

UNIVERSIDADE ESTADUAL DE CAMPINