil realized his goals were out of reach.

²⁴ The cost of the unlikely structure is balanced by the benefits of having a simple structure.

transitivo na voz ativa, por exemplo, muito comum em todas as línguas, seria sempre uma estrutura fácil de processar” (TRAXLER, 2012, p. 157), em seu texto, Traxler discute, inclusive a influência da animacidade, objeto de nossas investigações na presente pesquisa.

Alguns estudos vêm demonstrando claramente que a animacidade é acessada pelo *parser*, e um verbo na voz ativa, por exemplo, pode ter seu tempo de processamento comprometido, se, por exemplo, o sujeito for inanimado, já que a frequência muito superior de sujeitos animados já aponta para uma estrutura sintática em que o termo animado tenderá a ser o sujeito da sentença.

Visto assim, temos nesse contexto, pelo menos dois fatores fortemente aliados na construção de predições de que o *parser* pode lançar mão quando da análise de uma estrutura linguística: a semântica e a frequência.

A DLT, de Gibson (2000), prediz que o custo de integração e o custo de armazenamento de uma estrutura constituem fatores que interferem no tempo de leitura de uma estrutura dada. Além disso, a DLT prevê também que a distância de aposição entre itens sintáticos podem influenciar o custo de processamento. Kenedy (2015, p. 156), considera que “a integração diz respeito à tarefa do processador de relacionar sintaticamente um novo elemento introduzido no *input* linguístico à estrutura sintática já representada na memória de trabalho”.

Essa integração não acontece às cegas para a DLT, (GIBSON, 2000), como sugere a TGP (FRAZIER, 1979). Diante disso, assumimos a prerrogativa da DLT em nosso trabalho, uma vez que para descrição e explicação dos nossos experimentos tomaremos, para além da estrutura sintática, pelo menos um fator de natureza semântica – a animacidade – e, por conseguinte, a frequência desse fator em conjunto com os tipos de relativa, se de sujeito ou se de objeto, na língua Portuguesa nas suas variedades Brasileira e Europeia. Para Kenedy (2015, p. 154), “embora a experiência individual com a linguagem seja intensamente variável, é notável que certas palavras e construções possuam maior frequência de uso em determinadas comunidades em comparação com outras palavras e ou construções”.

Dito isso, assumimos o modelo interativo de Gibson (2000), a DLT, como modelo teórico basilar desta tese na análise de nossos dados. Consideraremos

a distância de aposição de termos relativizados em CRS e CRO, considerando as diferenças estruturais que facilitam as CRS quando comparadas às CRO. Todavia, procuraremos compreender o efeito que tem o traço da animacidade, um traço semântico, como argumentaremos na próxima seção, mas que pode provocar diferenças nos tempos de leitura dessas estruturas. Nossa análise será pautada também na frequência de ocorrência dessas estruturas, o que faremos a partir de um estudo de *corpus* que também é constitutivo desta tese, vide capítulo 3.

Em poucas palavras: a DLT será o modelo teórico por nós assumido na análise de relativas de sujeito e de objeto, estruturas sintaticamente diferentes entre si, e que carregam no termo relativizado o traço da animacidade, um traço semântico com frequências distintas conforme a condição analisada.

2.5 A animacidade: uma breve definição

Desde o início deste texto temos nos referido, genericamente, à animacidade, sem uma definição mais detalhada nos moldes que este trabalho exige. Nesta seção, tomaremos o termo para uma breve discussão com o fim de definir o que se entende por animacidade no âmbito da pesquisa linguística, e mais especificamente nos propósitos desta tese. A questão que se nos apresenta é o *status* do traço da animacidade por sua natureza, assumida neste trabalho, como semântica, ao que mais adiante argumentaremos.

Nairne et al. (cf. 2013) argumentam que a animacidade representa uma variável potencialmente sem controle na pesquisa cognitiva. Tal argumento, em algum nível, coaduna com nossas hipóteses, visto que a análise que se dá nas diferenças de processamento entre as CRS e as CRO é fortemente enviesada devido à ausência de controle dessa variável na maioria das pesquisas que tratam do tema.

Nairne et al. (cf. 2003) ainda afirmam que a animacidade provavelmente melhorou a aptidão inclusiva em algum momento do passado ancestral do ser humano. Noutras palavras, a evolução da espécie humana, mais especificamente da capacidade mnemônica, esteve fortemente atrelada à distinção entre seres animados e inanimados. Essa distinção é particularmente importante aos seres humanos por sua condição de presa, pela necessidade

inerente de se proteger dos predadores, como ainda por sua condição de predador na busca por suas presas.

Os principais dicionários da língua Portuguesa ainda não registraram o termo, embora o mesmo já venha sendo usado em diversas áreas da linguística, pelo menos, desde meados do século passado. Diante disso, procuramos construir um significado para a palavra a partir dos cognatos – ânimo, animar, animação – dos quais a palavra animacidade teria se derivado.

Segundo o Michaelis *On-line*, “ânimo: SM (lat. animu) 1 Alma, espírito, mente (...) desejo, intenção, consciência individual” ou seja, de acordo com o verbete, ter ânimo é basicamente ter alma, em algum nível, ter vida.

Ainda segundo esse mesmo dicionário, “animado *adj* (*part* de *animar*) 1 Dotado de animação”. Dito isso, fomos ao significado de animação, e encontramos “animação, (*animar*+ção) 1 Ação de animar. 2 Alegria, entusiasmo. 3 Movimento. 4 Calor, vivacidade”. Visto dessa forma, a animacidade, substantivo ainda sem definição em alguns dicionários, está fortemente atrelada à sua origem latina, *animu*, alma, vida.

O universo é constituído por seres animados e inanimados. A distinção, embora dicotômica, não é simples, seres animados não são o oposto de seres inanimados e mais, há gradações, há, entre os seres animados, aqueles que são mais ou menos animados, o mesmo ocorre entre os inanimados.

A princípio, palavras como “*casa, pedra, computador, vento e água*” são termos inanimados, em contraste com menino, “*lagarta, girafa, Pedro e árvore*”, por exemplo. Mas não se pode negar que “*vento*” pode ser, em algum aspecto, mais animado que “*pedra*”, mesmo não sendo um ser vivo. Como “*árvore*”, que embora seja um ser vivo, não possui o mesmo grau de animacidade que “*girafa*”. Mesmo entre os animais, os humanos talvez sejam referidos no discurso com um grau de animacidade distinto de outros animais.

Cabe-nos então considerar essa gradação: as coisas, os objetos do mundo real são, em graus distintos, mais ou menos animados. A animacidade é portanto o que define um ser como dotado de vida, em seus muitos aspectos.

A compreensão sobre o mundo se dá, inclusive, pela compreensão da animacidade das coisas que nos cercam. Nairne et al. (2013), já citados nesta seção, discutem o valor mnemônico da animacidade, e argumentam a partir de seus estudos que o reconhecimento de seres animados e inanimados é de

fundamental importância para a evolução das espécies. Em argumentação semelhante, Santos (2013, p.37), reportando um estudo experimental em que patos e gansos foram testados ao nascer, afirma que “esses animais são geneticamente programados para se depararem com sua mãe, a reconhecerem e a acompanharem; ou, na ausência de sua mãe, fazem isso com outro ser animado que esteja com eles no momento do nascimento”.

No âmbito da psicolinguística, em um trabalho experimental, Nairne et al. (2013) programaram um experimento em que foram testadas palavras cujos referentes eram animados ou inanimados. Os participantes foram expostos a um grupo de 24 palavras, metade animadas e metade inanimadas. Em seguida tiveram que lembrar as palavras em qualquer ordem. A quantidade de palavras animadas lembradas foi significativamente mais alta do que as palavras inanimadas. Segundo os autores, é possível que os estímulos animados sejam bem lembrados em razão de prioridades perceptivas, mais do que por razão simplesmente mnemônica.

Ainda para Nairne et al. (cf. 2013), será igualmente importante desconstruir a dimensão da animacidade. A categoria animada pode ser definida de diversas maneiras. Em seu nível mais abrangente, é possível considerar qualquer criatura viva como animada, incluindo animais não-humanos, plantas e talvez até bactérias e vírus. De outro modo, predisposições cognitivas podem requerer movimento, estados mentais ou mesmo habilidade de se comunicar a fim de que algo possa ser considerado animado.

Nesta tese, consideramos animado ou inanimado um item com base na sua representação de “coisa viva”. Contudo, preferimos não incluir as plantas como itens animados em nossos experimentos, por compreender que há gradações e que os participantes da pesquisa poderiam, por exemplo, considerar as plantas menos animadas do que os animais. Parte de nossos experimentos foi feita a partir de nomes próprios, o que possibilitou com que trabalhássemos apenas com termos animados e humanos.

Santos (2013), em estudo que procurou discutir a animacidade do ponto de vista estrutural, pretendia delimitar o traço da animacidade para além de sua abordagem clássica que o define como traço semântico.

O autor diz que alguns trabalhos apontam o fato de a animacidade, em algumas línguas, estar presente na concordância (*agree*), contudo, ao analisar

13 diferentes línguas (Pérsia, Búlgaro, Japonês, Kotiria (família Tukano), Ruwund, Kiriimi, Maasai, Swahili, Ndengeleko, Russo, Hindi-urdu, Navajo e Português), constatou que a animacidade está envolvida com outros fenômenos linguísticos de natureza sintática, como a atribuição de caso, nas estruturas argumentais, na hierarquia de papéis *theta*, no apagamento de constituinte entre outros fenômenos.

Para Carvalho (2013, p.118), porém, no programa Minimalista “traços podem ser considerados a substância da sintaxe”, uma vez que “todas as operações básicas da sintaxe dependem exclusivamente de traços para serem ativadas”. Perceba-se, as operações sintáticas são ativadas por traços de diversas naturezas, isso não faz com que esses traços sejam necessariamente sintáticos. Ou seja, o que Santos (2013) demonstra são implicações do traço da animacidade, essencialmente semântico, em operações sintáticas que ocorrem em algumas línguas naturais, talvez todas.

Santos (2013, p. 12) afirma que “o traço de animacidade é um traço bastante especial, pois é um traço sintático/formal, isto é, participa de operações sintáticas de fenômenos de várias ordens”. O autor diz ainda que, apesar disso, “o traço da animacidade é saturado na sintaxe, havendo ainda necessidade de que ele seja semanticamente interpretado”. Ou seja, admitindo sua característica de traço interpretável. Ainda para Santos (2013, p. 36):

Estando a estrutura argumental das sentenças relacionada à função sintática dos argumentos, resulta daí a dedução de que muitas vezes a animacidade pode ser refletida na maneira como são organizados esses argumentos, nas restrições ou obrigatoriedades que podem determinar sobre a sintaxe dos argumentos.

Na presente tese, discutimos o traço da animacidade por sua condição de traço semântico, mas é inegável sua relação direta com os fenômenos aqui estudados, que são de natureza sintática. Não pretendemos, como Santos (2013) fez em seu trabalho, discorrer sobre as considerações que levam alguns autores a incluírem esse traço como de natureza sintática, mas o fato de esse traço ser analisado por sua influência na atribuição de papéis *theta* nos é importante, já que a saturação dos argumentos de um predador verbal está intrinsecamente relacionada aos papéis *theta*.

Um item lexical, para Chomsky (cf. 1995), é formado por traços semânticos, fonológicos e formais. Carvalho (2012, p. 118), ao discutir os traços no programa minimalista, afirma que “o léxico é uma lista de itens lexicais (e suas propriedades idiossincráticas) que uma dada língua possui”. Carvalho afirma ainda que “traços semânticos têm a função de transmitir o conteúdo semântico do item lexical”.

Lopes (2006, p. 162) afirma que “há traços semânticos intrínsecos a itens lexicais, como animacidade, por exemplo, e interpretações semânticas computadas composicionalmente a partir de um conjunto de diferentes traços da sentença”. A autora complementa sua explicação com um exemplo bastante pertinente à nossa discussão. Segundo ela, “em menino há um traço [+ animado] intrínseco ao item lexical, mas sua interpretação como específico ou genérico, por exemplo, dependerá do predicado em que ocorrer”. Posto assim, a interpretação de “o menino” ou “um menino” depende da composição da sentença, mas o traço da animacidade, em si, é um traço de natureza semântica, intrínseco à palavra e à composição do seu significado.

Visto assim, também a partir da versão minimalista da gramática gerativa, Lopes (2006, p. 161) esclarece que “entende-se o acionamento paramétrico durante o processo de aquisição como a seleção de traços no léxico. Assume-se que há traços de duas naturezas: traços não-interpretáveis (formais) e interpretáveis (semânticos). Compreendemos, pois, que o traço da animacidade é um traço interpretável.

No trabalho de Santos (2013), acima referenciado, o autor argumenta que há implicações do traço da animacidade na estrutura sintática, e, de alguma maneira isso o inclui na categoria de traços formais. Discordamos do posicionamento de Santos (2013), e assumimos nesta tese a animacidade como um traço semântico, por compor o significado de um item lexical por sua presença ou ausência. Essa composição do significado em si nos parece suficiente para que o traço da animacidade seja um traço semântico por definição.

Por fim, vale destacar o fato de que a causatividade está intrinsecamente relacionada à animacidade. Um substantivo animado (com poder de causar uma ação) é o candidato típico a sujeito de um verbo. Inversamente, um substantivo inanimado é incapaz de causar uma ação verbal

e assim é o grande candidato a objeto da ação verbal. Aguiar (2013, p.33), ao tratar do tema, afirma que há uma “tendência natural para o sujeito/agente ser mais animado” no sentido oposto, o substantivo inanimado tende a ser objeto ou alvo da ação.

Se, conforme argumenta Santos (2013, p. 36), “a animacidade de fato deve ter influência sobre as estruturas sintáticas das sentenças, bem como deve ter importância para o estabelecimento das funções sintáticas”, isso não muda o *status* de interpretável a esse traço imputado, mas questionamos, afinal, quando a animacidade é acessada pelo *parser*? Quando ela passa a influenciar o marcador frasal a ponto de influenciar o tempo de processamento das estruturas frasais. E mais especificamente, essa influência, se houver, abrange que tipo de estruturas?

Dito isso, no próximo capítulo, procederemos a uma análise de *corpus*, em que contraporemos dados do Português Brasileiro e do Português Europeu, nas modalidades escrita e falada, com o fim de compreender se a distribuição de CRS e CRO nessas duas variedades da Língua Portuguesa é de alguma maneira influenciada pelo traço da animacidade.

Liberdade, essa palavra que o sonho humano alimenta que não há ninguém que explique e ninguém que não entenda
– Cecília Meireles, Romanceiro da Inconfidência.

CAPÍTULO 3

Estudo de *Corpus*

Neste capítulo, procederemos a uma análise de *corpora* de textos orais e escritos da língua Portuguesa nas variedades Brasileira e Europeia. O intuito desse estudo é averiguar se a distribuição de relativas de sujeito e de objeto é afetada pela animacidade do antecedente relativizado nessas estruturas. Os resultados do presente estudo serão fundamentais à discussão que faremos a partir dos dados experimentais constitutivos desta tese.

3.1 Sobre os *corpora*

É preciso que façamos alguma consideração sobre a representatividade do *corpus* aqui analisado, para os fenômenos ora discutidos. Esta tese, constituída no âmbito da psicolinguística experimental, tem suas bases teóricas estabelecidas nas discussões do capítulo anterior, e, no capítulo 4, posterior a este, faremos as análises dos dados experimentais que são o foco principal de nossas investigações.

Estabelecer o que é representativo de um dado universo linguístico é sempre uma tarefa árdua com a qual a linguística de *corpus* lida. Dito isso, reconhecemos que nosso *corpus*, ou nossos *corpora*, se estivéssemos empreendendo um trabalho no campo da linguística de *corpus*, seria, quantitativamente, pouco representativo. Entretanto, não é nossa intenção estabelecer, por meio dos *corpora* aqui levantados, parâmetros de uso das construções elencadas no estudo.

Nossa intenção foi, tão somente, traçar uma comparação entre os nossos dados com os dados apresentados para o Alemão, no estudo de Zubin (1979) e para o Alemão e o Holandês no estudo de Mak, Vonk e Schriefers (2002). Particularmente neste último estudo, como compararemos no final deste capítulo, o *corpus* levantado para essas línguas também foi constituído

por um universo pouco significativo, se tomadas as proporções da língua como um todo. Contudo, as tendências encontradas por esses autores, naquelas línguas, foram também encontradas nas duas variedades da língua Portuguesa, aqui estudadas. Dito isso, é na condição de tendência de uso que tomaremos os dados discutidos neste capítulo, para, a partir deles, compreender os fenômenos verificados nos dados experimentais, verdadeiro foco de nossas motivações investigativas.

Uma vez que nos estudos experimentais, que estarão devidamente reportados e discutidos no próximo capítulo, analisamos as variedades Brasileira e Europeia da língua Portuguesa, faz-se necessário que o estudo de *corpus* também se dê nessas duas modalidades.

Por considerarmos ainda que os fenômenos linguísticos, muitas vezes, possuem comportamentos distintos nas modalidades oral e escrita, optamos por constituir nossos *corpora* a partir das duas modalidades. Apresentaremos os resultados dos dados, respectivamente, da língua Portuguesa Brasileira e Europeia .

No capítulo anterior, referimo-nos a Zubin (1979), que estudou se há um efeito da animacidade na distribuição de cláusulas relativas em que os protagonistas diferem em animacidade. Em seu trabalho, Zubin estabeleceu que regularidades na produção da língua são conduzidas pela saliência e pela animacidade. Uma entidade que é mais saliente terá mais chances de ser o sujeito da sentença. Um dos fatores que influencia a saliência é a disponibilidade do termo em questão.

Em cláusulas relativas, o antecedente está disponível no primeiro momento, e o termo dentro da cláusula relativa é novo. Observe-se a estrutura que se segue:

43. A menina que ganhou o livro.

Em 37, “a menina” é o termo antecedente, que será retomado pelo pronome relativo. Segundo os pressupostos de Zubin (1979), é o termo mais saliente e por isso candidato natural a argumento externo da cláusula relativa. Por sua vez, “o livro” é o termo novo, menos saliente e por isso com menor probabilidade de ser o sujeito da sentença.

Assim, o antecedente é mais saliente, e, portanto, provavelmente será usado como o sujeito da cláusula relativa. Além disso, um termo animado será preferido em relação a um termo inanimado como o sujeito de uma sentença. Como resultado, há duas forças atuantes na produção de cláusulas relativas em que os protagonistas diferem em animacidade: a primeira é a tendência a usar o substantivo antecedente como sujeito da cláusula relativa, e a segunda é a tendência a usar uma entidade animada como sujeito da cláusula relativa.

Zubin (1979) encontrou evidência para esta consideração em um estudo de *corpus* de cláusulas relativas em Alemão com uma entidade animada e outra inanimada. Ele dividiu as cláusulas relativas do seu *corpus* em quatro categorias, a saber: sujeito animado, sujeito inanimado, objeto animado e objeto inanimado.

Partindo dos pressupostos de Zubin (1979), Mak, Vonk e Schriefers (2002) também realizaram um estudo de *corpus* do Alemão e do Holandês, estudo que não procurou replicar o estudo de Zubin, mas ampliar seu escopo.

Esses estudos de *corpus*, conforme discutimos na seção 2.3, do capítulo anterior, e conforme indicado na Tabela 1 daquela seção, demonstraram que a animacidade influencia a distribuição de cláusulas relativas. Além disso, os autores mostraram que as cláusulas relativas de objeto animado que foram usadas em muitos estudos de compreensão e produção de relativas anteriores são muito pouco frequentes.

A partir desses fatos, empreendemos um estudo de *corpus* das variedades Brasileira e Europeia da língua Portuguesa, com o fim de averiguar se a distribuição de CRS e CRO também é influenciada pela animacidade no Português, ou melhor dizendo, qual é a tendência de distribuição de relativas de sujeito e de objeto quando considerado o traço da animacidade do termo antecedente dessas relativas.

3.2 Português Brasileiro

3.2.1 *Corpus* escrito – PB

O *corpus* escrito do Português Brasileiro foi constituído por 69 textos escritos, entre crônicas, notícias, artigos de opinião e entrevistas. Sendo 39 do

jornal Tribuna do Norte, da cidade de Natal, no Rio Grande do Norte, e 30 da Revista Época, de circulação nacional.

Os textos do Jornal Tribuna do Norte foram todos veiculados entre 31 de agosto e 06 de setembro de 2014 e estão disponíveis em: <http://www.tribunadonorte.com.br/>.

Os textos da Revista Época foram todos veiculados entre 20 de março e 25 de dezembro de 2015 e estão disponíveis em: <http://epoca.globo.com/>.

O *corpus* escrito totalizou 36.962 palavras, a partir dele procedemos a uma análise e consideramos apenas as relativas com verbos transitivos. Havia, nessas condições, 411 relativas. Dentre as quais, 117 eram de sujeito animado, 222 de sujeito inanimado, 1 de objeto animado e 71 de objeto inanimado, conforme o gráfico¹.

Fizemos, por meio do *Action*²⁵, um teste qui-quadrado e encontramos uma diferença significativa de ocorrência entre essas estruturas no *corpus* do PB escrito ($\chi^2=30,23$, $p<0,001$).

Corpus escrito - PB

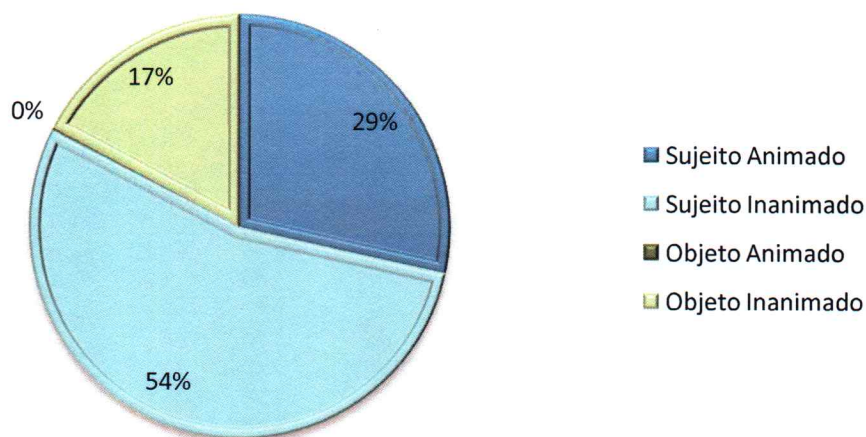


Figura 3: Gráfico representativo do *corpus* escrito do PB

²⁵ O Action é um sistema estatístico que conecta o R, um programa estatístico, com o Excel, por meio dele temos que fazer uma série de teste estatísticos.

3.2.2 *Corpus* falado – PB

O *corpus* falado foi constituído por 03 depoimentos espontâneos coletados no Brasil no ano de 1993. Os depoimentos transcritos estão disponíveis em: <http://www.discursoeagramatica.lettras.ufrj.br/download/natal.pdf> e foram organizados pela professora Maria Angélica Furtado da Cunha. Todos os depoimentos foram de falantes nativos do Português Brasileiro.

O *corpus* foi composto por 36.716 palavras, consideramos os mesmos critérios do *corpus* escrito e encontramos um total de 452 relativas, sendo 113 de sujeito animado, 177 de sujeito inanimado, 9 de objeto animado e 153 de objeto inanimado, conforme reportado no gráfico 2.

Fizemos, também, um teste qui-quadrado e encontramos uma diferença significativa de ocorrência entre essas estruturas no *corpus* do PB falado, ($\chi^2=57,19$, $p<0,001$).

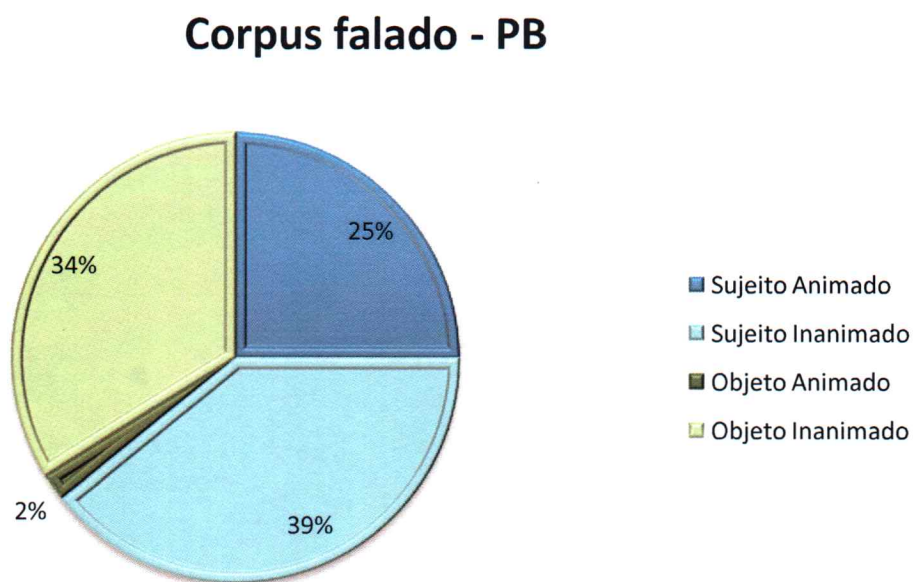


Figura 4: Gráfico representativo do *corpus* falado do PB

Não há grande discrepância entre os números da modalidade escrita para os números da modalidade oral, mas, claro, precisamos considerar as diferenças em nossa análise. Por uma questão didática, optamos por fazer o

gráfico da Figura 5 em que contabilizamos os dois *corpora*, o escrito e o falado, constituindo, pois, o nosso *corpus* escrito e falado da modalidade Brasileira da língua Portuguesa para o fim de nossas análises.

A soma das relativas dos dois *corpora* totalizou **863** relativas, sendo **411** do *corpus* escrito e **452** do *corpus* falado. Foram contabilizadas apenas as relativas de sujeito e de objeto com o pronome “que”. Outros pronomes relativos, assim como outras funções não foram contabilizados. No *corpus* falado havia exemplos de relativas cortadoras e resumptivas que também foram contabilizadas. Os resultados estão na Figura 5, abaixo:

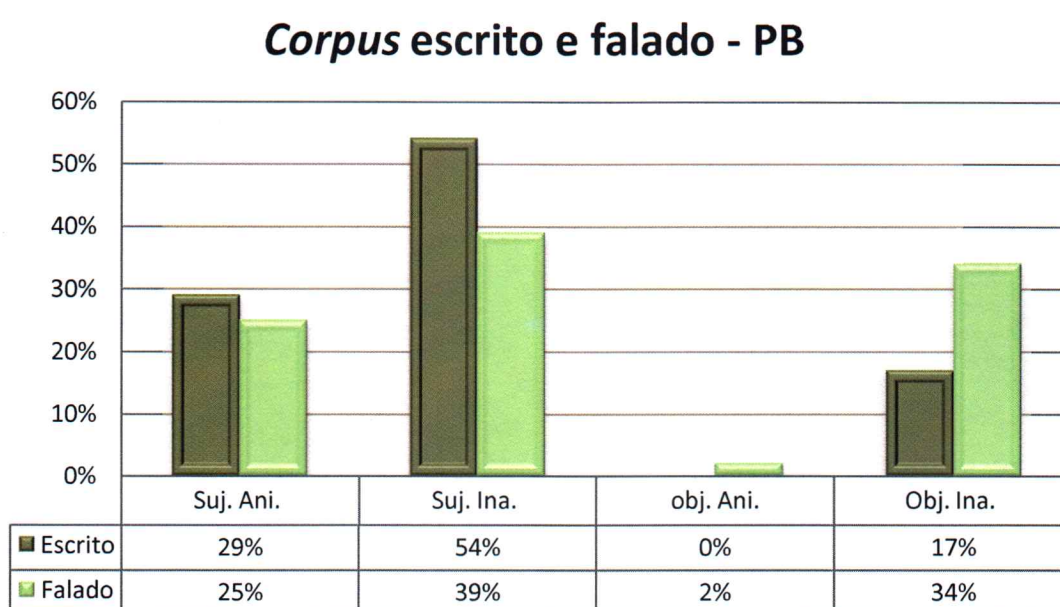


Figura 5: Gráfico representativo do *corpus* escrito e falado do PB

Como se pode observar, as maiores diferenças entre as duas modalidades do PB aqui encontradas estão no sujeito inanimado, 15 pontos percentuais; e no objeto inanimado, 17 pontos percentuais. Parece ter havido um crescimento da quantidade de objetos inanimados na oralidade em detrimento de uma diminuição de sujeitos inanimados na mesma modalidade.

As outras condições permaneceram relativamente estáveis na comparação entre a modalidade oral e escrita em PB.

Cabe apontar ainda a raridade de ocorrência de objetos animados, essa frequência consideravelmente baixa, também encontrada por Zubin (1979) e

Mak, Vonk e Schriefers (2002), será de fundamental importância para a construção de nossos argumentos nesta tese.

Antes, porém, de tecermos algumas considerações importantes à nossa discussão, traremos os dados do *corpus* da variedade Europeia do Português na seção que se segue.

3.3 Português Europeu

3.3.1 *Corpus* escrito – PE

O *corpus* escrito do Português Europeu foi constituído por 28 artigos de opinião do Jornal de Notícias *on-line* de Portugal. Os textos foram acessados no dia 7 de fevereiro de 2016 na página <http://www.jn.pt/paginainicial/>.

Os textos do *corpus* totalizaram 10.412 palavras. Obedecendo aos mesmos critérios do estudo de *corpus* do Português Brasileiro, encontramos um total de 161 relativas, dentre elas, 37 com sujeito animado, 86 com sujeito inanimado, 2 com objeto animado e 36 com objeto inanimado, conforme disposto no gráfico da Figura 6, logo abaixo.

Aqui também fizemos um teste qui-quadrado e encontramos uma diferença significativa de ocorrência entre essas estruturas no *corpus* do PE escrito, ($\chi^2=8,43$, $p<0,003$).

Corpus escrito - PE

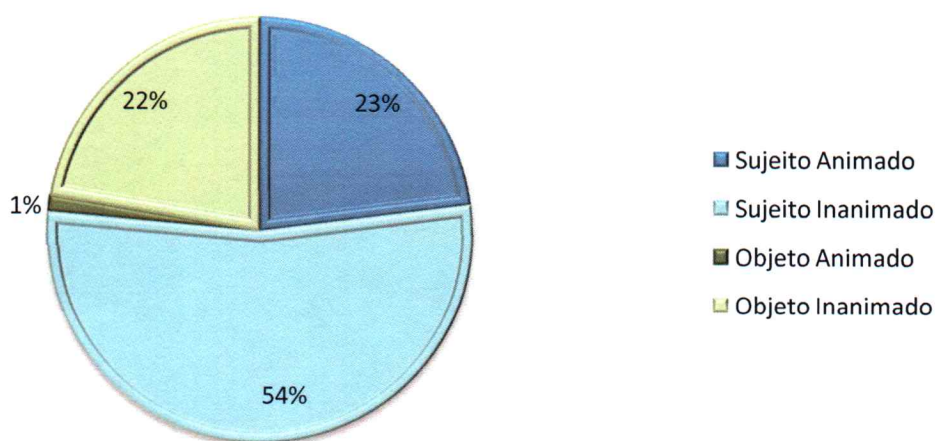


Figura 6: Gráfico representativo do *corpus* escrito do PE

3.3.2 *Corpus* Falado – PE

O *corpus* falado constitui-se de 12 depoimentos espontâneos coletados em Portugal entre os anos de 1989 e 1997. Os depoimentos transcritos estão disponíveis em <http://www.clul.ul.pt/pt/recursos/183-reference-corpus-of-contemporary-portuguese-crpc> e fazem parte do *Corpus* de Referência do Português contemporâneo da Universidade de Lisboa. Todos os depoimentos foram de falantes nativos do Português Europeu.

O *corpus* totalizou 14.139 palavras, e, também obedecendo aos mesmos critérios de seleção das análises anteriores, encontramos 178 relativas, sendo de 54 sujeito animado, 80 de sujeito inanimado, 1 de objeto animado e 43 de objeto inanimado, conforme demonstra o gráfico da Figura 7.

Por fim, Fizemos neste *corpus* também um teste qui-quadrado e encontramos uma diferença significativa de ocorrência entre essas estruturas no *corpus* do PB escrito, ($\chi^2=30,25$, $p<0,001$).

***Corpus* falado- PE**

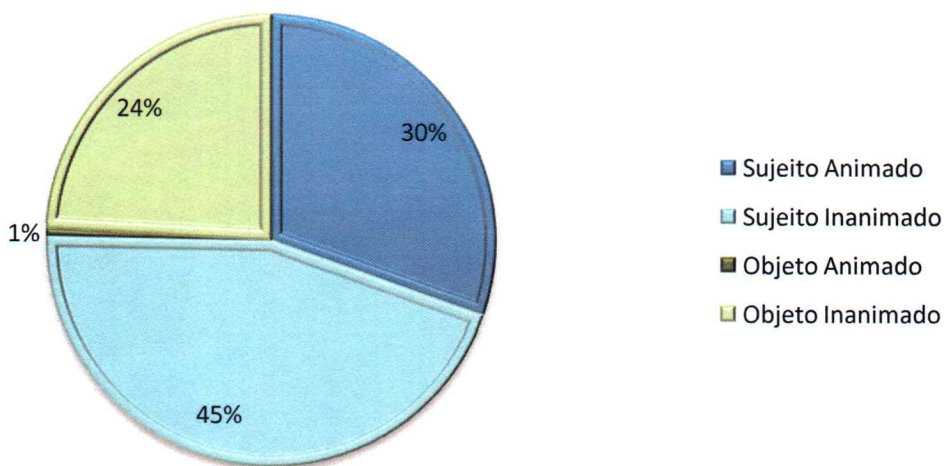


Figura 7: Gráfico representativo do *corpus* falado do PE

A soma das relativas dos dois *corpora* totalizou **339** relativas. Foram contabilizadas apenas as relativas de sujeito e de objeto com o pronome “que”. Outros pronomes relativos, assim como outras funções, não foram contabilizados. Os *corpora* escrito e falado totalizaram 24.551 palavras, e

também obedecendo aos mesmos critérios de seleção das análises anteriores, encontramos 339 relativas. Sendo 91 de sujeito animado, 163 de sujeito inanimado, 3 de objeto animado e 79 de objeto inanimado, conforme representamos no gráfico 6, abaixo.

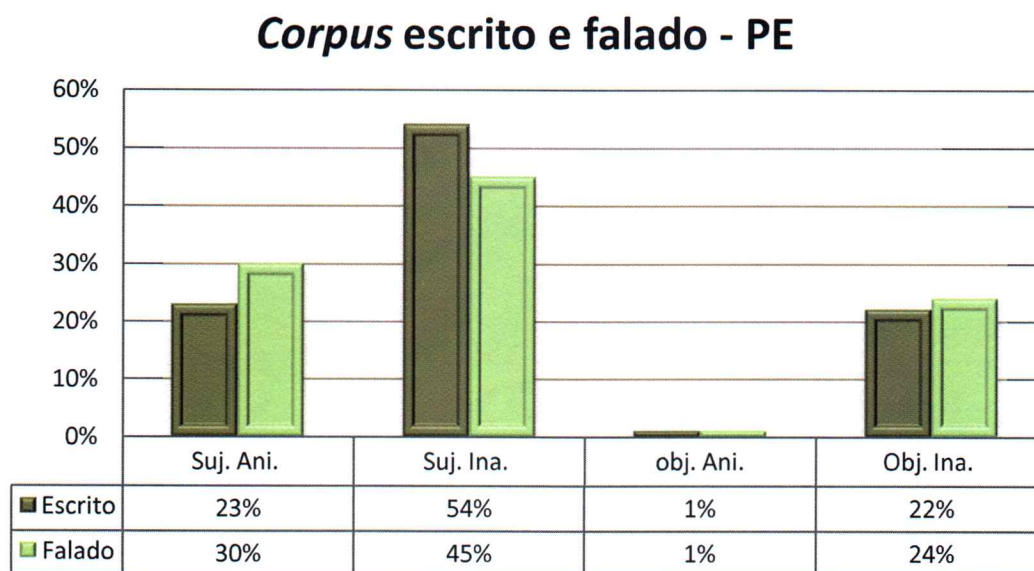


Figura 8: Gráfico representativo do *corpus* escrito e falado do PE

3.4 Comparação PB/PE

As diferenças entre as duas modalidades em PE estão distribuídas entre duas das condições analisadas. Há uma diferença de 7 pontos percentuais nas relativas de sujeito animado, de 9 pontos percentuais no sujeito inanimado.

As outras condições também permaneceram relativamente estáveis na comparação entre a modalidade oral e escrita em PE. Assim como no PB, nas duas modalidades da variedade Europeia da língua as estruturas de objeto animado são extremamente raras.

Por fim, na Tabela 7, abaixo, fazemos uma comparação entre todas as condições elencadas no estudo. Ou seja, reunimos estruturas de sujeito e objeto, com antecedentes animados e inanimados nas modalidades escrita e oral das modalidades Brasileira e Europeia da língua Portuguesa.

Tabela 7: *Corpora* – escrito e falado, em PB e PE.

Corpora escrito e falado								
Relativa	Suj. Ani.		Suj. Ina.		Obj. Ani.		Obj. Ina.	
Modalidade	Escrito	Falado	Escrito	Falado	Escrito	Falado	Escrito	Falado
PB	29%	35%	54%	39%	0%	2%	17%	34%
PE	23%	30%	45%	45%	1%	1%	22%	24%

Percentual de ocorrência de estruturas relativas animadas e inanimadas em PB e BE

Nessa Tabela, apresentam-se as diferenças entre as duas variedades da língua Portuguesa aqui estudadas. Como se pode perceber, de maneira geral nas duas variedades linguísticas apresentadas, ao menos no que tange ao *corpus* trabalhado, as estruturas variam de maneira relativamente uniforme em PB e PE, apresentando tendências semelhantes.

O índice de ocorrência das estruturas que varia entre uma variedade e outra da língua Portuguesa, como também entre as modalidades escrita ou falada, mas o mais relevante é apontar para o fato de que os testes de qui-quadrado apontaram, que, entre uma estrutura e outra, há diferenças significativas, conforme aponta o p-valor da variedade Brasileira do Português na modalidade escrita: ($\chi^2=30,23$, $p<0,001$); na modalidade falada: ($\chi^2=57,19$, $p<0,001$); na variedade Europeia na modalidade escrita: ($\chi^2=8,43$, $p<0,003$); e na modalidade falada: ($\chi^2=30,25$, $p<0,001$).

Na seção 2.3, transcrevemos, na Tabela 1, os números e porcentagens trazidos no estudo de *corpus* realizado por Mak, Vonk e Schiefers (2002). Na Tabela abaixo, trazemos os mesmos números, acrescentamos, porém, os números encontrados na presente pesquisa, nas duas variedades e modalidades pesquisadas, para uma comparação entre as línguas.

As estruturas mais comuns do Português foram as CRS inanimado, já em Holandês e Alemão foram as CRS animado. No mais, nas três línguas elencadas na Tabela que se segue, é preciso que ressaltemos que a raridade das estruturas de objeto animado é explícita, do que se pode depreender a anti-naturalidade da leitura de sentenças que as contenham, e isso pode acarretar diferenças no custo de processamento. Ao que mais adiante comentaremos.

Tabela 8: *Corpora* do Holandês, do Alemão e do Português

Relativa	Holandês	Alemão	Português
Sujeito Animado	140(49%)	82(49%)	321(27%)
Sujeito Inanimado	66(23%)	62(37%)	565(47%)
Objeto Animado	02 (1%)	2 (1%)	13(1%)
Objeto Inanimado	78 (27%)	22(13%)	303(25%)
Relativas em números absolutos dos <i>Corpora</i> – Porcentagem entre parêntesis.			

O estudo de *corpus* do Português, por nós realizado, conforme já apontamos anteriormente, foi instaurado com o fim de ser tomado como base nas discussões que faremos no próximo capítulo. Em consonância com Kenedy (2009, p.33), “as análises de *corpus* apresentam limitado poder explanatório para as pesquisas que pretendem fazer generalizações sobre a competência linguística dos falantes”, contudo, para esta pesquisa, as considerações feitas a partir dos dados que coletamos serão parte da argumentação que pretende descrever e explicar os fenômenos que ora investigamos. Alguns dos questionamentos feitos por Kenedy (2009) estão em torno da constituição do *corpus*, da variedade de textos, do perfil sociocultural dos produtores de tais textos, entre outros fatores.

Noutras palavras, a constituição de um *corpus* verdadeiramente representativo, inclusive quando fazemos um trabalho nos moldes desta tese, em que comparamos variedades linguísticas de um mesmo idioma, pode ser algo delicada e não representar de fato o que ocorre na língua em termos de processamento. Para Kenedy (2009, p. 33), a linguística experimental pode dar conta dessa questão, uma vez que esse tipo de pesquisa pode, por exemplo, permitir “a comparação controlada de falantes do PB e do PE que possuem perfis socioculturais equivalentes”.

Perceba-se, esse controle do perfil dos participantes que efetivamente produziram os textos dos *corpora* analisados não aconteceu na nossa pesquisa, e seria muito difícil de ser feito com o rigor que se espera quando temos medidas muitas vezes minuciosas. Na pesquisa experimental, porém, o controle do perfil dos participantes pode se dar de maneira muito mais efetiva, como fizemos e pretendemos demonstrar no próximo capítulo.

E como poderá uma sociedade de cegos organizar-se para que viva, Organizando-se, organizar-se já é, de uma certa maneira, começar a ter olhos.

– José Saramago, Ensaio sobre a cegueira.

CAPÍTULO 4

Estudos Experimentais

O processamento de frases encontra, na psicolinguística experimental, diversos métodos distintos de análise, com arcabouço teórico-metodológicos distintos entre si, mas que complementam de alguma forma a visão sobre fenômenos naturais envoltos em mecanismos bastante complexos.

A linguística experimental, e mais especificamente a psicolinguística, tem encontrado respostas interessantes que revelam, paulatinamente, como os mecanismos da mente humana agem ao processar a linguagem.

Nesta tese, apropriamo-nos de metodologias experimentais, com a pretensão de discutir o processamento de relativas de sujeito e de objeto sob a influência da animacidade. Para tanto, elaboramos e aplicamos 5 experimentos a partir dos quais pretendemos explicar algumas questões que circundam o processamento dessas estruturas.

Organizamos os 5 experimentos em três seções, e a discussão em torno deles em outras duas, organizadas da seguinte forma:

Na primeira seção, 4.1, apresentamos dois experimentos, que nomearemos de experimento 1 e 2. O primeiro deles em Português Brasileiro e o outro, na realidade sua réplica, em Português Europeu. Cada um desses experimentos possui duas condições experimentais. São sentenças com cláusulas relativas encaixadas com extração de sujeito e de objeto. Todas as frases experimentais possuem termos antecedentes relativizados animados. A variável independente é, portanto, o tipo de relativa, se de sujeito (CRS), ou se de objeto (CRO). A variável dependente é o tempo de processamento do segmento crítico, S2. Isso na aferição do processamento *on-line*, uma vez que aferimos também o tempo de resposta e o índice de acertos à questão proposta, o que nos dá duas medidas *off-line*. As frases experimentais foram feitas com nomes próprios, isso lhes dá o caráter de cláusulas relativas explicativas, o que ficará explicado na descrição do experimento.

Na segunda seção, **4.2**, mais dois experimentos serão apresentados, são os experimentos 3 e 4. O experimento 3 foi feito em PB e o experimento 4 é uma réplica dele em PE. Assim, as variáveis independentes são o tipo de relativa e a animacidade; temos, pois, experimentos com 4 condições experimentais, são elas: CRSa, CRSi, CROa e CROi. Nesses experimentos, as cláusulas não serão mais explicativas, mas restritivas, na descrição dos experimentos isso será devidamente esclarecido.

Na terceira seção, **4.3**, apresentaremos o último experimento, feito apenas em PB. O experimento de Leitura Automonitorada, feito a partir de palavras isoladas. Assim, temos substantivos em duas condições experimentais, a saber: animados e inanimados.

Na quarta seção, **4.4**, fazemos uma análise geral dos dados encontrados com considerações com base no modelo teórico assumido para tal, a DLT.

Por fim, na seção **4.5**, discutiremos alguns modelos teóricos de processamento e o modo com que eles concebem nosso objeto de estudo.

Diante da discussão teórica que empreendemos no segundo capítulo da presente tese e dos dados elencados no estudo de *corpus* do terceiro capítulo, tomaremos para análise os estudos experimentais que constituem o foco principal desta tese.

Para nossa empreitada, elaboramos e aplicamos os experimentos supracitados em PB e PE e pretendemos, nas próximas seções, apresentar a metodologia de trabalho que adotamos e os resultados dos referidos experimentos.

Apresentaremos, por questões de cunho didático, os experimentos que primeiro se fizerem interessantes à discussão, independente da ordem cronológica em que foram realizados. Nas próximas seções, conforme já afirmamos, pretendemos apresentar os 5 experimentos por nós realizados no decorrer dessa investigação em torno das CRS e das CRO.

A nossa opção metodológica nos 5 estudos experimentais foi pela Técnica de Leitura Automonitorada. A partir dessa técnica temos condições de fazer medidas *on-line*, ou seja, medidas que são realizadas no momento em que o processamento de uma estrutura acontece. Além dos experimentos com Leitura Automonitorada a partir de frases com perguntas do tipo SIM/NÃO no

final, também fizemos um experimento Leitura Automonitorada de palavras isoladas.

Não podemos deixar de considerar que a Técnica Experimental de Leitura Automonitorada recebe algumas críticas por parte de alguns psicolinguistas que procuram aferir o processamento *on-line*. Tais críticas são no sentido de haver outras Técnicas Experimentais, a exemplo da Técnica de Rastreamento Ocular, que pode ser mais precisa no aferimento de tempos de processamento. Não refutamos essas críticas, sabemos de sua pertinência. Contudo, por meio da Leitura Automonitorada, alcançamos resultados interessantes e robustos que nos pareceram suficientes às explicações dos fenômenos aqui investigados. Tais resultados podem ser ampliados a partir da aplicação de da Técnica de Rastreamento Ocular e por isso mesmo o empreendimento de futuras pesquisas por meio dessa técnica é um caminho a ser seguido após a análise ora em curso.

4.1 Experimentos 1 e 2 – Cláusulas relativas de sujeito e objeto com antecedentes animados em PB e PE

Como temos afirmado desde o início deste texto, a maior parte dos trabalhos cujo foco era a assimetria entre CRS e CRO analisou o processamento de relativas em que os antecedentes retomados na cláusula relativa eram termos animados.

Também a maior parte dessas pesquisas não controlou o traço da animacidade. Ou seja, embora tenhamos encontrado termos animados nos exemplos dados no corpo desses trabalhos a que nos referimos, nem sempre se controlou, sistematicamente, o traço da animacidade.

Nossos argumentos, conforme ficará mais claro neste capítulo, partem da hipótese de que há uma influência desse traço no processamento de relativas. Já parece estabelecido, até o presente momento, que as relativas de sujeito são estruturas mais fáceis de processar em diversas línguas. Advogamos a ideia de que o fato de as estruturas aferidas serem, se não na totalidade, mas pelo menos na maior parte, animadas, pode ter enviesado os resultados de tais estudos.

Para uma visão mais apropriada do fato, decidimos começar nossos estudos experimentais por uma aferição sistemática disso. Controlamos, rigorosamente, o traço da animacidade. Construimos estruturas frasais com termos antecedentes sempre animados para que pudéssemos verificar se, nessas condições, tanto em PB quando em PE, a assimetria entre CRS e CRO se mantém.

Muitos fatores podem concorrer para um processamento mais rápido ou mais lento de uma sentença. Um desses fatores está relacionado à métrica da estrutura em si, o tamanho da palavra ou da sequência de palavras medidas no experimento. Sobre isso, nos deteremos com o devido afinho quando tratarmos das condições experimentais de cada experimento.

Outro fator que pode interferir no tempo de processamento de uma sentença é a frequência de uso da palavra, para tanto, nestes primeiro e segundo experimentos, nossa estratégia foi de usar nomes próprios, com isso resolvemos duas questões relacionadas à nossa pesquisa. Sendo nomes próprios os termos antecedentes relativizados, garantimos que todos os termos fossem não só animados, mas ainda humanos. Para que a frequência de uso dos nomes próprios escolhidos não interferisse nos resultados, no Brasil, procuramos selecionar todos os nomes entre os 50 nomes mais comuns do país com base no Cadastro de Pessoa Física – CPF²⁶ – e em Portugal, procuramos em sites especializados²⁷ os nomes mais usados nos cartório de registro de nascimento nas duas últimas décadas.

4.1.1 Técnica Experimental de Leitura Automonitorada

Na Técnica de Leitura Automonitorada, os participantes da pesquisa são convidados a ler frases na tela de um computador e, na sequência, respondem a uma pergunta simples sobre a frase do tipo SIM/NÃO.

²⁶ Os nomes mais comuns no Brasil estão disponíveis nos seguintes endereços: <http://respostasdastrivias.blogspot.com.br/2014/01/50-nomes-mais-populares-do-brasil.html>, <http://g1.globo.com/jornal-da-globo/noticia/2011/11/confira-lista-completa-com-os-50-nomes-mais-populares-do-brasil.html>.

²⁷ Os nomes mais comuns em Portugal estão disponíveis nos seguintes endereços: <http://nomesportugueses.blogspot.com.br/2015/01/nomes-populares-em-portugal-top-100-de.html>.

As frases não são apresentadas integralmente, mas segmentadas, também de acordo com os propósitos de cada experimento. O fenômeno analisado, geralmente, leva a diferenças de tempo de processamento de um desses segmentos. Quanto maior o tempo de leitura, provavelmente, maior será a dificuldade em torno da estrutura.

Contudo, além do segmento crítico, medimos também o pós-crítico, a fim de averiguar a possível ocorrência de um efeito *spillover*, que, segundo Farias, Leitão e Ferrari-Neto (2012, p. 101), é “definido como sendo a captação de um efeito presente em um dado segmento que se reflete nos segmentos seguintes que o sucedem na sequência da sentença”.

Como dito, a sentença aparece fragmentada, o participante deve acionar a tecla “espaço” do computador, de maneira a aparecer o primeiro segmento, ao apertar pela segunda vez, esse segmento desaparece da tela dando lugar ao segundo e assim sucessivamente até o último segmento. No final, aparece uma pergunta interpretativa a respeito da frase lida, a que o participante deve responder.

Admitimos, pois, que a rapidez da resposta está diretamente relacionada com o nível de dificuldade encontrado. Quanto mais rápida é a leitura, mais fácil terá sido o processamento, e o contrário também é verdadeiro. Ao final, a média dos tempos de leitura do segmento crítico será contabilizada a fim de conseguirmos inferir como se deu o processamento *on-line* das estruturas investigadas.

Para os dois experimentos de que tratamos nesta seção, as sentenças foram divididas em 5 segmentos, como os que se seguem:

44. João / **que encontrou José** / naquele dia / não quis / pedir perdão.
 S1 S2 S3 S4 S5
45. João / **que José encontrou** / naquele dia / não quis / pedir perdão.
 S1 S2 S3 S4 S5

O segundo segmento, S2, em negrito, contém a cláusula relativa encaixada²⁸, é, portando, o segmento crítico. Em 44, temos uma CRS e em 45, temos uma CRO. Como são nomes próprios, portanto nomes que possuem o

²⁸ Consideremos aqui a definição de encaixamento e Gibson (2001), segundo esse autor, numa livre tradução, uma categoria sintática *A* é dita sendo encaixada dentro de outra categoria *B* se *B* contém *A*, um elemento à esquerda de *A*, e um elemento à direita de *A*.

traço animado, esperamos que o tempo de processamento de S2 seja mais rápido em 44 do que em 45.

É preciso que façamos algumas considerações importantes quando à escolha de nomes próprios como termos relativizados. A leitura de 44 ou 45, certamente, será mais natural se considerada como relativa explicativa. A implicação dessa consideração é que essas estruturas não são idênticas às estruturas que foram testadas nos vários estudos anteriores aqui citados. Tais estudos partiram de relativas restritivas, mas nossa hipótese é de que isso não acarretará diferença entre este e os experimentos citados anteriormente. Ou seja, também as explicativas apresentarão a assimetria dos outros estudos, já que os termos antecedentes das relativas de nosso experimento também são todos animados. Por fim, cabe ressaltar que, ao nos valermos de relativas explicativas, se encontrarmos os mesmos resultados, ou seja, se CRS obtiver média de tempo de leitura menor do que CRO, teremos ampliado os resultados antes encontrados.

As explicações que a literatura especializada vem atribuindo a essa maior facilidade nas estruturas de sujeito do que nas de objeto orbitam pressupostos de ordem sintática. Oliveira (2013, p. 77), por exemplo, afirma que “a extração de sujeito é sempre mais fácil de processar”. Segundo esse autor, “esse tipo de extração descreve uma cadeia linearmente e estruturalmente menor”.

Costa et al. (2009, 2012) argumentam que “o papel temático do objecto atravessa o papel temático do sujeito.” Outros autores endossam essas afirmações, que, aliás, são pertinentes e possuem comprovações empíricas consistentes. Não obstante a pertinência dessas afirmações, que são de ordem sintática, acreditamos que a animacidade do termo antecedente, relativizado na cláusula encaixada, pode exercer influência no processamento dessas sentenças. Se isso for verdadeiro, o encapsulamento sintático de que trata, por exemplo, a teoria do *Garden-Path*, precisa ser discutido, pelo menos no que tange a essas estruturas relativas aqui investigadas, uma vez que a animacidade é um traço semântico.

Por ora, não aferiremos o traço da animacidade, já que todas as frases experimentais trazem o traço animado em seus antecedentes, assim a animacidade não é uma variável independente neste experimento.

Como todas as construções deste experimento carregam o traço positivo da animacidade, como aconteceu na maioria dos experimentos que testaram essa assimetria, nossa expectativa é que as CRS serão lidas mais rapidamente do que as CRO em construções explicativas, repetindo-se o que foi encontrado nas restritivas.

Conforme já apontamos no estudo de *corpus*, com as duas variedades do Português aqui estudadas, tanto em PB quanto em PE, estruturas relativas de objeto animado são extremamente raras, correspondendo a cerca de 1% das ocorrências nas modalidades oral e escrita. Dito isso, é de se supor algum estranhamento por parte do leitor quando da leitura de CRO com antecedente animado. Sobre isso, ainda não teceremos maiores considerações, uma vez que nos experimentos 3 e 4 teremos dados mais detalhados para tratar dessa questão.

Como já foi posto, o experimento foi feito no Brasil e replicado em Portugal. As diferenças entre um e outro experimento se deram apenas nas diferenças dialetais existentes entre as duas variedades linguísticas. Ou seja, toda a programação do experimento foi idêntica, apenas modificamos parte das frases experimentais, modificações mínimas, porém necessárias para que os participantes pudessem ler as frases o mais naturalmente possível. Alguns dos nomes próprios também foram diferentes obedecendo aos critérios já descritos no início deste capítulo.

Os participantes, todos estudantes de nível técnico ou superior, foram submetidos a um questionário em que pudemos constatar entre eles um nível de educação formal semelhante, além de garantir terem a língua Portuguesa, nas suas respectivas variedades, Brasileira ou Europeia, como língua materna. Para Kenedy (2009, p. 42), “a observação desse perfil é necessária para que desempenhos assimétricos na tarefa não possam ser atribuídos a diferenças no nível de letramento/escolarização dos sujeitos”.

4.1.2 Participantes do PB

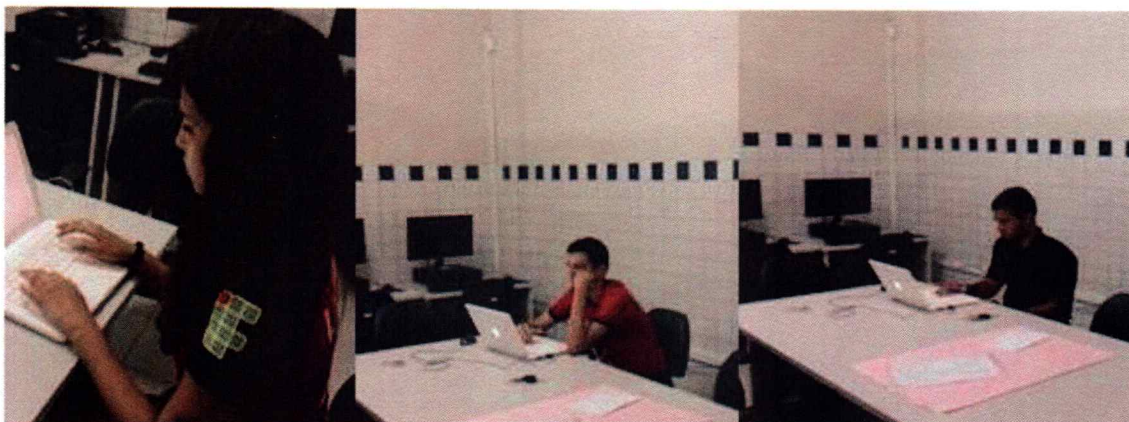


Figura 9: Participantes voluntários da pesquisa durante a realização do experimento, alunos do Instituto Federal do Rio Grande do Norte, em Natal, RN, Brasil.

Os participantes do primeiro experimento são todos falantes do PB, foram selecionados 30 estudantes do ensino tecnológico, sendo 13 homens e 17 mulheres, com idade média de 18,4 anos. Todos eles são alunos do Instituto Federal do Rio Grande do Norte, do *campus* Natal Central, na cidade de Natal. A pesquisa aconteceu com a colaboração da professora Maria Tânia Florentino de Sena Nascimento, que nos cedeu o espaço necessário na referida instituição e ainda nos auxiliou na seleção dos alunos que participaram da pesquisa.

4.1.3 Participantes do PE



Figura 10: Participantes voluntários da pesquisa durante a realização do experimento, alunos da Universidade do Minho, em Braga, Portugal.

Na Universidade do Minho, na cidade de Braga, Portugal, contamos com a colaboração da professora Maria do Carmo Lourenço-Gomes, que não só nos cedeu o espaço para realização do experimento e auxiliou na seleção dos

alunos participantes, como ainda nos orientou na programação do experimento, que, dadas as diferenças das duas variedades linguísticas investigadas, precisou se adequar ao PE.

Os participantes do segundo experimento são todos falantes nativos do PE, foram selecionados 32 estudantes entre alunos de graduação e pós-graduação da Universidade do Minho, sendo 13 homens e 19 mulheres, com idade média de 22,1 anos.

4.1.4 Materiais

O experimento foi composto por 12 frases experimentais e 24 distratoras. Entre as experimentais, 6 eram CRS e 6 CRO. Desta maneira, cada participante foi exposto 6 vezes às duas condições do teste. Para que um participante não tivesse acesso à mesma frase na versão CRS e CRO, adotamos um *design* do tipo quadrado latino.

As 24 frases distratoras tiveram a função de impedir que o participante percebesse um dado padrão nas estruturas frasais e com isso formulasse algum tipo de estratégia de leitura enviesando os resultados do experimento.

Todas as frases experimentais deveriam receber um “sim” como resposta à pergunta subsequente. Balanceamos para que metade das respostas fosse “sim” e metade fosse “não” com as frases distratoras. Dessa forma, cada participante foi exposto a 36 frases, entre distratoras e experimentais.

Foi controlada a frequência, a animacidade e o tamanho dos itens investigados. O segmento crítico analisado foi o segundo (S2), contendo um pronome relativo (que), um verbo trissílabo na terceira pessoa do singular, um nome próprio na condição de sujeito ou de objeto, conforme a frase. Os nomes próprios, no PB, conforme já citado, foram escolhidos entre os 50 nomes mais comuns no Brasil segundo o Cadastro de Pessoas Físicas (CPF), dos quais foram retirados apenas os dissílabos. No PE, a partir de sites especializados, tivemos acesso aos nomes mais comuns registrados em cartórios nas últimas duas décadas.

A expectativa era de que o segmento crítico (S2) de estruturas como 44, acima, fosse lido mais rapidamente que de estruturas como 45. Também

mensuramos o segmento seguinte (S3), a fim de averiguar a possível ocorrência de um efeito *spillover*. Para tanto, também foi controlado o tamanho do segmento 3.

4.1.5 Procedimentos

Os sujeitos foram instruídos individualmente sobre as tarefas constitutivas do experimento. As instruções foram apresentadas oralmente pelo experimentador, e, posteriormente, o participante lia as tarefas na tela do computador em que o experimento foi rodado.

Depois de ter declarado que havia entendido a tarefa, o participante se submetia a um treinamento prático constituído de 4 frases distratoras, com o mesmo número de segmentos das frases experimentais. Com isso, asseguramos que os participantes compreenderam perfeitamente as tarefas a serem realizadas, que só eram iniciadas quando do consentimento do participante que demonstrava ter de fato compreendido seu papel frente às tarefas experimentais.

Ao pressionar a tecla “espaço”, aparecia o primeiro segmento no centro da tela do computador, apertando novamente aparecia o segundo segmento, no lugar do primeiro, que, por sua vez, desaparecia. O procedimento deveria ser repetido por cinco vezes, até aparecer a pergunta a que ele tinha que responder acionando a tecla “sim”, devidamente adaptada no teclado do computador, na cor verde, ou “não”, destacada no teclado na cor vermelha.

A fonte das sentenças lidas aparecia na cor preta na tela do computador, ao passo que a pergunta sobre a sentença aparecia na cor vermelha.

No Brasil, o experimento foi realizado no mês de outubro de 2014, conforme já indicamos, no Instituto Federal do Rio Grande do Norte, *campus* Natal Central, em Natal, RN. Em Portugal, realizamos o experimento no mês de outubro desse mesmo ano na Universidade do Minho, na cidade de Braga.

O experimento foi rodado em um MacBook, com uma tela de 13”. G3, 233Mhz. Utilizamos o programa *Psyscope* (COHEN; MACWHINNEY; FLATT & PROVOST, 1993). As frases foram randomizadas e o tempo de leitura foi aferido em milésimos de segundo.

4.1.6 Resultados do Português Brasileiro

4.1.6.1 Dados *off-line* - PB

Conforme já expusemos, há medidas *off-line* e *on-line*, e ambas são importantes quando da análise dos dados. As medidas *off-line* revelam fatos sobre o processamento mais tardio, numa fase mais interpretativa, depois de ele ter ocorrido.

No que concerne às medidas *off-line* deste experimento, aferimos o tempo de resposta, bem como o índice de acertos das questões referentes à sentença lida.

Ainda no que diz respeito a essas medidas, encontramos um índice de acertos maior nas relativas de sujeito do que nas de objeto. Houve, nas 360 questões que totalizaram o experimento, 13 erros nas relativas de sujeito e 37 erros nas relativas de objeto. Essas diferenças foram submetidas a um teste de qui-quadrado, e a proporção de erros para uma e outra condição apresentou uma diferença significativa: ($\chi^2=12,28$, $p<0,001$). A proporção de erros e acertos estão representados na Tabela a seguir:

Tabela 9: Proporção de erros e acertos para as relativas de sujeito e objeto em PB.

Índice de erros		
Variedade	Relativa de Sujeito	Relativa de Objeto
PB	13 (7,2%)	37 (20,5%)
P-valor: ($\chi^2=12,28$, $p<0,001$) – 180 questões por condição		

O índice maior de erro para as relativas de objeto será discutido mais adiante, cabe aqui ponderar, porém, que a menor ocorrência dessas estruturas na língua, apontada por alguns autores e confirmada por nós nos estudos de *corpus* apresentado no capítulo anterior, pode ter contribuído para essa dificuldade maior, aferida nas medidas *off-line* das CRO quando relacionadas às CRS.

Contabilizamos também a média do tempo de resposta, as CRS foram respondidas em 2.264 milésimos de segundo (ms) e as CRO em 2.983 ms.

Cabe ser posto ainda que os tempos muito acima ou abaixo da média foram descartados da contagem geral, isso foi feito por meio do *Box-Plot*, ferramenta disponível no Excel que descarta as discrepâncias com base no desvio padrão.

Nesse segmento, especificamente, contabilizadas as duas condições, foram encontradas 13 medidas discrepantes, os *outliers*. Essa “limpeza” nos dados é necessária, e portando, será realizada em todos os experimentos constitutivos desta tese. A finalidade é descartar aquelas medidas que se excederem, para mais ou para menos, devido a algum problema, por exemplo, de falta de atenção do participante no decorrer do experimento.

Submetemos nossos dados, relativos ao tempo de resposta, por meio do *Action* (versão 2.9.29.368.534), ao teste-T pareado e encontramos uma diferença significativa, o P-valor ($t(179) = 7,24p < 0,001$). Os resultados estão representados no gráfico que se segue em milésimos de segundo.

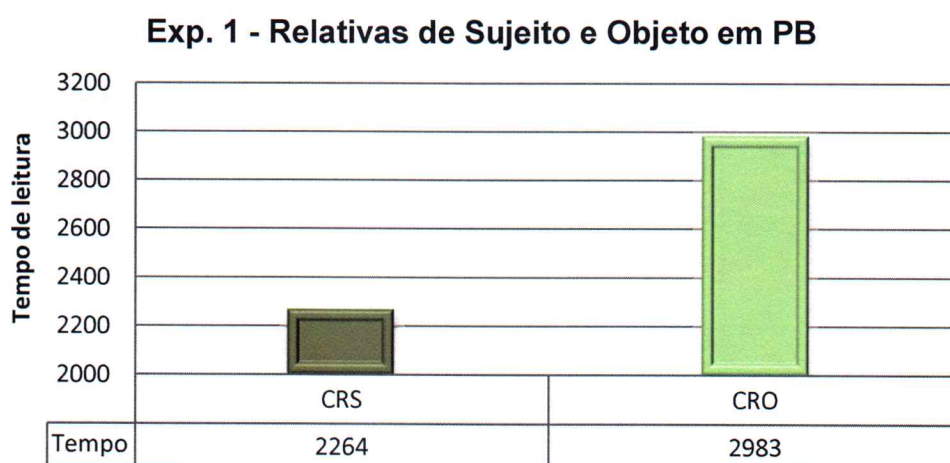


Figura 11: Resultado do tempo de resposta - PB

De antemão, podemos afirmar que, segundo as medidas *off-line*, as relativas de sujeito são mais fáceis de processar, quer pelo fato de as respostas terem sido mais rápidas, quer pelo índice de acerto, que foi maior quando comparadas às relativas de objeto. Houve diferenças significativas do ponto de vista estatístico nas duas medidas *off-line*. Tendo em conta que, conforme apontamos no início da descrição do experimento, as relativas trazidas nestes dois experimentos aqui reportados, são explicativas, nossos resultados confirmam o que foi encontrado nos estudos anteriores em que foram testadas relativas restritivas no que se refere às medidas *off-line*.

4.1.6.2 Dados *on-line* - PB

Na computação dos dados *on-line*, aferimos o segmento 2, e confirmando as medidas *off-line*, realmente as estruturas de sujeito foram mais rápidas do que as de objeto. Também por meio do *Box-Plot*, eliminamos 25 *outliers*, e posteriormente procedemos à análise estatística. O segmento crítico, S2, foi lido, em média, em 1,194 ms nas CRS, ao passo que as CRO tiveram esse mesmo segmento lido em média em 1.372 ms. Submetido ao teste-T pareado, encontramos um P-valor ($t(179) = 4,51$; $p < 0,001$) ou seja, uma diferença estatisticamente significativa. Os dados *on-line* confirmam, pois, que as relativas explicativas possuem a mesma assimetria entre CRS e CRO quando aferidas as medidas *on-line* de relativas restritivas, sem deixar de considerar o fato de os termos relativizados são todos animados.

Tabela 10: Resultados do tempo de processamento do segmento crítico do PB.

Experimento 1 - Relativas de Sujeito e Objeto em PB Segmento Crítico		
Variedade	Relativa de Sujeito	Relativa de Objeto
PB	1194 (416,52)	1372 (549,43)
Desvio Padrão entre Parêntesis – Diferença significativa: P-valor ($t(179) = 4,51$; $p < 0,001$)		

O primeiro experimento confirma nossas hipóteses de que as CRS são estruturas mais fáceis de processar em PB do que as CRO. Essa prerrogativa também foi considerada pela maior parte dos estudos que procuraram compreender o processamento de relativas de sujeito e de objeto em muitas línguas naturais. Contudo, consideraremos, mais adiante neste trabalho, o traço da animacidade como uma possível influência no *parsing* dessas relativas.

4.1.7 – Resultados do Português Europeu

Por ora, tomaremos os dados do PE na réplica do experimento que reportamos acima nessa variedade da língua Portuguesa. É mister considerar que, sendo este experimento uma réplica, utilizamos as mesmas estruturas

frasais, modificando-as, apenas, quando as diferenças dialetais que há entre as duas variedades pudessem prejudicar a fluidez da leitura diante de uma possível anti-naturalidade da estrutura em questão.

Ajustamos também as questões ortográficas pertinentes a cada variedade. No mais, obedecemos aos mesmos procedimentos anteriormente reportados na coleta e tratamento dos dados.

Por fim, o segmento crítico no PE foi mais extenso também porque continha um artigo definido, as estruturas experimentais do PB seguiram o padrão de 46, e do PE seguiram o padrão de 47, abaixo. Essa diferença se mostrou necessária porque, na cidade de Natal, onde foram rodados os experimentos do PB, não se usa o artigo definido antes de nomes próprios, diferentemente do que acontece no PE.

46. João / **que encontrou José** / naquele dia / não quis / pedir perdão.

S1 S2 S3 S4 S5

47. O João / **que encontrou o José** / naquele dia / não quis / pedir perdão.

S1 S2 S3 S4 S5

4.1.7.1 Dados *off-line* - PE

A quantidade de erros nas respostas dadas, também em PE, foram muito superiores nas CRO do que nas CRS. Nas relativas de sujeito houve 8 erros, enquanto nas relativas de objeto foram 33 questões com respostas erradas. Por meio de um teste qui-quadrado, encontramos uma diferença significativa entre as proporções de erros das CRS e CRO, tendo sido as CRO as estruturas que apresentaram o maior número de erro. ($\chi^2=14,38$; $p < 0,001$).

Tabela 11: Proporção de erros e acertos para as relativas de sujeito e objeto em PE.

Índice de erros		
Variedade	Relativa de Sujeito	Relativa de Objeto
PE	8 (4,4%)	33 (18,3%)
P-valor: ($\chi^2 = 14,38$; $p < 0,001$) – 180 questões por condição		

O tempo de resposta também foi menor nas CRS, sendo 1.837 ms nessas construções, e 2.687 para as CRO. Feito o teste-T, encontramos um P-

valor significativo ($t(191) = 9,27$ $p < 0,001$). Ou seja, diferença significativa entre as duas estruturas. Esse resultado também foi obtido depois de termos retirado os *outliers* apontados pelo *Box-Plot*, que neste segmento foi num total de 20 itens excluídos. As médias, em ms, estão representadas no gráfico abaixo:

Tabela 12: Resultados do tempo de processamento do tempo de resposta em PE

Experimento 2 - Relativas de Sujeito e Objeto em PE		
Variedade	Relativa de objeto	Relativa de Sujeito
PE	1837 (533,3)	2687 (407,86)
Desvio Padrão entre Parêntesis		

4.1.7.2 Dados *on-line* - PE

Nas medidas *on-line*, encontramos as CRS mais rápidas de processar em PE, ou seja, o mesmo resultado encontrado no PB. O segmento crítico das CRS foi lido em 1.410 ms, significativamente mais rápido do que o segmento crítico das CRO, que foi de 1.665 ms. Feito o teste-T pareado, encontramos um P-valor ($t(191) = 4,59$ $p < 0,001$). Neste segmento, por meio do *Box-Plot*, eliminamos 18 *outliers*. Representamos esses resultados no gráfico 13, abaixo.

Tabela 13: Resultados do tempo de processamento do segmento crítico do PE

Experimento 2 - Relativas de Sujeito e Objeto em PE Segmento Crítico		
Variedade	Relativa de Sujeito	Relativa de Objeto
PE	1410(529,14)	1665(748,32)
Desvio Padrão entre Parêntesis - Diferença significativa: P-valor ($t(191) = 4,59$; $p < 0,001$).		

Também medimos o segmento pós-crítico, S3, neste experimento do PE, como no anterior do PB, com o intuito de por ventura captar um efeito *spillover*, não encontramos, porém, efeitos significativos em S3 em nenhuma das variedades duas variedades do Português aqui estudadas.

Diante de tais resultados, percebemos que, tanto em PB quanto em PE as CRS são mais fáceis de processar do que as CRO. As explicações sobre

esse fato são dadas em muitos estudos experimentais a partir das diferenças sintáticas constitutivas dessas estruturas.

Nesta tese, não negamos tais diferenças, assumimos juntamente com a literatura pertinente ao tema que essas diferenças provocam sim níveis distintos de dificuldade entre as duas estruturas, o que pode acarretar a assimetria vista nas duas variedades da língua Portuguesa, aqui investigadas, porém, admitimos que informações de cunho semântico, bem como a frequência de ocorrência dessas estruturas na língua, por exemplo, podem interferir no processo. É preciso ressaltar ainda que, conforme discutimos na seção 2.5, assumimos o traço da animacidade como um traço semântico.

As diferenças *on-line* do experimento, respectivamente em PB e PE, estão representadas no gráfico abaixo disposto.

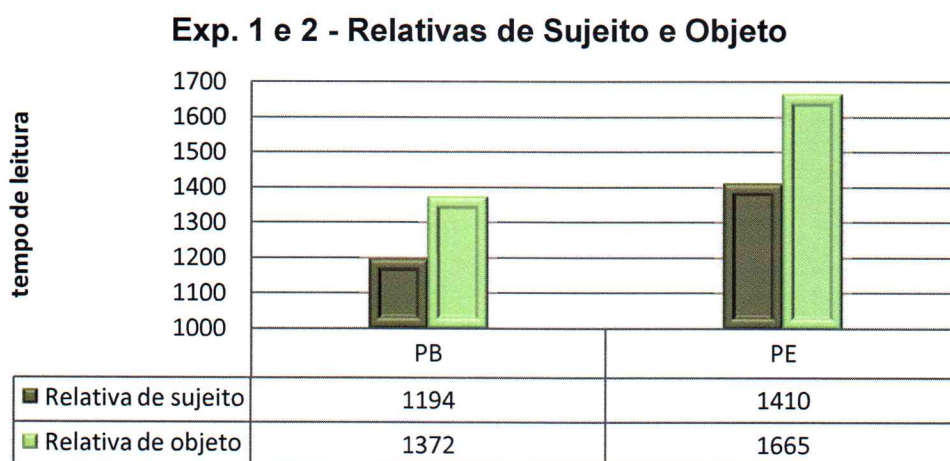


Figura 12: Resultado do tempo de processamento do segmento crítico - PB e PE

4.1.7.3 – Comparação PB/PE

Note-se que em todas as condições a leitura do segmento crítico foi mais rápida em PB do que em PE, atribuímos isso ao fato de haver diferenças, ainda que sutis, entre as estruturas experimentais, que no PE possuem um artigo definido e em PB não possuem, conforme justificamos no início desta seção.

Como foi apontado anteriormente, o segmento 3, pós-crítico, foi devidamente controlado quanto ao seu tamanho e mensurado nos dois

experimentos acima reportados. Não encontramos, porém, nenhuma diferença significativa nas medidas aferidas. Concluímos, então, que os efeitos aqui investigados são antes dos segmentos em que se encontra a estrutura relativa, não havendo, pois, um efeito *spillover*.

Para concluirmos a exposição dos dados desses dois experimentos, cabe-nos apontar para o fato que, conforme já discutimos anteriormente, as CRO são todas compostas de termos antecedentes relativizados animados. Ou seja, estruturas muito pouco frequentes nas duas variedades da língua aqui analisadas. Posto isso, aventamos a possibilidade de essa raridade da estrutura influenciar de alguma forma o *parsing* da sentença, o que discutiremos com mais propriedade a partir dos dados que traremos na próxima seção.

Se isso for verdade, para além da animacidade, também a frequência pode influenciar o processamento de cláusulas relativas. Tal pressuposto é condizente com a DLT. Mais adiante retomaremos essa questão. Ao tratarmos de frequência, outros modelos teóricos, a princípio, podem responder aos questionamentos que levam às explicações dos fenômenos aqui estudados. Contudo, como discutimos no final da seção 2.4.4, a DLT não apenas admite a frequência como um fator possivelmente visto durante o *parsing*, mas ainda o considera dentro das estruturas com que lidamos aqui.

4.2 Experimentos 3 e 4 – CRSa, CRSi, CROa e CROi em PB e PE

Os dois experimentos que reportaremos nesta seção trazem os dados que são o cerne de nossa discussão nesta tese. Temos discutido, desde o início deste trabalho, sobre o acesso que o *parser* tem ao processar uma estrutura relativa nas línguas humanas.

Os modelos teóricos que elencamos para nossa discussão não são unânicos nessa questão, e a partir dos dados que serão apontados nestes dois próximos experimentos, pretendemos aclarar essa problemática, apresentando algumas evidências empíricas que, em algum nível, parecem contribuir para a constituição de uma teoria psicolinguística que retrate o processamento de sentenças relativas num momento inicial da leitura.

Assumimos, no capítulo 2, a DLT (GIBSON, 2000), como uma proposta teórica que pode abarcar os fenômenos por nós investigados.

Como temos reafirmado, a TGP (FRAZIER, 1979) descreve o *parser* como modular, serial, encapsulado, que, grosso modo, afirma que o processador sintático é cego às informações que não fizerem parte da estrutura sintática. Se visto por esse modelo, a animacidade, traço de natureza semântica, não poderia ser acessada pelo *parser* nos momentos iniciais do *parsing*. A TGP, na realidade, não prevê que a leitura de relativas possa ser afetada pelo traço da animacidade, uma vez que esse traço só seria considerado nos momentos mais tardios do processamento.

Por outro lado, para Gibson (2000), não só a animacidade, mas outras informações de natureza semântica, pragmática e discursiva, por exemplo, podem guiar o *parser* nos primeiros momentos do *parsing* de sentenças. Mais que isso, a DLT admite o uso da frequência como fator relevante no processamento *on-line*, e isso nos será fundamental neste estudo.

O modelo de Gibson concebe uma arquitetura interativa do *parser*, pois considera, por exemplo, a atuação da maior complexidade das CRO na comparação com as CRS, já que o custo de integração e armazenamento das duas construções é distinto, pois a distância de aposição maior de CRO é um fator que demanda custo operacional. Contudo, diferente do que assume a TGP, a DLT não concebe o *parser* como encapsulado, ou seja, o *parser* não é cego a informações não estruturais. Para esse modelo a complexidade sintática, pode, inclusive não ser o peso de maior valia frente a outros fatores, como, por exemplo, a frequência.

Assim, sendo um modelo interativo, informações sintáticas são utilizadas pelo *parser* nos momentos iniciais da sentença, mas essas informações podem e são guiadas por outras informações que estejam disponíveis no curso do processamento. Se a DLT estiver correta, a frequência de distribuição do traço da animacidade entre sujeito e objeto de relativas poderá guiar o *parser* na leitura das frases experimentais dos dois experimentos que se seguem. Sendo assim, haverá efeito principal de animacidade, ou seja, esse traço mudará o tempo de processamento das estruturas experimentais. As frases experimentais deste estudo seguem o padrão de 48, 49, 50 e 51, que trazem

relativas encaixadas restritivas, respectivamente, com extração de sujeito animado, de sujeito inanimado, de objeto animado e de objeto inanimado.

48. O gato / **que encantou a menina** / de tarde / estava brincando / na grama do jardim.
S1 S2 S3 S4 S5
49. O livro / **que encantou a menina** / de tarde / era sobre a vida / das aves do cerrado.
S1 S2 S3 S4 S5
50. O livro / **que a menina encontrou** / de tarde / na minha casa / era muito interessante.
S1 S2 S3 S4 S5
51. O gato / **que a menina encontrou** / de tarde / naquela praça / comeu a ração dele.
S1 S2 S3 S4 S5

Como se pode observar, controlamos a animacidade dos termos antecedentes retomados pela relativa, bem como o tipo de estrutura, se de sujeito ou se de objeto. Com isso, temos 4 condições experimentais, a saber: CRSa, CRSi, CROa e CROi.

Temos aqui também que ressaltar o fato de essas estruturas serem restritivas, diferentemente do que fizemos nos experimentos 1 e 2, em que as estruturas eram explicativas. Assim, testaremos agora estruturas idênticas às que foram testadas nos muitos estudos que atestaram a assimetria entre CRS e CRO, e diante do controle da animacidade dos termos antecedentes, poderemos afirmar se essa assimetria é confirmada ou refutada.

Posto assim, há duas variáveis independentes, a animacidade (animado ou inanimado) e o tipo de relativa (com extração de sujeito ou de objeto). Como variáveis dependentes temos o tempo de processamento do segmento crítico, destacado em negrito nos exemplos acima, e o tempo de processamento da resposta à pergunta feita posteriormente à leitura, bem como o índice de erros e acertos das perguntas feitas após a leitura da sentença.

O experimento 4 é uma réplica do experimento 3. As condições experimentais foram as mesmas, contudo, ajustes foram feitos no sentido de adequar as frases experimentais à variedade Europeia da língua Portuguesa.

A expectativa é que o controle da animacidade, com base nos pressupostos constitutivos da DLT, seja vista pelo *parser* e isso possa provocar alguma influência no processamento de relativas de sujeito e de objeto.

Mak, Vonk e Schiefers (2002, 2006), encontraram evidências, no Alemão e no Holandês, de que a animacidade influencia a distribuição de relativas nessas línguas e ainda demonstraram que o controle do traço em estudos

experimentais fez com que o tempo de processamento de relativas de objeto, tradicionalmente mais lento que o de sujeito, fosse relativizado.

Segundo os pressupostos da DLT, há uma maior complexidade nas CRO em relação às CRS, contudo, essa maior complexidade não é o único fator acessado pelo *parser*. Nesse modelo, outras informações podem guiar o *parser*, com base, por exemplo, na frequência de realização dos itens lexicais em determinadas estruturas da língua. Itens lexicais carregam traços semânticos, a exemplo da animacidade. Dito isso, no modelo proposto por Gibson (2000), a frequência de realização de uma dada estrutura que possui itens animados ou inanimados pode ser um dos fatores que guia o *parser* quando da leitura de sentenças. Ou mais especificamente, a frequência de ocorrência de CRS e CRO, tendo em conta se seus antecedentes são animados ou inanimados, pode acarretar maior ou menor custo de processamento para essas estruturas.

Diante dessas considerações, conduzimos nossos dois próximos experimentos com base nos trabalhos de Mak, Vonk e Schiefers (2002, 2006), e supomos que em PB e PE o mesmo que ocorreu no Holandês vai acontecer nessas duas variedades do Português.

Cabe aqui um adendo. De acordo com o que expusemos no capítulo 2 desta tese, nossos experimentos não são uma réplica dos trabalhos feitos no Holandês, haja vista haver diferenças estruturais consideráveis entre o Holandês e o Português que impedem que possamos traçar uma comparação direta entre essas línguas.

4.2.1 Participantes do PB

Os participantes do terceiro experimento são todos falantes do PB, foram selecionados 32 estudantes do ensino tecnológico, sendo 16 homens e 16 mulheres, idade média de 17,3 anos. Todos eles são alunos do Instituto Federal do Rio Grande do Norte, do *campus* Zona Norte, e no *campus* Cidade Alta, na cidade de Natal.

A pesquisa aconteceu com a colaboração do professor Fábio Alexandre Araújo, no campus Zona Norte, e do professor Dayvyd Lavaniery Marques de Medeiros, no campus Cidade Alta. Esses professores nos cederam o espaço

necessário na referida instituição e ainda nos auxiliaram na seleção dos alunos que participaram da pesquisa.

Todos os alunos foram submetidos a um questionário em que declaram ser falantes do Português Brasileiro como primeira língua. O questionário também nos permitiu aferir o nível de letramento dos alunos e dos familiares com quem convivem.

4.2.2 Participantes do PE

Também esse experimento, assim como o experimento 2, foi feito na Universidade do Minho, na cidade de Braga, Portugal, mais uma vez tivemos a colaboração da professora Maria do Carmo Lourenço-Gomes. Ela nos cedeu o espaço para realização do experimento e ajudou na programação do mesmo, garantindo assim que as frases estivessem ajustadas às especificidades do PE. Os participantes do quarto experimento são todos falantes nativos do PE, foram selecionados 32 estudantes entre alunos de graduação e pós-graduação da Universidade do Minho, sendo 16 homens e 16 mulheres, idade média de 21,5 anos.

4.2.3 Materiais

Cada participante foi exposto a um conjunto de 16 frases experimentais e 32 distratoras. Entre as experimentais, 4 eram CRSa, 4 CRSi, 4 eram CROa e 4 CROi. Ou seja, o participante foi exposto 4 vezes a cada uma das quatro condições do experimento. Para que um participante não tivesse acesso à mesma frase em cada uma das versões apresentadas, constituímos um quadrado latino como o que se segue:

Tabela 14: Quadrado latino dos experimentos 3 e 4

1 A	1 B	1 C	1 D
2 B	2 C	2 D	2 A
3 C	3 D	3 A	3 B
4 D	4 A	4 B	4 C
5 A	5 B	5 C	5 D
6 B	6 C	6 D	6 A
7 C	7 D	7 A	7 B
8 D	8 A	8 B	8 C
9 A	9 B	9 C	9 D
10 B	10 C	10 D	10 A
11 C	11 D	11 A	11 B
12 D	12 A	12 B	12 C
13 A	13 B	13 C	13 D
14 B	14 C	14 D	14 A
15 C	15 D	15 A	15 B
16 D	16 A	16 B	16 C

Cada letra – A, B, C e D– representa, respectivamente, CRS animado, CRS inanimado, CRO animado, e CRO inanimado. Cada coluna representa um grupo de frases disponíveis a um participante. Ou seja, distribuído dessa maneira, um participante em momento algum foi exposto à outra versão da frase que já foi lida.

As 32 frases distratoras tiveram a função de impedir que o participante percebesse um dado padrão nas estruturas frasais e com isso formulasse algum tipo de estratégia de leitura enviesando os resultados do experimento.

Todas as frases experimentais deveriam receber um “sim” como resposta à pergunta subsequente. Balanceamos para que metade das respostas fosse “sim” e metade fosse “não” com as frases distratoras. Dessa forma, cada participante foi exposto a 48 frases, entre distratoras e experimentais.

Foi controlado o tamanho dos itens investigados. O segmento crítico analisado foi o segundo (S2), contendo um pronome relativo (que), um verbo trissílabo na terceira pessoa do singular, um substantivo comum, também trissílabo na condição de sujeito ou de objeto conforme a frase. O substantivo antecedente era dissílabo, e foi controlado quanto à animacidade, tendo sido 50% animado e 50% inanimado.

Assim como fizemos nos dois primeiros experimentos, mensuramos o segmento seguinte (S3) a fim de averiguar a possível ocorrência de um efeito *spillover*. Para tanto, controlamos também o tamanho do segmento pós-crítico. Adiantamos que, também aqui não encontramos efeitos significativos em S3, descartando, pois, a possibilidade de um efeito *spillover* para estes experimentos.

4.2.4 Procedimentos

Os participantes foram instruídos individualmente sobre as tarefas constitutivas do experimento. As instruções foram apresentadas oralmente pelo experimentador, e posteriormente, o participante lia as tarefas na tela do computador em que o experimento foi rodado, seguindo o mesmo padrão dos experimentos 1 e 2.

Depois de ter declarado que entendeu a tarefa, o participante se submetia a um treinamento prático constituído de 4 frases distratoras. Com isso asseguramos que os participantes compreenderam perfeitamente as tarefas a serem realizadas, que só eram iniciadas quando do consentimento do participante que demonstrava ter de fato compreendido seu papel.

Ao pressionar a tecla “espaço”, aparecia o primeiro segmento no centro da tela do computador, apertando novamente parecia o segundo segmento, no lugar do primeiro, que conseqüentemente desaparecia. O procedimento deveria ser repetido por cinco vezes, até aparecer a pergunta a que ele tinha que responder acionando a tecla “sim”, devidamente adaptada no teclado do computador, na cor verde, ou “não”, destacada no teclado na cor vermelha.

No Brasil, o experimento foi realizado no mês de dezembro de 2015, conforme já indicamos, no Instituto Federal do Rio Grande do Norte, *campus* Zona Norte, e no *campus* Cidades Alta, ambos em Natal, RN. Em Portugal, realizamos o experimento no mês de novembro desse mesmo ano na Universidade do Minho, na cidade de Braga.

O experimento foi rodado em um MacBook, com uma tela de 13". G3, 233Mhz. Utilizamos o programa *Psyscope* (COHEN; MACWHINNEY; FLATT & PROVOST, 1993). As frases foram randomizadas e o tempo de leitura foi aferido em milésimos de segundo.

4.2.5 Resultados do Português Brasileiro

4.2.5.1 – Dados *off-line* – PB

Começaremos nossa descrição pelas medidas *off-line* deste experimento, aferimos o tempo de resposta, ou seja, a duração do segmento 6, bem como o índice de acertos das questões referentes à sentença lida.

As CRSi apresentaram o menor índice de erros, num total de 10 respostas erradas, seguidas das CROa, com 13 erros, e depois pelas CRSi, com 15 erros, por fim, as CROi apresentaram o maior número de erros, um total de 18. Cada uma dessas condições possuía um total de 128 perguntas. Submetemos os índices de erros das 4 condições a um teste de qui-quadrado, mas não houve resultado significativo quanto a esse índice, o p-Valor foi de ($\chi^2=2,72$; $p < 0,43$).

Tabela 15: Proporção de erros e acertos para as relativas de sujeito e objeto em PB.

Índice de erros				
Variedade	Suj. Ani.	Suj. Ina.	Obj. Ani.	Obj. Ina.
PB	10 (7,8%)	15 (11,7%)	13 (10,1%)	18 (14%)
P-Valor ($\chi^2=2,72$, $p<0,43$) – 128 questões por condição				

Submetemos nossos dados relativos ao tempo de resposta ao ezANOVA, programa por meio do qual obtivemos a Análise da Variância (ANOVA).

No tocante ao tempo de resposta, respectivamente na análise por participante (F1) e por item (F2), houve efeito de interação entre o tipo de relativa e a animacidade ($F(1,31) = 9,11$; $p < 0,005$) e ($F(2,1,3) = 41,2$; $p < 0,007$). Não houve efeitos principais no tipo de relativa ($F(2,1,31) = 0,272$; $p < 0,6$) e ($F(2,1,3) = 0,648$; $p < 0,47$) e nem na animacidade $F(1,31) = 0,151$; $p < 0,7$) e ($F(2,1,3) = 0,779$; $p < 0,44$).

Houve, entre as condições experimentais, algumas diferenças significativas, pois o tempo de resposta de CROi foi significativamente mais rápido do que CROa ($t(31)=2,20$; $p < 0,035$). A resposta referente à estrutura

com CROi foi também mais rápida do que a resposta da estrutura com CRSi ($t(31)=2,40$; $p < 0,022$).

A CRSa, por sua vez, foi mais rápido do que o CROa, ($t(3)=9,18$ $p < 0,002$). As médias de tempo de resposta deste experimento estão representadas na Tabela abaixo.

Tabela 16: Resultados do tempo de resposta – PB

Experimento 3 - Tempo de Resposta – Português Brasileiro

Relativa	Súj. Ani.	Súj. Ina.	Obj. Ani.	Obj. Ina.
PE	2561(533,3)	2764 (508,08)	2770 (407,86)	2470 (487,53)
Desvio Padrão entre Parêntesis				

Perceba-se que, curiosamente, as relativas com objeto inanimado foram as que obtiveram o tempo de resposta mais rápido, tendo sido significativamente mais rápido do que o tempo do objeto animado. Cabe lembrar que a condição CROa é a condição utilizada na maior parte experimentos que trata dessa assimetria, e o objeto inanimado é desconsiderado na maior parte das pesquisas. Ainda como mostra a Tabela 16, o sujeito animado foi mais rápido do que o objeto animado, nessas que são as condições normalmente testadas nos experimentos que tratam da assimetria entre CRS e CRO.

Como já posto, nosso foco neste trabalho é, sobretudo, o processamento *on-line*, já que pretendemos saber se a animacidade influencia o *parsing* nesse primeiro momento. Dito isso, retomaremos os dados desta seção mais adiante, por ora, fica registrado que os tempos de resposta já dão sinais de que o processamento se dá de forma diferente quando da computação de relativas de sujeito e de objeto se controlada a animacidade dessas estruturas.

4.2.5.2 – Dados *on-line* – PB

Os dados da computação *on-line* nos trouxeram resultados condizentes com nossas previsões. Houve diferenças significativas entre os tempos médios

de leitura do segmento crítico, S2, o que discutiremos a partir do gráfico que se segue.

Cabe ainda ser posto que, assim como nos experimentos anteriores, também aqui, em todas as análises, utilizamos o *Box-Plot* com o fim de eliminar os *outliers*. Nas condições abaixo relacionadas, foram encontrados respectivamente 10, 7, 11 e 8 *outliers* em cada condição²⁹.

Tabela 17: Resultados do tempo de leitura do segmento crítico – S2 – em PB.

Experimento 3 - Tempo de leitura do Segmento Crítico – Português Brasileiro

Relativa	Suj. Ani.	Suj. Ina.	Obj. Ani.	Obj. Ina.
PB	1520 (322,63)	1393 (320,15)	1672 (332,96)	1327(220,71)
Desvio Padrão entre Parêntesis				

Fizemos a ANOVA por meio do *Action* (versão 2.9.29.368.534), que usa a base de cálculo do R (versão 3.0.2). Encontramos efeito principal de animacidade ($F(1,62) = 19,51$; $p < 0,001$) e ($F(2,30) = 17,7$; $p < 0,001$), e efeito de interação ($F(1,62) = 4,19$; $p < 0,042$) na análise por participante, mas não houve efeito de interação na análise por item, $F(2,30) = 3,15$; $p < 0,081$).

Além disso, como podemos observar na Tabela 17, quando olhamos para as diferenças entre as médias do tempo de leitura das condições, vemos que o tempo de leitura de CROi foi significativamente mais rápido do que CROa ($t(31)=6,58$; $p < 0,001$) repetindo-se aqui o que ocorreu na medida *off-line*, acima apresentada. É muito importante ser posto que as CROi alcançaram as médias mais baixas de todas as condições, tendo sido, inclusive, mais rápidas do que todas as CRS.

É importante ressaltarmos, mais uma vez, que as estruturas tradicionalmente trazidas à baila quando da discussão da assimetria entre CRS e CRO são as de sujeito animado e de objeto animado. Ao se incluírem as outras estruturas no cômputo dos tempos aferidos, com base nos pressupostos teóricos de que nos valem, é de se supor que haverá uma relativização dos resultados tradicionalmente encontrados.

²⁹Os dados amostrais passaram pelo teste de normalidade Shapiro-wilk que indicou padrão normal: P-valor = 0,147, o que permite o uso dos testes paramétricos.

Mais que isso, nossos resultados coadunam com os achados de Mak, Vonk e Schiefers (2002, 2006), trabalhos segundo os quais, quando controlamos o traço da animacidade, as relativas de sujeito não necessariamente serão as estruturas processadas mais rapidamente, conforme a literatura predominante que temos à disposição.

Ora, diante desses resultados, parece-nos que a frequência de ocorrência de estruturas relativas com o traço da animacidade positivo ou negativo de alguma forma guia o *parsing* de cláusulas relativas, uma vez que, como visto, encontramos efeito principal de animacidade em nossos estudos experimentais. Precisamos, pois, a partir de uma análise mais acurada, compreender os fatores que levariam a isso. Tal análise dar-se-á a partir do próximo experimento, em que replicamos o experimento ora descrito na variedade Europeia da Língua Portuguesa.

4.2.6 Resultados do Português Europeu

4.2.6.1 – Dados *off-line* – PE

Tomaremos agora as medidas *off-line* deste experimento em PE, aferimos o tempo de resposta, bem como o índice de acertos das questões referentes à sentença lida.

As CRSa, apresentaram 13 erros; as CRSi, 28 erros; as CRSa, 22 erros; e as CROi, 17 erros. Cada uma dessas condições possuía um total de 144 perguntas. Submetemos os índices de erros das 4 condições a um teste de qui-quadrado, mas não houve resultado significativo quanto a esse índice, o p-Valor foi de ($\chi^2=7,31$, $p<0,06$), muito embora tenha apresentado uma tendência.

Tabela 18: Proporção de erros e acertos para as relativas de sujeito e objeto em PE.

Índice de erros				
Variedade	Suj. Ani.	Suj. Ina.	Obj. Ani.	Obj. Ina.
PE	13 (9%)	28 (19,4%)	22 (15,2%)	17 (11,8%)
P-Valor ($\chi^2=7,31$, $p<0,06$) – 144 questões por condição				

Submetemos nossos dados relativos ao tempo de resposta ao ezANOVA, programa por meio do qual obtivemos a Análise da Variância (ANOVA). Quanto ao tempo de resposta, na análise por participante (F1), houve efeito de interação entre o tipo de relativa e a animacidade ($F(1,70) = 63,4$; $p < 0,001$). Houve efeito principal no tipo de relativa $F(1,70) = 6,10$; $p < 0,015$), mas não houve efeito principal de animacidade.

Houve, entre as condições experimentais, algumas diferenças significativas, o tempo de resposta de CROi foi significativamente mais rápido do que CROa $t(35)=11,06$ $p < 0,0001$. CRSa, por sua vez, foi mais rápido do que o objeto animado, $t(3)=5,35$ $p < 0,001$) e do que o Sujeito Inanimado $t(3)=4,16$ $p < 0,001$). As médias de tempo de resposta deste experimento estão representadas na Tabela abaixo.

Tabela 19: Resultados do tempo de resposta – PE

Experimento 4 - Tempo de Resposta – Português Europeu

Relativa	Suj. Ani.	Suj. Ina.	Obj. Ani.	Obj. Ina.
PE	1953(337,99)	2286 (458,94)	2550 (578,77)	2198 (488,88)
Desvio Padrão entre Parêntesis				

Em suma, os dados dos dois experimentos que acabamos de reportar mostram-nos que a assimetria CRS/CRO, nos termos que vem sendo discutida na literatura que trata do tema, não se sustenta nos dados de nossos experimentos.

4.2.6.2 – Dados *on-line* – PE

Traremos agora as medidas *on-line*, do PE, para que possamos fazer uma comparação com os resultados *on-line* apresentados do PB. Os dados da computação *on-line* do PE mostram evidências semelhantes às que foram encontradas no PB. Houve diferenças significativas entre os tempos médios de leitura do segmento crítico S2 quando da manipulação da animacidade, o que confirma nossa hipótese de que esse traço pode guiar o *parser* em suas decisões.

Neste segmento, assim como nos experimentos anteriores, também utilizamos o *Box-Plot* com o fim de eliminar os *outliers*. As CRS com antecedente animado apresentaram 9 *outliers*; com antecedente inanimado, 6 *outliers*; nas CRO com antecedente animado e inanimado encontramos, respectivamente, 7 e 11 *outliers*.

Tabela 20: Resultados do tempo de processamento segmento crítico – S2 – em PE.

Experimento 4 - Tempo de leitura do Segmento Crítico Português Europeu

Relativa	Suj. Ani.	Suj. Ina.	Obj. Ani.	Obj. Ina.
PE	1504 (291,15)	1311 (191,32)	1637 (267,52)	1281 (231,38)
Desvio Padrão entre Parêntesis				

Também com o PE, Fizemos a ANOVA por meio do *Action* (versão 2.9.29.368.534), que usa a base de cálculo do R (versão 3.0.2).

Encontramos efeito principal de animacidade ($F(1,62) = 13,2$; $p < 0,001$) e ($F(1,30) = 30,69$; $p < 0,001$), o mesmo que aconteceu no PB. Não houve, porém, efeito de interação.

Com base na Tabela 20, podemos verificar que o tempo de leitura do segmento 2 referente à CROi foi significativamente mais rápido do que do CROa ($t(35)=5,08$; $p < 0,001$), também da mesma forma que aconteceu no PB. CRSi também foi significativamente mais rápido do que o CROa ($t(35)=2,82$; $p < 0,007$) Outra medida que coincide com o experimento feito em PB.

Como se pode observar, embora tenham ocorrido algumas diferenças pontuais, os dados do Português Europeu confirmam o que encontramos em Português Brasileiro, e confirmam ainda nossas hipóteses de que no Português, assim como aconteceu no Holandês, a manipulação do traço semântico da animacidade mudou o tempo de leitura das estruturas relativas, do que se pode depreender que esse traço pode, de alguma maneira, ser acessado pelo *parser* nos momentos iniciais da sentença.

Nossos dados mostram que o tempo de leitura dessas estruturas é significativamente modificado a partir da manipulação de um traço semântico, se por um lado isso não desfaz a assimetria, o processamento dessas estruturas continua sendo diferente quando comparadas umas às outras; por

outro lado, outros fatores, como a frequência de ocorrência do traço semântico manipulado – a animacidade – nessas estruturas pode provocar a assimetria de forma ainda mais contundente do que as diferenças estruturais propriamente ditas.

Pretendemos discutir e explicar essas afirmações com base nos pressupostos da DLT, o que faremos na seção 4.4, depois que descrevermos o próximo experimento.

4.3 Leitura Automonitorada de palavras isoladas

A necessidade deste experimento se deu frente ao fato de querermos mensurar o tempo de leitura das palavras isoladas, sem a interferência do contexto sintático. Para a realização deste experimento, queríamos que a frequência das palavras escolhidas também não interferisse no tempo de processamento, para tanto, usamos as mesmas palavras que compuseram as frases dos experimentos 3 e 4. A princípio, consideramos o fato de termos encontrado nos experimentos anteriores, em que utilizamos palavras animadas e inanimadas, diferenças significativas no que tange às estruturas com antecedentes controlados quando à animacidade.

Trata-se de uma tarefa de Leitura Automonitorada, mesmo consistindo na leitura de palavras e não de frases, já que o participantes automonitoram sua leitura. Albuquerque (2008) utilizou-se dessa metodologia para fins parecidos com os nossos.

Queremos saber se as mesmas palavras dos experimentos 3 e 4, lidas isoladamente, são processadas em tempos distintos quando da manipulação do traço da animacidade, assim como aconteceu nas sentenças relativas de sujeito e de objeto.

Nossa hipótese é de que um item lexical, isoladamente, não terá seu tempo de leitura alterado apenas por conta de um traço semântico. Se nossas previsões se confirmarem, e o tempo de leitura de substantivos animados e inanimados não for significativamente diferente, teremos indícios empíricos que não é a animacidade das palavras em si que provoca a mudança no tempo de processamento das relativas, mas fatores de outras naturezas, como, por exemplo, a frequência de uso de estruturas relativas constituídas de itens

léxicos animados e inanimados nas línguas naturais, conforme já discutimos.

Para a DLT, os traços semânticos são acessados quando disponíveis no processamento corrente de frases. É preciso haver uma estrutura frasal, e as previsões inferidas a partir da frequência dos *inputs* lexicais poderão guiar o *parser*. Mas não havendo uma estrutura frasal, ou seja, em palavras isoladas, é provável que não haja uma mudança do tempo de leitura apenas com base no traço da animacidade.

Se, por outro lado, houver indícios de que a animacidade muda o tempo de processamento das palavras, mesmo isoladamente, temos que reconsiderar o modelo de processador sintático a partir do qual empreendemos esta pesquisa.

Dito isso, apresentaremos agora o experimento 5, em que testamos o palavras isoladas animadas e inanimadas também por meio da Técnica de Leitura Automonitorada.

4.3.1 Participantes do PB

Os participantes do quinto experimento são todos falantes do PB, foram selecionados 24 estudantes do ensino tecnológico, sendo 10 homens e 14 mulheres, com idade média de 24,6 anos. Todos eles são alunos do Instituto Federal do Rio Grande do Norte, do *campus* Santa Cruz, na cidade de Santa Cruz, RN. A pesquisa aconteceu com a colaboração do professor Iradilson Ferreira da Costa, que nos cedeu o espaço necessário na referida instituição e ainda nos auxiliou na seleção dos alunos que participaram da pesquisa.

A variável dependente foi o tempo de leitura da segunda palavra lida de uma sequência de duas palavras idênticas. A variável independente foi a animacidade da palavra (animada ou inanimada).

Seguem abaixo dois exemplos das condições experimentais. As duas palavras experimentais eram lidas, e posteriormente o participante tinha que responder se tais palavras eram iguais ou não, pressionando a tecla “sim” ou “não”. Foram medidos os tempos de leitura da segunda palavra e da resposta.

52. Ator / Ator

53. Faca / Faca

4.3.2 Materiais

O experimento foi composto por 30 palavras experimentais e 60 distratoras. Entre as experimentais, 15 eram animadas e 15 inanimadas. Desta maneira, cada participante foi exposto 15 vezes às duas condições do teste.

As 30 palavras distratoras tiveram a função de impedir que o participante percebesse um dado padrão e com isso formulasse algum tipo de estratégia de leitura enviesando os resultados do experimento.

Todas as palavras experimentais deveriam receber um “sim” como resposta quanto ao questionamento final. Balanceamos para que metade das respostas fosse “sim” e metade fosse “não” com as frases distratoras. Dessa forma, cada participante foi exposto a 90 palavras, entre distratoras e experimentais.

Foi controlada a frequência, uma vez que usamos as mesmas palavras do experimento anterior, a animacidade e o tamanho dos itens investigados. Todas as palavras eram dissílabas.

4.3.3 Procedimentos

Os sujeitos foram instruídos individualmente sobre as tarefas constitutivas do experimento. As instruções foram apresentadas oralmente pelo experimentador, posteriormente, o participante lia as tarefas na tela do computador em que o experimento foi rodado.

Depois de ter declarado que havia entendido a tarefa, o participante se submetia a um treinamento prático constituído de 4 palavras distratoras. Com isso asseguramos que os participantes compreenderam perfeitamente as tarefas a serem realizadas, que só eram iniciadas quando do consentimento do participante que demonstrava ter de fato compreendido seu papel frente às tarefas experimentais.

Ao pressionar a tecla “espaço”, aparecia a primeira palavra (alvo) no centro da tela do computador, apertando novamente parecia a segunda, no lugar da primeira, que, por sua vez, desaparecia. Depois de apertar a segunda vez o participante deveria responder “sim” no caso de a segunda palavra ser idêntica à primeira, e “não” se não fosse idêntica. A tarefa experimental

consistia, portanto, na Leitura Automonitorada da palavra alvo, seguida de uma decisão do tipo igual/diferente em relação à segunda palavra.

O experimento foi realizado no mês de maio de 2015, conforme já indicamos, no Instituto Federal do Rio Grande do Norte, *campus* Santa Cruz, RN.

O experimento foi rodado em um MacBook, com uma tela de 13". G3, 233Mhz. Utilizamos o programa *Psyscope* (COHEN; MACWHINNEY; FLATT & PROVOST, 1993). Os pares de palavras foram randomizados e o tempo de leitura de cada uma delas, bem como o tempo de resposta, foi aferido em milésimos de segundo.

4.3.4 Resultados

Não houve diferenças significativas entre as condições. Ou seja, a média do tempo de leitura das palavras animadas não apresentou diferença estatisticamente significativa da média do tempo de leitura das palavras inanimadas nos dois segmentos aferidos. Abaixo apresentamos, separadamente, os tempos de leitura da primeira palavra lida e do tempo de resposta.

Como havia somente duas condições experimentais, realizamos teste-T e, a partir dos resultados, obtivemos os gráficos que se seguem:

Tabela 21: Tempo de leitura da primeira palavra

Exp. 5 – Tempo de Leitura da 1ª palavra		
	Substantivo Animado	Substantivo Inanimado
Tempo de leitura	360 (111,39)	372 (99,71)
Desvio Padrão entre Parêntesis		

Na palavra aferida, a média dos tempos de leitura foi praticamente igual para substantivos animados e inanimados, da mesma forma que fizemos com os experimentos anteriores, aqui também retiramos os dados discrepantes por meio do *Box-Plot*. Na condição de substantivo animado, foram 18 itens discrepantes, e na condição de substantivo inanimado, foram 31 itens.

Não houve diferença significativa entre os tempos de leitura dos substantivos animados e inanimados; os animados foram lidos em 360 ms e os inanimados em 372 ms.

Fizemos o teste-T e o P-valor foi de $(t(359) = 1,44 \text{ } p < 0,14)$. Ou seja, a partir da medida *on-line*, a animacidade parece não afetar o tempo de leitura de substantivos quando isolados, fora do contexto frasal.

Tabela 22: Tempo de resposta

Exp. 5 – Tempo de Resposta		
	Substantivo Animado	Substantivo Inanimado
Tempo de resposta	780 (206,26)	764 (184,44)
Desvio Padrão entre Parêntesis		

Posteriormente medimos o tempo de resposta. Ou seja, depois de ler a palavra alvo e a segunda palavra, o participante respondia “sim”, caso as duas fossem iguais, e respondia “não” caso elas fossem diferentes. Todos os pares de palavras experimentais eram formados por duas palavras iguais e, portanto, as respostas deveriam ser “sim” para todas elas. Contudo, com as palavras distratoras, nivelamos de maneira a termos metade das respostas “sim” e metade “não. Duas medidas *off-line* puderam ser computadas, o tempo da resposta em si e o índice de erro. Ambas as medidas foram bastante equilibradas. Houve 14 erros nas respostas dos substantivos animados e 13 erros nos substantivos inanimados, essa diferença não é significativa estatisticamente.

Quanto ao tempo, foram 780 ms para os animados e 764 ms para os inanimados. Ou seja, as medidas *off-line* confirmam a semelhança que há no processamento de palavras animadas e inanimadas. As diferenças de tempo não são significativas. Feito o teste-T, encontramos um P-valor $(t(359) = 0,20 \text{ } p < 0,83)$. Também nesse segmento, antes de submetermos os dados ao teste estatístico, retiramos, por meio do *Box-Plot*, os tempos discrepantes, os *outliers*, que foram 18 no caso dos substantivos animados e 31 nos inanimados.

O fato de não haver diferença significativa no tempo de processamento de substantivos animados e inanimados nos confirma que a diferença

encontrada nas estruturas frasais é pertinente à frase e não à animacidade do substantivo em si.

Visto dessa maneira, os dados do presente experimento nos munem de argumentos no sentido de afirmar que há um acesso à estrutura sintática quando da mudança de custo de processamento das relativas quando considerado o traço semântico da animacidade, já que, *per si*, o traço não muda o processamento de palavras isoladas, fora de um contexto sintático. Logo, a frequência de uso discutida na seção precedente a esta, refere-se ao uso da estrutura sintática cujo antecedente é animado, e não da animacidade vista isoladamente.

No próximo capítulo, diante do estudo de *corpus* empreendido no capítulo 3 desta tese, com base nos dados discutidos neste capítulo e dos pressupostos da DLT, discutiremos o efeito da animacidade no processamento de cláusulas relativas à guisa de uma explicação aos fatos descritos por esta pesquisa.

4.4 Discussão geral

A DLT, conforme assumimos desde o início deste texto, é o modelo teórico de processamento basilar de nossas análises. Um modelo interativo que prevê, no decurso temporal do processamento *on-line*, a interação entre elementos estruturais e não estruturais. A partir da DLT, analisaremos os dados de nossos 5 experimentos do estudo de *corpus* que fizemos.

Para a discussão que iremos fazer, inicialmente repetimos os resultados encontrados para as médias dos tempos de processamento do segmento crítico, tanto para PB, quanto para PE, agora o fazemos em forma de gráfico e juntando os dados de ambas as variedades do português, como podemos ver na figura 13.

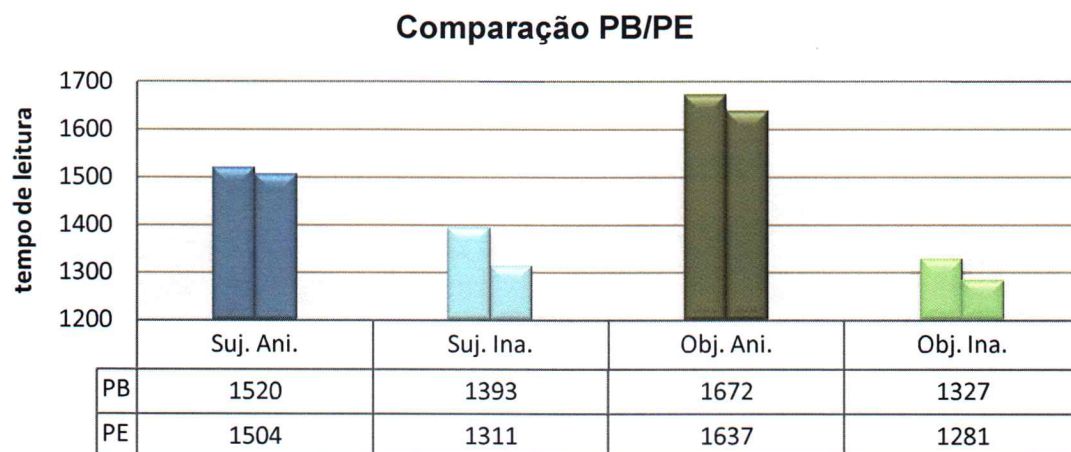


Figura13: Gráfico comparativo – Tempo de Leitura – Segmento Crítico – PB e PE

A seguir, repetimos os resultados dos estudos de *corpora* realizados também em ambas as variedades e as descrevemos em forma de gráfico, como podemos observar na figura 14.

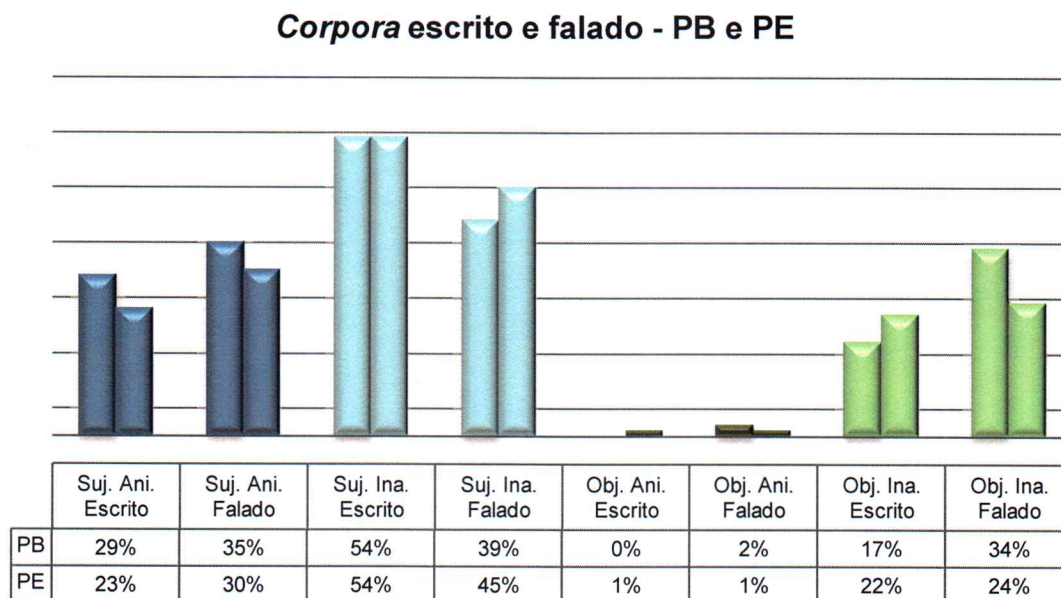


Figura 14: Corpora – escrito e falado, em PB e PE – Comparativo

A primeira constatação que podemos fazer observando ambas as figuras, é que os resultados tanto de PB, quanto de PE são muito semelhantes,

seja nos estudos experimentais, seja no estudo dos *corpora*, apontando uma mesma direção de análise para ambas as variedades.

Como se pode perceber na figura 13, as estruturas CRSi e CROi são processadas mais rapidamente do que as outras condições de forma significativa, tanto em PB, quanto em PE. Ou seja, as estruturas com antecedente inanimado, que geralmente não são usadas nos estudos experimentais que constatarem assimetria, assim como testamos no tipo de relativa dos experimentos 1 e 2, são as mais rápidas em termos de processamento e as CRSa e CROa que são utilizadas nesses mesmos estudos são as mais lentas.

Posto assim, vimos também que não houve efeito principal de tipo de relativa nesses dois experimentos, do que se depreende que, quando somamos as médias dos tempos de todas as CRS e todas as CRO, as referidas relativas não apresentam diferenças significativas entre si. ($F(1,62) = 0,66$; $p < 0,41$) em PB e ($F(1,71) = 1,08$; $p < 0,28$) em PE. Os tempos de leitura do segmento crítico das CRS e CRO estão abaixo representados.

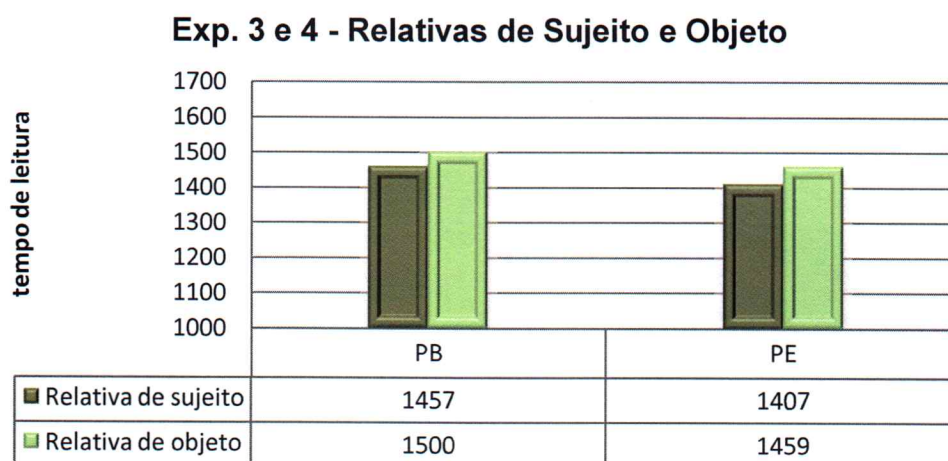


Figura 15- Relativas de sujeito e de objeto

Como se vê, controlada a quantidade de relativas de sujeito e objeto no que tange à animacidade, não mais há uma diferença de tempo de leitura entre essas estruturas. A assimetria em termos de custo de processamento a essas estruturas atribuída pode estar, então, relacionada a fatores diversos e não, necessariamente, às diferenças sintáticas.

Dito isso, nossos dados refutam o que tem sido considerado assimetria entre CRS e CRO. A partir desses mesmos dados, estamos afirmando que, o fato de o traço da animacidade não ter sido controlado em tais estudos, não só enviesou os resultados antes encontrados, como fez da assimetria, tradicionalmente considerada nos estudos aqui citados, uma falsa assimetria, assim como a assimetria que encontramos nos experimentos 1 e 2. O controle de informações não estruturais aqui empreendidos faz com que as considerações preditas por outros estudos pareçam não mais se sustentar.

Se as CRS realmente levassem vantagem por serem mais fáceis de processar, conforme os argumentos da literatura vigente, então por que as CROi teriam sido as estruturas mais rápidas quando comparadas às CRS?

Focalizando agora os estudos dos *corpora*, se observarmos o gráfico da figura 14, percebemos que no caso das relativas de objeto, a CROi é muito mais frequente do que a CROa, quase nula em termos de uso. Já no caso das orações relativas de sujeito, há um percentual de uso significativamente maior para a condição CRSi do que para a condição CRSa, entretanto, ambas são frequentes nos *corpora*.

Se compararmos os gráficos da figura 13 e da figura 14, notamos sem muito esforço que há uma correlação, as condições que são mais frequentes nos *corpora* são mais rápidas em termos de processamento. Isso nos permite levantar a hipótese que em vez da complexidade sintática, o que pode estar influenciando os resultados é a frequência de uso desses tipos de relativa.

Para entendermos melhor a motivação do levantamento dessa hipótese, em termos teóricos, estamos assumindo que uma estrutura sintática é construída no momento inicial do processamento, e que o marcador frasal trabalha com base em previsões sintáticas constituídas no decorrer do processo, serialmente. Mais que isso, em consonância com a DLT, consideramos que a complexidade sintática, que depende, inclusive, da distância ou localidade entre os dois elementos a serem integrados, pode ser relativizada quando guiada por informações não estruturais. Nossos dados, na realidade, apontam para a possibilidade de essa complexidade poder ser até mesmo desconsiderada, quando a frequência de uso for mais relevante na configuração da expectativa de integração do *input* corrente. Sobre expectativa e integração, discorreremos mais adiante, na próxima seção.

Ou seja, o fato de as CRO serem sintaticamente mais complexas não define o cômputo geral do custo de processamento dessas estruturas. A frequência de ocorrência de CRS e de CRO, no que tange à animacidade do item lexical antecedente, pode equilibrar e relativizar o custo total de processamento da cláusula.

Com base nos nossos resultados, há pouco apresentados, podemos perceber com alguma clareza esse equilíbrio do custo de processamento. A DLT prevê que, controladas as frequências de uso de dadas estruturas, a localidade e a distância de aposição serão determinantes no custo de processamento. Os experimentos que trataram do tema não controlaram essas frequências, ao contrário, tais experimentos foram feitos apenas com relativas com sujeito e objeto animados.

Fizemos o mesmo nos experimentos 1 e 2, ao utilizarmos apenas frases com nomes próprios na condição de termos antecedentes relativizados, controlando apenas a frequência e a animacidade dos itens léxicos constitutivos das estruturas aferidas, pois os nomes lá utilizados têm, entre si, mais ou menos a mesma frequência na língua. Algo parecido ocorre nos estudos que focalizam a assimetria em estruturas relativas de sujeito e de objeto em PE e PB já mencionados.

Oliveira (2013, p.77) explica a assimetria com o argumento de que as relativas com extração de sujeito são mais fáceis de processar porque “esse tipo de extração descreve uma cadeia linearmente e estruturalmente menor”. Já Mangas (2011, p. 107) investigou o processamento de relativas com crianças com deficiência auditiva afirmou em seu trabalho que os tempos maiores para a leitura de CRO “parecem estar relacionadas com dificuldades na realização do movimento A-barra. Argumento que corrobora o trabalho de Costa et al (2009), trabalho que, inclusive, é citado por ela.

Ou seja, esses autores explicam a assimetria a partir da estrutura sintática das relativas, no entanto, assim como nos experimentos 1 e 2 desta tese, só utilizaram relativas com antecedentes animados. Mas os dados dos *corpora* mostram que não há uma equivalência na frequência de CRS e CRO na língua como um todo no que tange à animacidade. Logo, utilizamos estruturas com antecedentes animados, que no caso das CRO são estruturas muito raras na língua.

Diante disso, as CRS foram, nas duas variedades do Português testadas naqueles experimentos, significativamente mais rápidas do que as CRO, mas o que estava em jogo, provavelmente, não era complexidade da estrutura, mas a frequência superior das relativas de sujeito animado em detrimento das de objeto animado.

Conforme vimos na seção anterior, nos resultados do 5º experimento, em que fizemos um teste com Leitura Automonitorada de palavras isoladas, não há efeito de animacidade no tempo de leitura de palavras animadas e inanimadas quando lidas isoladamente. Ou seja, como também prevê a DLT, a expectativa de aposição na estrutura em curso é que pode de alguma maneira gerar o custo de processamento. O traço da animacidade, isoladamente, parece não ter esse poder.

Baseando-se nos resultados dos 5 experimentos e nos estudos dos *corpora*, podemos inferir que há claramente acesso à informação sintática relacionada ao tipo de estrutura processada, se relativa de sujeito, ou relativa de objeto, mas também há acesso a informação semântica relacionada à animacidade, já que os resultados apontam diferenças quando se manipula esse traço semântico. Ao acessar essas informações, o *parser* é capaz de também utilizar informação probabilística em termos da frequência de uso das estruturas testadas, o que, como argumentamos, parece influenciar os resultados dos tempos de processamento, seja nos experimentos 1 e 2, em que só temos estruturas com traço animado, seja nos experimentos 3 e 4, em que temos a variável independente animacidade manipulada.

O uso da informação probabilística utilizada pelo *parser* explica os resultados encontrados. CROi é muito mais frequente do que CROa e, por isso, mais rápida nas relativas de objeto, CRSi é mais frequente do que CRSa e, por isso, também é mais rápida em termos de tempo de leitura.

Uma questão que também deve ser explicada é o fato de as estruturas CROi serem mais rápidas de serem lidas do que CRSa. Isso pode ser explicado mais uma vez via frequência de uso, já que enquanto ambas as estruturas relativas de sujeito testadas têm uso frequente na língua, apenas as que não contêm o traço animado nas relativas de objeto testadas têm uso frequente na língua, ou seja, enquanto no processamento das estruturas CRS há competição via frequência, no processamento das CRO, essa competição

se dá de forma diferente, pois, praticamente, como já observamos, só a CROi é usada, por isso a rapidez na leitura desse tipo de estrutura.

A partir dessa proposta de análise, na seção seguinte trataremos de articulá-la com os modelos teóricos de processamento, buscando mostrar porque a explicação via DLT é mais adequada em relação aos resultados que encontramos na tese.

4.5 Nossa análise e os modelos teóricos de processamento

Com base nesses dados apresentados e discutidos, é preciso ser posto que modelos estruturais, a exemplo da TGP (FRAZIER, 1979), não preveem a interferência de traços semânticos na construção da estrutura sintática em curso. Ou seja, nos momentos iniciais, apenas fatores sintáticos são vistos pelo *parser*. Dito isso, a TGP não atende aos questionamentos desta tese quando da análise dos dados que obtivemos em nossos experimentos, já que encontramos efeitos on-line referentes ao traço semântico de animacidade.

A AFS, Estratégia do Antecedente Ativo (FRAZIER, 1987; FRAZIER & FLORES D'ARCAIS, 1989) analisa diferenças entre CRS e CRO exclusivamente a partir de suas respectivas estruturas sintáticas, na distância entre o Antecedente e a lacuna a que esse antecedente endereçado. Em cláusulas relativas, isto significa que o antecedente é relacionado à posição do sujeito, pois é a posição mais anterior possível. Segundo Oliveira (2013, p. 30), “quando o *parser* se depara com um elemento QU movido para uma posição A', na periferia esquerda da oração, esse elemento fica ativo na memória e o *parser* engatilha ou ativa imediatamente uma busca pela primeira posição onde o antecedente ativo possa ser interpretado”. Esta estratégia parece correta para cláusulas relativas de sujeito, mas é problemática em cláusulas relativas de objeto, pois a posição do sujeito já está ocupada. Para nós, é especialmente importante apontar que não há qualquer menção na AFS sobre o acesso a traços de natureza semântica, ou mesmo a interveniência da frequência de uso desses termos. Mak, Vonk e Schriefers (2002) inclusive citam esse princípio e apontam para sua inadequação frente ao fato de a manipulação de um traço semântico acarretar mudanças de processamento.

Frazier e Clifton (1996, 1997) apresentam a hipótese de *Construal*. A princípio, o modelo procura abarcar estruturas não primárias, constituídas por meio de adjunção. Nessa perspectiva, essa hipótese poderia abarcar a análise nas construções relativas aqui estudadas. Inclusive porque, nesse modelo, admite-se que traços semânticos sejam vistos pelo *parser* nas estruturas não primárias. Contudo, a descrição de como esse acesso a fatores distintos dos sintáticos acontece não é clara no modelo, e mesmo se esse acesso é feito *on-line*, ou se numa fase posterior, mais reflexiva. Para além disso, a frequência de uso não é explicitamente admitida nesse modelo, o que o inviabiliza para a análise dos dados presentes nesta tese.

Os modelos de satisfação de restrição (MACDONALD, PEARLMUTTER, SEIDENBERG, 1994) consideram a semântica e a frequência fatores relevantes no processamento linguístico. Visto assim, tais modelos poderiam ser satisfatórios na análise de nossos dados. Todavia, não nos parece que apenas esses fatores estejam em voga. Os modelos de satisfação de restrição concebem que o processamento aconteça paralelamente. Assim, por exemplo, em estruturas ambíguas, o *parser* constituiria múltiplas árvores e no final, vence aquela que melhor se adequar à resolução da ambiguidade. Levy e Gibson (2013) discutem tais modelos e apresentam a pertinência de seus princípios. Contudo, afirmam que não é precisa a localização das dificuldades de processamento na frase. Essa não localização compromete a confirmação empírica de que o processamento possa ocorrer paralelamente, conforme afirmam os modelos de satisfação de restrição.

Embora estejamos, nesta tese, fazendo nossas análises considerando a semântica, ou mais especificamente o traço da animacidade, e a frequência de uso, fazemos isso a partir da estrutura sintática das relativas. Consideramos que, no início do processo, há uma previsão estabelecida. Essa previsão se dá a partir de informações de naturezas distintas, mas existe

Em nota, Kenedy (2015b, p.15) argumenta que Crain e Steedman (cf. 1985) e Altmann e Steedman, (cf. 1988) descaracterizam “a essência do connexionismo como paradigma psicolinguístico, pois com ela se assume que o processamento sintático preceda o processamento semântico”. Segundo esses autores, nos modelos connexionistas, o processador gera todas as representações sintáticas possíveis, para uma posterior inspeção do

processador semântico. Como as frases em um experimento, por exemplo, são apresentadas sem um contexto maior, para o descarte das representações inadequadas, o processador semântico adota uma interpretação padrão, do tipo sujeito/predicado. Se a decisão, numa frase ambígua, se mostrar incorreta, entra-se no *Garden-Path*. “A defesa de um processamento paralelo em sintaxe é a maior fraqueza do paradigma conexionista. Na verdade, a realidade psicológica do processamento em paralelo de indefinidas representações sintáticas complexas nunca foi demonstrada empiricamente” (KENEDY, 2015b, p.15).

Dito isso, não cabe, a esta tese, uma comprovação da pertinência de dos modelos conexionistas, ao contrário, compreendemos que uma representação sintática é necessária nos instantes iniciais da compreensão/produção de frases. O processamento paralelo, portanto, não explica nossos dados porque só temos confirmado o efeito da frequência do traço da animacidade nas estruturas oracionais. Ou seja, o *parser* acessa a frequência e a o traço da animacidade em si na estrutura sintática da relativa e somente nela. O traço da animacidade ou a sua frequência de ocorrência na língua não é acessada isoladamente, ou pelo menos, não temos dados que comprovem isso em nossos experimentos, já que, conforme Levy e Gibson (2013), tais modelos não conseguem explicitar em que lugar da frase ocorre a maior ou menor dificuldade de processamento. Diante dessa realidade, os modelos de satisfação de restrição não atendem às necessidades de nossas análises.

Gibson (1998) propôs a Teoria da Predição Sintática Local, SPLT, teoria segundo a qual, há uma diferença no número de elementos requeridos entre cláusulas relativas de sujeito e cláusulas relativas de objeto. Essa diferença se assemelha ao que se pressupõe na AFS, e tem a ver com a quantidade de elementos que precisam ser mantidos na memória no processamento de CRS e CRO. Para a SPLT, as CRS precisam de apenas dois elementos, e as CRO de três. Como o *parser* preferiria a interpretação da estrutura mais simples, as CRS seriam as preferidas. Mak, Vonk e Schriefers (2002) consideram que essa teoria não prevê a influência de traços semânticos e por isso mesmo dizem que ela não dá conta dos fenômenos estudados por eles. Nossos fenômenos são da mesma natureza, por isso, a SPLT também não nos é interessante como

modelo de processamento. Contudo, a partir da SPLT, Gibson revisa alguns pressupostos importantes e constitui a DLT, Teoria da Dependência Local, modelo que vem ao encontro da nossa análise.

A DLT é um modelo interativo, como tal, concebe um *parser* serial e não modular. Ou seja, sendo serial, considera que o indivíduo constrói estruturas linguísticas conforme os *inputs* vão aparecendo na linearidade do discurso. Por outro lado, não sendo modular, o *parser* faz previsões sintáticas de acordo o material linguístico que vai aparecendo durante o *parsing*, sem, porém, fazer isso de forma encapsulada. Ou seja, as previsões sintáticas não se baseiam apenas em informações estruturais.

“A DLT propõe que o processamento de estruturas sintáticas possua dois componentes fundamentais: *integração* e *expectativa* (KENEDY, 2015, p. 156). Na formulação da teoria, a integração e a expectativa são operações sintáticas constitutivas do processamento e geram custo de processamento. Gibson (2000) admite, porém, que outros fatores estejam dispostos no cardápio de possibilidades, dentre eles, a frequência de uso de itens com seus respectivos traços semânticos. Assim, pode ocorrer, inclusive, um efeito *Garden-Path* via frequência, segundo a DLT. O marcador frasal poderia, por exemplo, prever uma dada estrutura com base em frequência e ela não se confirmar, gerando maior custo de processamento.

Para Gibson (2000), como dissemos, existem os fatores integração e expectativa. “A integração diz respeito à tarefa do processador de relacionar sintaticamente um novo elemento introduzido no *input* linguístico à estrutura sintática já representada na memória de trabalho” (KENEDY, 2015, p. 156). Além disso, o *parser* precisa prever o que vem pela frente, é o componente a que Gibson (2000) se refere como expectativa, que, para Kenedy (2005, p. 156) “relaciona-se à busca por certos elementos sintáticos (verbos, sintagmas nominais etc.)”.

Visto assim, ao que parece, quando o marcador frasal encontra uma relativa com o termo antecedente animado, ele precisa integrar esse item lexical. A distância de aposição até pode ser levada em conta e por isso, constituir uma CRS seria preferível. Lembremos que para a DLT, o *parser* prevê uma e apenas uma estrutura no curso do processamento. Todavia, a frequência de uso da estrutura pode direcionar em sentido diferente, e o que

parece mais fácil do ponto de vista estrutural, aposição mais local, por exemplo, pode ser menos previsível, uma vez que a experiência linguística do falante diante dessa estrutura pode mostrar outro caminho. Ou seja, uma vez que a frequência de uso mostra que o mais fácil pode ser improvável.

Na argumentação de Gibson (2000), ele se vale da assimetria de CRS e CRO, demonstrando a maior rapidez de CRS. Contudo, como fizeram outros autores, ele também elabora seus experimentos e consequentemente seus argumentos, com base na análise de CRSa e CROa. Ou seja, ele prevê que o controle da frequência seja necessário, mas não o faz nos seus experimentos. Visto assim, nossos resultados são analisados com base em seus pressupostos teóricos, mas experimentalmente manipulamos um fator, previsto pela teoria, mas não manipulado no referido trabalho. Nossos resultados, portanto, ampliam os resultados por Gibson (2000) apresentados.

Mais especificamente, estamos afirmando que, embora a DLT admita explicitamente em seus pressupostos que o *parser* acessa informações semânticas nos momentos iniciais do processamento, para esta tese, avaliamos a frequência de uso da estrutura relativa como um todo, e nessa estrutura há um item lexical que varia quanto ao traço da animacidade. Ou seja, nossos dados não apontam para uma interferência direta do traço *per se* no processamento da estrutura, mesmo porque, conforme nos apontou o experimento 5, isoladamente esse traço não gera maior dificuldade ou facilidade de processamento. Mas sim uma influência da frequência de uso dessas estruturas, é a frequência da estrutura na língua que guiaria o *parser*, e não o traço semântico, visto isoladamente.

Talvez pela própria natureza complexa da DLT, no sentido de transitar entre a serialidade e a não modularidade, o modelo não possui, em termos quantitativos, muitos trabalhos que comprovem seus pressupostos nos Brasil. Alguns trabalhos pontuais apresentam o modelo e dispõem sobre a pertinência de sua arquitetura, a exemplo de Kennedy (2015), mas esta seria, a princípio, a primeira tese defendida num programa de Doutorado em Psicolinguística que toma os pressupostos dessa teoria como base teórica de análise na explicação de fenômenos linguísticos.

A frequência, a animacidade e a estrutura da relativa são fatores que estão em discussão para que possamos compreender o processamento

dessas estruturas na mente humana. Mas estamos convictos, também, de que outros fatores não considerados por nós, no presente trabalho, contribuem para o maior ou menor custo de processamento de CRS e de CRO. Isso precisa ser rediscutido, repensado e por meio de outras metodologias experimentais, passar por melhor escrutínio em estudos que ainda hão de se instaurar diante dos resultados ora apresentados.

*O real não está na saída nem na chegada:
ele se dispõe para a gente é no meio da travessia*
– Guimarães Rosa, Grande sertão: veredas.

CAPÍTULO 5

Considerações Finais

5.1 “The cost of the unlikely structure is balanced by the benefits of having a simple structure”

“O custo da estrutura improvável é equilibrado pelos benefícios de ter uma estrutura simples” (TRAXLER, 2012, p. 157), – livre tradução – A partir dessa consideração de Traxler, tomaremos nossos dados para aventarmos algumas inferências relevantes na constituição das explicações que circundam os fenômenos por nós investigados na presente tese.

Uma estrutura será improvável quando sua realização na língua for pouco frequente. Uma estrutura será mais simples quando o custo de integração e armazenamento, nos termos de Gibson (2000), com base na distância de aposição dos itens em questão, for menor. Frente a isso, um modelo que concebe o processamento de uma sentença com base em forças distintas que constituem um algoritmo complexo, como a DLT, descreve e explica os fatos constitutivos de nossos resultados. Esse algoritmo é equacionado e, quando do acesso a vários elementos de diferentes naturezas, o *parsing* da sentença acontece de maneira mais ou menos custosa. Complementamos a assertiva de Traxler supracitada afirmando que a frequência pode ser algo surpreendente na maneira que age sobre o processamento. Quando o autor trata de equilíbrio, consideramos que a frequência pode mesmo anular uma possível opção do *parser* por uma estrutura sintaticamente mais simples.

Trazendo a discussão para a empiria empreendida em nossos experimentos, confrontada com o estudo de *corpus* também constituído para esta tese, trazemos, a seguir, algumas considerações que confirmam os pressupostos da DLT e consubstanciam, em forma de síntese, as suas predições.

1. Estruturas relativas com antecedentes inanimados, tanto de sujeito quanto de objeto são mais frequentes na língua portuguesa em ambas as variedades estudadas (PB e PE);
2. As CRO são, na grande maioria, constituídas por antecedentes inanimados;
3. As CROa são muito raras na língua portuguesa: (1%) das ocorrências;
4. Relativas de sujeito são mais fáceis de processar do que as de objeto quando os termos relativizados são todos animados, mas isso não está relacionado à estrutura propriamente dita;
5. Quando o traço da animacidade é controlado, a facilidade maior de CRS comparada à CRO é revista;
6. As CROi são as estruturas mais fáceis de processar dentre as que foram testadas experimentalmente nesta tese;
7. Parece haver uma relação direta entre frequência de ocorrência de relativas com itens animados e inanimados e a rapidez de processamento dessa estruturas.
8. Palavras isoladas não são lidas mais ou menos rapidamente por conta do traço da animacidade;
9. A assimetria entre CRS e CRO, nos moldes apontados pela literatura vigente, é uma ilusão provocada pelo não controle do traço da animacidade nos experimentos.

Feitas essas afirmações, resta-nos considerar que sintaxe e frequência atuam em sentenças relativas de sujeito e de objeto, tendo em conta o traço semântico da animacidade.

Como a mente humana acessa o léxico e os traços constitutivos dos itens lexicais ainda não está claro, mas o avanço da Psicolinguística nas últimas décadas vem abrindo novos horizontes. Podemos afirmar que, paulatinamente, aproximamo-nos mais, a cada dia, de explicações empíricas que sustentem ou refutem fatos teóricos desenvolvidos a partir de línguas diversas e sob métodos e teorias diferentes.

A DLT (GIBSON, 2000) nos permite compreender fatos da língua de uma maneira bastante produtiva, por considerar a integração de informações de naturezas distintas entre si.

Ao analisar CRS e CRO, chegamos à conclusão de que a assimetria apontada em muitos estudos linguísticos e psicolinguísticos nas últimas décadas pode ter sido enviesada pela não manipulação do traço da animacidade.

É temerário dizer que tais estudos estavam equivocados, mas é lícito afirmar que as considerações feitas por eles constituem apenas parte dos fatos envolvidos no processamento de relativas de sujeito e de objeto. Dito isso, a assimetria sustentada por tais estudos é posta em xeque frente a nossos resultados no presente trabalho e isso traz implicações diretas na forma como se concebe o processamento de frases, já que as fronteiras entre os modelos mais estruturais, que se apoiam em pressupostos sintáticos, e os modelos mais connexionistas, que se apoiam em fatos de natureza lexical são atenuadas.

A interação entre fatores sintáticos, semânticos ou mesmo de natureza não linguística, como a frequência parece avançar a compreensão que temos em torno das cláusulas relativas de sujeito e de objeto.

O traço da animacidade, quando manipulado, muda os rumos da assimetria entre CRS e CRO. Na realidade, quando equilibradas as quantidades de frases experimentais, e CRSa, CRSi, CROa e CROi são distribuídas equitativamente, a assimetria de que tratam esses estudos desaparece e não há diferença significativa entre a leitura de CRS e CRO.

Um dos fatores que mais pesou nos estudos que abordaram essas estruturas foi o fato de se trabalhar, quase que na totalidade, com estruturas com antecedentes animados. Não se considerou que as CROa podem ser estruturas muito raras nas línguas, pelo menos o é no Alemão, no Holandês e no Português. Ao se trabalhar com estruturas pouco típicas, a frequência de uso parece ter enviesado os resultados.

Nossos experimentos mostraram que a raridade de CROa frente a ocorrência das outras estruturas na língua foi determinante no processamento dessas estruturas. Mais que um simples estranhamento, por se tratar de uma estrutura rara, a frequência de ocorrência parece direcionar, de maneira robusta, as decisões do *parser*.

Temos tido o cuidado em falar de frequência da estrutura, e não do traço da animacidade. A DLT prevê a constituição de uma estrutura sintática, e por isso mesmo, o acesso do parser não é feito, isoladamente, ao traço, mas à estrutura como um todo. Então, é a frequência de ocorrência de CRSa, CRSi, CROa e CROi que faz com que CRSi sejam, dentre elas, as estruturas mais rápidas de se processar, a partir de um algoritmo complexo em que o uso mais ou menos frequente determina o custo de processamento.

Conforme argumentamos no capítulo anterior, a referida frequência acaba por equilibrar quais são as estruturas mais rápidas de serem lidas como também as mais lentas. Dito isso, uma estrutura mais simples, como a relativa de sujeito, que faz uma aposição mais local do que a relativa de objeto, pode ser lida mais lentamente, se a relativa de objeto for mais provável naquele contexto. O traço da animacidade e a frequência que varia de acordo com o tipo de estrutura de que faz parte, acaba por ser um elemento determinante no equilíbrio entre o que é mais ou menos provável e o que é mais ou menos complexo.

Concluimos que a possível assimetria entre CRS e CRO precisa ser observada levando-se em conta a animacidade para que se possa de fato compreender melhor que tipo de informação está sendo utilizada no processamento *on-line* dessas estruturas sintáticas.

Sabemos que as questões levantadas por essa tese estão longe de esgotar as investigações sobre possíveis assimetrias de sujeito e objeto em estruturas relativas, mas também em outros tipos de estruturas sintáticas. Faz-se necessária a ampliação das investigações, e essa ampliação pode se dar a partir da aplicação de técnicas experimentais diferentes da Leitura Automonitorada, como, por exemplo, pelo Rastreamento Ocular, técnica que nos permitiria aferir, com mais precisão, o lugar da frase em que ocorre o custo maior de algumas estruturas em detrimento de outras. A aplicação dos testes aqui executados com a manipulação da variável animacidade em outras línguas também nos parece um caminho produtivo, uma vez que o traço da animacidade parece estar, de alguma forma, ligado à constituição das sentenças nas línguas humanas, e a compreensão disso pode não só ampliar os dados aqui apresentados, como ainda a confirmar ou refutar os pressupostos da DLT em termos de processamento.

Além disso, em nossos experimentos, consideramos relativas explicativas e restritivas com encaixe central, aventamos a ampliação delas para estudos experimentais envolvendo também relativas de encaixe à direita e mesmo outros tipos de estrutura sintática em que têm se encontrado assimetria entre sujeito e objeto, como por exemplo, estruturas interrogativas. Por último, é fundamental estudos de *corpus* em outras línguas para que se possa estabelecer alguma correlação mais segura entre frequência de uso e

processamento de sentenças e em consequência observar se a hipótese explicativa aqui explorada se mantém.

Todas essas questões aqui não respondidas apontam para novos caminhos a serem seguidos na psicolinguística. Desconstruída a noção de assimetria entre CRS e CRO com base em pressupostos exclusivamente sintáticos, acreditamos que outros fatores podem ser explorados, mas sabemos que a inclusão dos efeitos provocados pela frequência levando em conta o traço da animacidade na realização dessas estruturas precisa ser considerada em pesquisas vindouras.

A multiplicidade de fenômenos da mente humana ainda tem muito a nos oferecer e a nos surpreender. A língua guarda algumas das chaves que podem desvendar o emaranhado de construtos ainda a serem desvendados. À psicolinguística cabe um quinhão importante: compreender os microscópicos fenômenos da língua na construção de teorias que venham a explicar a mente humana em sua grandeza. Mesmo que nas limitações.

REFERÊNCIAS

- ABNEY, S. *Parsing by chunks*. In R. BERWICK, S. ABNEY, & C. TENNY (Eds.), *Principle-based parsing*. Dordrecht, The Netherlands: Kluwer, 1991, p. 257-278.
- _____. *Chunks and dependencies: Bringing evidence to bear on syntax, computational linguistics and linguistic theory*. Stanford, CA: CSLI. 1995.
- AGUIAR, A. G. G. *Marcação de Definitude em Sintagmas Nominais de Línguas da Família Tupí-Guaraní*. Dissertação de Mestrado. Universidade de Brasília. 2013.
- ALBUQUERQUE, G. S. *Processamento da Linguagem no Déficit de Atenção e Hiperatividade*. Tese de Doutorado. Universidade Federal do Rio de Janeiro. 2008.
- ALTMANN, G. & STEEDMAN, M. Interaction with context during human sentence processing. *Cognition*, v. 30, 1998, p. 191-238.
- BEVER, T. G. The cognitive basis for linguistic structures. In: HAYES, R. (Ed.). *Cognition and language development*. New York: Wiley & Sons, 1970, p. 279-362.
- BEZERRA, G. *O processamento de argumentos e adjuntos em estruturas sintáticas sem ambiguidade de em estruturas com correferência*. Dissertação de Mestrado, UFPB, 2013.
- BEZERRA, G; LEITÃO, M. M. O processamento de argumentos e adjuntos sem ambiguidade estrutural. *Veredas on-line – atemática – Juiz de Fora*, v. 17, 2013, n.2, p. 60-82.
- BRAGA, M. L; KATO, M. A & MIOTO, C. As Construções-Q no Português Falado. In: KATO, M. A; NASCIMENTO, M. *Gramática do Português Culto Falado no Brasil*. Campinas: Editora da Unicamp, 2009.
- BUSNELLO, R. H. D. Passos e Listas na Investigação do *Priming* Ortográfico. *Psicologia: Reflexão e Crítica*, v. 24, 2010, p. 352-354.
- CAPLAN, D. & FUTTER, C. Assignment of thematic roles by na agrammatic aphasic patient. *Brain and Language*, v. 27, 1986, p. 117-135.
- CARAMAZZA, A. & ZURIF, E. Dissociation of algorithmic and heuristic processes in sentence comprehension: Evidence from aphasia. *Brain and Language*. V. 3, 1976, p. 572-583.
- CARVALHO, D. S. Traços. In: FERRARI-NETO, J.; SILVA, C. R. T. (Org.). *O Programa Minimalista em Foco*. Curitiba: CRV, 2012. p. 113-132.
- CHOMSKY, N. *O programa minimalista*. Cambridge, MA: MIT Press, 1995.

CHURCH, K. *On memory limitations in natural language processing*. Dissertação de mestrado. MIT, 1980.

COHEN, J. D., MACWHINNEY, B., FLATT, M., & PROVOST, S. Psyscope: A New Graphic Interactive Environment for Designing Psychology Experiments. *Behavioral Research Methods, Instruments & Computers*. v.25(2), 1993, 257-271.

Corpus Discurso & Gramática. A língua Falada e Escrita na Cidade do Natal. CUNHA, M. A. F. (org). Disponível em: <http://www.discursoegramatican.letras.ufri.br/download/natal.pdf>. Acesso em 17 de maio de 2015.

COSTA, J. et. al. Produção e compreensão de orações relativas em português europeu: dados do desenvolvimento típico, de PEDL e do agramatismo. *XXIV Encontro Nacional da Associação Portuguesa de Linguística*, Lisboa, APL, 2009, pp. 211-224

CRAIN, S. & STEDDMAN, M. *On not being led up the garden-path: the use of context by psychological parser*. IN. DOWTY, D.; KARTUNNEN, L. & ZWICKY, A. (Eds.) *Natural language parsing: psychological, computational and theoretical perspectives*. Cambridge: CUP, 1985.

CRPC, *Corpus de Referência do Português Contemporâneo*. Universidade de Lisboa. Disponível em: <http://www.clul.ul.pt/pt/recursos/183-reference-corpus-of-contemporary-portuguese-crpc>. Acesso em 19 de novembro de 2015.

CUETOS, F.; MITCHELL, D. C. Cross-linguistic differences in *parsing*: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, v. 30, 1988, p. 73-105.

ELLIS, N. C. Frequency Effects in Language Processing. A Review with Implications for Theories of Implicit and Explicit Language Acquisition. *SSLA*, v. 24, 2002, p. 143-188

FARIAS, S. C.; LEITÃO, M. M.; FERRARI-NETO, J. Gênero e número no processamento da anáfora conceitual com nomes coletivos em português brasileiro. *ReVEL*, edição especial n. 6, 2012, 82-109.

FODOR, J. A. *The modularity of the mind*. Cambridge, MA: MIT Press, 1983.

FORD, M. A method for obtaining measures of local *parsing* complexity throughout sentences. *Journal of Verbal Language and Verbal Behavior*, v. 22, 1983, p. 203-218.

FRAUENFELDER, U. ;SEGUI, J.; & MEHLER, J. (1980). Monitoring around the relative clause. *Journal of Verbal Learning and Verbal Behavior*, v. 19, 1980, p. 328-337.

FRAZIER, L. *On comprehending sentences: Syntactic parsing strategies*. Unpublished doctoral dissertation, University of Connecticut, 1979.

_____. Sentence Processing: A Tutorial Review. In: COLTHEART, M (Ed.). *Attention and Performance XII: The Psychology of Reading*. Mahwah, NJ: Erlbaum, 1987, p. 559-586.

FRAZIER, L & CLIFTON, C. *Construal*. MIT Press Cambridge, 1996

_____. *Construal*: Overview, Motivation, and Some New Evidence. *Journal of Psycholinguistic Research*, v. 26, n. 3, 1997, p. 277-295.

FRAZIER, L. & FLORES D'ARCAIS, G.B. Filler-driven *parsing*: A study of gap filling in Dutch. *Journal of Memory and Language*, 28, 1989, 331-344.

FRAZIER, L.; FODOR, J.- The Sausage Machine: A New Two-Stage *Parsing* Model. *Cognition*, v. 6, 1978, p. 291-325.

GARNSEY, S. M., PEARLMUTTER, N. J., MYERS, E., & LOTOCKY, M. A. The Contributions of Verb Bias and Plausibility to the Comprehension of Temporarily Ambiguous Sentences. *Journal of Memory and Language*, v. 37, 1997, 58-93.

GIBSON, E.. Linguistic complexity: Locality of syntactic dependencies. *Cognition*, v. 68, 1998, 1-76.

_____. The dependency locality theory: a distancebased the teory of linguistic complexity. In: MARANTZ, Alec; MIYASHITA, Yasushi; O'NEIL, Wayne. *Image, language, brain*: papers from the first mind-articulation project symposium, Cambridge, Massachusetts, MIT Press, 2000, p. 95-126.

GIBSON, E.; FEDORENKO, E. Weak quantitative standards in linguistics research. Trends in *cognitive sciences*, v. 14, n. 6, 2011, p. 233-234.

GILBOY, E.; SOPENA, J., CLIFTON, C. Jr.; FRAZIER, L. Argument structure and association preferences in Spanish and English complex NPs. *Cognition*, 54, 1995, p. 131-167. GORRELL, Paul. *Syntax and parsing*. Cambridge University Press, 1995.

GRAVINA. A. Sentenças "Garden Path": Orações Relativas Ambíguas e o Princípio do Late Closure. *Revista Gatilho*. V4, 2009.

GRODZINSKY, Y. Agrammatic comprehension of relative clauses. *Brain and Language*. V. 31, 1989, p. 480-499.

HAKES, B., EVANS, J. & BRANNON, L. Understanding sentences with relatives clauses. *Memory and Cognition*. v. 4, 1976, p. 283-296.

HEMFORTH, B., KONIECZNY, L., & SCHEEPERS, C. *Syntactic and anaphoric processes in modifier attachment*. Poster presented at the Ninth Annual CUNY Conference on Human Sentence Processing, New York. 1996.

HICKOK, G., ZURIF, E. & CANSECO-GONZALEZ, E. Structural description of agrammatic comprehension . *Brain na Language*. v. 45, 1993, p. 371-395.

HOLMES, V. M., & O'REGAN, J. K. Eye fixation patterns during the reading of relative-clause sentences. *Journal of Verbal Learning and Verbal Behavior*, v. 20, 1981, p. 417-430.

HOLMES, V. Order of Main and Subordinate Clauses in Sentence Perception. *Journal of Verbal Language na Verbal Behavior*. v. 12, 1973, 285-293.

JORNAL DE NOTÍCIAS. Disponível em: <http://www.jn.pt/paginainicial/>. Acesso em 07 de fevereiro de 2016.

JORNAL TRIBUNA DO NORTE. Disponível em: <http://www.tribunadonorte.com.br/>. Acesso em 06 de setembro de 2014.

JÚNIOR, J. C. L. & CORRÊA, L. S. A natureza do custo computacional na compreensão de passivas: um estudo experimental com adultos. *Letras de Hoje*, Porto Alegre, v. 59, n. 1, 2015, p. 91-101.

JUST, M. A., CARPENTER, P. A., & KELLER, T. A. The capacity theory of comprehension: New frontiers of evidence and arguments. *Psychological Review*, v. 1996, 103, 773-780.

KENEDY, E. Análise de *corpus*, a intuição do linguística e metodologia experimental na pesquisa sobre as orações relativas do PB e do PE. *Revista Linguística*, Rio de Janeiro, v. 4, 2009, p. 30-51.

KENEDY, E. Léxico e computações lexicais. In: FERRARI-NETO, J.; SILVA, C. R. T. (Orgs.). *Programa minimalista em foco: princípios e debates*. Curitiba, PR: CRV, 2012, p. 41-69.

KENEDY, E. *Curso Básico de Linguística Gerativa*. São Paulo: Contexto, 2013.

KENEDY, E. Modelos Interativos no Processamento de Frases: a teoria da dependência local. *Revista ProLingua*, v.10, 2015, p. 150-162.

KENEDY, E.; *Sintaxe, Sintaxes: uma introdução*. São Paulo: Contexto, 2015.

KENEDY, E. *Acesso e integração de informações cognitivas no processamento da compreensão de frases*. Projeto de Pesquisa. UFF. 2015b.

KIMBALL, J. Seven Principles of Surface Structure Parsing. *Cognition*, v. 2, n. 1, 1973, p. 15-47.

KING, J. W., & KUTAS, M. Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, v.7, 1995, p. 376-395.

KING, J., & JUST, M. A. Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, v. 30, 1991, p. 580-602.

LEVY, R., & GIBSON, E. Surprisal, the PDC, and the primary locus of processing difficulty in relative clauses. *Frontiers in Psychology*. v. 4, 2013, p. 01-04.

LOPES, R. V. Traços semânticos na aquisição da linguagem. *Letras de Hoje*. Porto Alegre. v. 41, nº 1, 2006, p. 161-178.

MACDONALD, M. C.; PEARLMUTTER, N.J. & SEIDENBERG, M.S. The Lexical nature of syntactic ambiguity resolution. *Psychological Review*, v. 101, 1994, p. 676-703.

MAIA, M. *Psicolinguística, psicolingüísticas: uma introdução*. São Paulo: Contexto, 2015.

MAK, W. M., VONK, W., & SCHRIEFERS. The influence of animacy on relative clause processing. *Journal of Memory and Language*, v. 47, 2002, p. 50–68.

_____. H. Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*. v. 54, 2006, p. 466-490.

MANGAS, V. L. H. *Compreensão e Produção de Orações Relativas em Crianças Falantes do Português Europeu Portadoras de Deficiência Auditiva*. Dissertação de Mestrado. Universidade Nova Lisboa, 2011,

MARCUS, M., HINDLE, D., & FLECK, M. D-Theory: Talking about trees. In *Proceedings of the Twenty-first Annual Conference of the Association for Computational Linguistics*, Cambridge, MA, 1983, p. 129-136.

MECKLINGER, A., SCHRIEFERS, H., STEINHAUER, K., & FRIEDERICI, A. D. Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition*, v. 23, 1995, 477–494.

MICHAELIS. *Dicionário on-line*. 1998-2009 Editora Melhoramentos Ltda. Disponível em <http://michaelis.uol.com.br/moderno/portugues/index.php?type=Pag=sobre&languageText=portugues-portugues>. Acesso em 15 de dez. 2015.

MITCHELL, D. C. In M. Carreiras & C. Clifton, Jr. (Eds), *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond*. New York, NY: *Psychology*, 2004, p.15-32.

NAIRNE, J. S ET AL. Adaptative Memory: the Mnemonic Value of Animacy. In: *Psychological Science*. Sage, 2013, P. 2099-2105.

OLIVEIRA, F. L. *O processamento da assimetria sujeito-objeto em construções do tipo QU no Português brasileiro: interrogativas e relativas*. Dissertação de Mestrado – Universidade Federal do Rio de Janeiro, 2013.

PERFETTI, C. The cooperative language processors: Semantic influences in na autonomous syntax. In D. A. BALOTA, G.B. FLORES D'ARCAIS, & K.

RAYNER (Eds.). *Comprehension processes in reading*. Hillsdale, NJ: Erlbaum, 1990, p. 205-230.

REVISTA ÉPOCA ON-LINE. disponíveis em <http://epoca.globo.com/>. Acesso em 26 de dezembro de 2015.

RIBEIRO, A. J. C. (1998) Um caso de não aplicação de *Late Closure* no Português do Brasil. Manuscrito. Rio de Janeiro: UFRJ/FL, 1998

UFRJ/FL. SANTOS, T. S. *Animacidade: um estudo entre línguas*. Dissertação de Mestrado Universidade Federal do Rio de Janeiro, 2013.

SCHAFER, A., CARTER, J., CLIFTON, C., JR., & FRAZIER, L. Focus in Relative Clause Construal. *Language and Cognitive Processes*, v. 11, 1996, 135-163.

SCHRIEFERS, H., FRIEDERICI, A. D., & KÜHN, K. The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language*, v. 34, 1995, 499-520.

SENDREINS, A. P. ; SIBALDO, M. A. Estrutura de Constituintes. In: FERRARINETO, J.; SILVA, C. R. T. (Org.). *O Programa Minimalista em Foco*. Curitiba: CRV, 2012. p. 71-111.

STROMSWOLD, K., CAPLAN, D., ALPERT, N. & RAUCH, S. Localization of Syntactic Comprehension by Emission Tomography. *Brain and Language*, v. 52, 1996, p. 452-473.

STURT, P., & CROCKER, M.. Monotonic syntactic processing: A cross-linguistic study of attachment and reanalysis. *Language and Cognitive Processes*, v. 11, 1996, 449-494.

TRAXLER, M. J. (2012). *Introduction to psycholinguistics: Understanding language science*. Chichester, West Sussex: Wiley-Blackwell, 2012.

TRUESWELL, J. C., TANENHAUS, M. K., & GARNSEY, S. M. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, v. 33, 1994, 285-318.

WANNER, E. & M. MARATSOS. Na ATN Approach in Comprehension. In *Linguistic Theory and Psychological Reality*. Ed. HALLE, M. BRESNAN, J & Miller, G. 119-1661. Cambridge, Mass.: MIT Press, 1978.

WATERS, G. S., CAPLAN, D., & HILDEBRANDT, N. Working Memory and Written sentence comprehension. In *Attention and Performance*, v. 12: *The Psychology of reading*, ed COLTHEART, M. 1987 – p. 414-419.

ZUBIN, D. A. Discourse function of morphology: The focus system in German. In T. Givón (Ed.), *Syntax and semantics: Discourse and syntax*. New York: Academic Press, 1979.

ANEXOS 1: Termo de consentimento assinado pelos participantes da pesquisa em Portugal. No Brasil, o mesmo termo foi assinado, porém, com as devidas adaptações ao Português Brasileiro.

TERMO DE CONSENTIMENTO INFORMADO

Eu, _____ fui convidado(a) para participar no estudo de psicolinguística no Instituto de Letras e Ciências Humanas da Universidade do Minho. O estudo em questão é parte integrante da Tese de Doutoramento do investigador Althiere Frank Valadares Cabral, da Universidade Federal da Paraíba (UFPB, Brasil), orientada pelo Professor Doutor Márcio Martins Leitão (UFPB), com a colaboração da Professora Doutora Maria do Carmo Lourenço-Gomes (ILCH-UMinho).

Fui informado (a) quanto ao seguinte:

1. A investigação que se encontra a decorrer está inserida na área da Psicolinguística Experimental, e tem como objectivo estudar o processamento da linguagem humana.

2. A minha participação poderá ocorrer em mais do que um teste, de acordo com o projeto experimental do autor, e segundo minha vontade, sendo que sou informado(a) antecipadamente sobre a minha participação

3. Antes da minha participação só serão revelados o objectivo geral da investigação, os procedimentos a serem realizados e as instruções sobre os mesmos, com o único propósito de não comprometer ou influenciar os resultados do estudo. As informações relativas ao fenómeno que está a ser analisado só podem ser fornecidas no final da investigação.

4. A minha participação é voluntária e isenta de cobranças, e tenho garantia de total acesso aos resultados.

5. Tenho garantida, a qualquer tempo, a liberdade de retirar o meu consentimento para utilização de material gerado pela minha participação e de poder abandonar o estudo em qualquer uma das suas fases sem qualquer prejuízo ou restrição.

6. Tenho conhecimento dos direitos de sigilo absoluto em relação à minha identificação, tornando-se esta, desde já, material confidencial sob responsabilidade do executor do estudo, ALTHIERE CABRAL, que se compromete a utilizar os dados recolhidos somente para a investigação, podendo ser veiculados através de artigos científicos em revistas especializadas e/ou em encontros científicos e congressos, sem tornar possível a minha identificação.

Para outras informações, por favor contacte o autor do estudo através do e-mail althiere@gmail.com .

Braga, ____/____/____

ANEXOS 2: Ficha de informações preenchida pelos participantes da pesquisa em Portugal. No Brasil, o mesmo termo foi assinado, porém, com as devidas adaptações ao Português Brasileiro

FICHA DE INFORMAÇÕES

Está a participar como voluntário num estudo psicolinguístico. Por favor, responda ao questionário abaixo. Todas as informações serão confidenciais. Se precisar de mais espaço, use o verso da folha. Caso cometa algum erro nas perguntas de escolha múltipla, faça um círculo na resposta correta.

Nome (apenas iniciais, se preferir):

Nacionalidade: Naturalidade:

Nacionalidade e naturalidade do pai: Nacionalidade e naturalidade da mãe:

Idade: Sexo: ☐ masculino ☐ feminino.

Está a fazer o curso de licenciatura? ☐ sim ☐ não

Se sim, indique o nome do curso:

Ano:

Já viveu noutra país? ☐ sim ☐ não

Se sim, descreva brevemente onde, quando e por quanto tempo

.....

Fala outra(s) língua(s) além do Português Europeu? ☐ sim ☐ não

Se sim, indique:

1) Língua: ☐ Elementar ☐ Independente ☐ Proficiente

2) Língua: ☐ Elementar ☐ Independente ☐ Proficiente

Já estudou numa escola/universidade na qual a língua falada não era o Português Europeu? Se sim, descreva brevemente onde, quando e durante quanto tempo

.....

Considera que:

Ouve bem em ambiente silencioso? ☐ sim ☐ não

Tem dificuldades para manter a atenção? ☐ sim ☐ não

É muito agitado? ☐ sim ☐ não

É muito quieto? ☐ sim ☐ não

Compreende bem uma conversa:

☐ em ambiente silencioso ☐ em ambiente ruidoso com uma pessoa ☐ em ambiente ruidoso com várias pessoas ☐ oscila independentemente do ambiente

Usa óculos ou lentes de contacto? ☐ sim ☐ não Usa aparelho auditivo? ☐ sim ☐ não

Tem/teve alguma dificuldade de:

(a) fala - ☐ sim ☐ não

(b) leitura e escrita - ☐ sim ☐ não

(c) escolares (de modo geral) ☐

(d) outras

Se respondeu "sim" numa das alíneas acima, recebeu ou recebe tratamento de ☐ terapeuta da fala ☐ neurologista ☐ outro

Por favor, verifique se não deixou questões em branco. Obrigado pela sua participação!

Frases experimentais do Experimento 3.

Frases para o 3º Experimento. São 16 grupos de 4 frases, cada um desses grupos é formado por uma sentença contendo uma relativa de sujeito inanimado, uma de sujeito animado, uma de objeto inanimado e uma de objeto animado. Todas as frases foram divididas em 5 experimentos, medimos o tempo de leitura dos segmentos 2 e 3.

O livro / que a aluna encontrou / esta tarde / na estante / não me serviu de nada.

O gato / que a aluna encontrou / esta tarde / na rua ainda / não comeu a ração dele.

O livro / que encantou a aluna / esta tarde / era sobre / a vida das aves do litoral.

O gato / que encantou a aluna / esta tarde / estava brincando / na grama de eu jardim.

A bola / que o menino procurou / pela manhã / antes do jogo / estava debaixo da cama.

A vaca / que o menino procurou / pela manhã / para tirar leite / tinha fugido da vacaria.

A bola / que magoou o menino / pela manhã / foi atirada pelo jogador / da outra escola.

A vaca / que magoou o menino / pela manhã / estava a / proteger a sua cria.

O texto / que o professor explicou / nessa aula / foi publicado no jornal / da escola desta semana.

O autor / que o professor explicou / nessa aula / com certeza foi / o mais importante.

O texto / que motivou o professor / nessa aula / não era nada interessante / para os alunos.

O autor / que motivou o professor / nessa aula / escreveu livros muito bons / para sua época.

O vinho / que o repórter mencionou / na revista / era de uma região / no sul da Itália.

O padre / que o repórter mencionou / na revista / tinha sido assassinado / há mais de um ano.

O tiro / que atingiu o repórter / da revista / foi disparado depois / da hora do almoço.

O padre / que atingiu o repórter / da revista / naquela igreja / foi inocentado.

A planta / que a aluna pesquisou / na escola / não é encontrada / em nenhum lugar.

A atriz / que a aluna pesquisou / na escola / era uma das mais / talentosas do teatro.

A planta / que inspirou a menina / na escola / era a mesma que / a sua avó / plantou há anos.

A atriz / que inspirou a menina / na escola / foi vista na cerimônia / de lançamento do filme.

O filme / que a cantora comentou / nesta noite / concorreu ao Oscar / de melhor canção.

O ator / que a cantora comentou / nesta noite / na entrevista foi / seu namorado há muito tempo.

O filme / que alegrou a cantora / nesta noite / foi considerado muito / mau pelos críticos.

O ator / que alegrou a cantora / nesta noite / trabalha naquela / telenovela.

O vento / que o marujo enfrentou / no sábado / não foi tão forte / quanto ele queria.

O peixe / que o marujo desejou / no sábado / no seu anzol foi / se embora sem ser fígado.

O vento / que ajudou o marujo / no sábado / a navegar na baía / era mais forte que o esperado.

O peixe / que ajudou o marujo / no sábado / a matar sua fome / foi a sua salvação.

A arma / que o soldado segurou / na verdade / pertencia / à corporação policial.

A gata / que o soldado segurou / na verdade / pertencia / à sua esposa.

A arma / que arranhou o soldado / na verdade / era considerada / inofensiva.

A gata / que arranhou o soldado / na verdade / não era assim tão agressiva.

A casa / que o médico conheceu / na cidade / é mais simples / do que a que ele mora.

A freira / que o médico conheceu / na cidade / estava esperando / alguém para ajudá-la.

A casa / que abrigou o médico / na cidade / neste verão ainda / está para ser alugada.

A freira / que abrigou o médico / na cidade / foi curada com / um tratamento que ele fez.

A caixa / que a vizinha devolveu / nessa noite / estava com / defeito de fabricação.

A cabra / que a vizinha devolveu / nessa noite / não dava / muito leite.

A caixa / que assustou a vizinha / nessa noite / foi enviada a ela / por brincadeira.

A cabra / que assustou a vizinha / nessa noite / estava atrás / de um arbusto.

A lição / que a criança recebeu / nessa hora / representa algo importante / no seu caráter.

A mulher / que a criança recebeu / nessa hora / fará parte de sua vida / para sempre.

A lição que / convenceu a criança / nessa hora / pode ser que não / a convença quando adulta.
A mulher / que convenceu a criança / nessa hora / será lembrada / por ela quando for adulta.

O beijo / que a médica desejou / no inverno / era uma / ilusão passageira.
O homem / que a médica desejou / no inverno / nunca apareceu / na vida real.
O beijo / que iludiu a médica / no inverno / não representará mais / que uma lembrança.
O homem / que iludiu médica / no inverno / queria levá-la / para sua casa.

O prédio / que o piloto avistou / nessa guerra / foi alvo de / muitas explosões.
O índio / que o piloto avistou / nessa guerra / ficou preso na tribo / durante alguns anos.
O prédio / que acolheu o piloto / nessa guerra / foi totalmente / destruído depois da batalha.
O índio / que acolheu o piloto / nessa guerra / ficou seu amigo / para o resto da vida.

O crime / que o professor criticou / esta noite / realmente deixou / todos indignados.
O filho / que o professor criticou / esta noite / era um adolescente / problemático.
O crime / que condenou o professor / esta noite / não foi investigado / como deveria.
O filho / que condenou o professor / esta noite / na verdade só / queria provocá-lo.

A carta / que a médica procurou / nessa quinta / tinha sido queimada / dias antes.
A moça / que a médica procurou / nessa quinta / trabalhava numa / lanchonete muito simples.
A carta / que acusou a médica / nessa quinta / tinha manchas / de sangue da vítima.
A moça / que acusou a médica / nessa quinta / não voltou ao hospital / no outro dia.

A faca / que a morena escondeu / na floresta / foi encontrada dias / depois pelos policiais.
A índia / que a morena escondeu / na floresta / tinha ficado grávida / pela segunda vez.
A faca / que machucou a morena / na floresta / desapareceu depois / do ocorrido.
A índia / que machucou a morena / na floresta / estava muito assustada / quando chegamos.

Distratoras

1. O ventilador / de teto / do meu quarto / está parado. Telhado
2. Aquele menino / andando de bicicleta / deve ter / uns oito anos. Idade
3. Gosto de colecionar / muitos livros / sobre assuntos / bem variados. Caderno
4. O banheiro do meu quarto / tem uma janela / que fica aberta / a noite inteira. Porta
5. A menina / queria ganhar / uma pônei rosa / naquela loja. Brinquedo
6. Hoje de manhã / o creme dental / tinha acabado / na farmácia. Dente
7. As flores / de plástico / não murcham / mas são belas. Rosas
8. Comprei um creme / hidratante / para pés ressecados. Pele
9. Pretendo viajar / para Portugal / nas férias / do fim do ano. Feriado.
10. Gosto de comprar / perfume francês / quando viajo / a Paris. Turista
11. Tenho um mapa / da Europa / na parede / do meu escritório. Continente
12. A tesoura / de cabo branco / está em cima / da escrivaninha. Escritório
13. O sorvete / de morango / combina / com calda de chocolate. Baunilha
14. A lâmpada / do meu quarto / não ilumina / o suficiente. Claridade
15. O campo de tulipas / é muito colorido / mesmo assim / me traz tristeza. Cores
16. A aranha / desce devagar / pela parede / procurando comida. Teias
17. Meu filho / tem um dinossauro / de brinquedo / verde e marrom. Brincadeira
18. Aquela toalha / que está no banheiro / ainda está / muito molhada. Sabão
19. Hoje de manhã / tinha uma senhora / muito elegante / na piscina. Biquíni
20. As novelas / sempre trazem / personagens / muito maus. Atriz
21. Na minha casa / moram quatro pessoas / todas elas / são felizes. Família
22. A capa do livro / nem sempre / traduz / seu conteúdo. Escrita
23. Quando ela chega / um perfume / muito bom / invade o ambiente. Cheiro.
24. Hoje vamos comprar / um par de sapatos / bem finos / para festa. Baile.
25. O árbitro / duvidou / ao aplicar / o cartão vermelho. Futebol.
26. O ator / cumpriu / o contrato / integralmente. Emprego.
27. Apesar de atrasado / o enfermeiro / atendeu / toda a demanda. Hospital.
28. A chuva / causou / danos irreparáveis / em toda cidade. Tempestade.
29. Os muros do condomínio / precisarão / ser reforçados / urgentemente. Segurança.
30. A cirurgia / foi realizada / em uma sala / improvisada. Doente.
31. O livro / ficou mais longo / quando a versão / impressa foi lançada. Impressão.

32. O aluno / procurou / a professora / após seu expediente. Escola.
33. O tempo seco / prejudica / as plantações / anuais de milho. Lavoura.
34. O centro comercial / será fechado / devido / ao mau tempo. Compras.
35. As tintas / perderam a validade / antes do previsto / por causa do frio. Inverno.
36. O incêndio / foi visto / a centenas de quilômetros / por causa da fumaça. Fogo.
37. A cômoda / foi montada / inesperadamente / pelo morador. Cômoda.
38. A coluna vertebral / do paciente / precisou ser imobilizada / para a cirurgia. Paciente.
39. Ao levantar voo / o piloto percebeu / uma instabilidade / na asa do avião. Piloto.
40. A poltrona do ônibus / foi danificada / devido ao mau uso / dos passageiros. Ônibus.
41. Ao quebrar a pia do banheiro / a faxineira / machucou-se / seriamente. Banheiro.
42. Considerando / as condições climáticas / o alpinista / decidiu adiar a escalada.
Alpinista.
43. O ataque de cupim / comprometeu / o telhado / de toda a casa. Cupim.
44. O calor excessivo / fez o patinador / passar mal / durante a apresentação. Patinador.
45. A cama / tem um lençol / de linho azul / com flores claras. Lençol.
46. A bicicleta / foi quebrada / ao ser / atingida pelo bonde. Bicicleta.
47. O volume / das caixas de som / do evento / precisou ser revisto. Som.
48. A banda / sentiu-se ofendida / pelo pequeno público / que foi ao show. Banda.

Para o experimento 4 usamos as mesmas frases, tendo feito apenas modificações relativas às diferenças entre dialetais do PB e do PE.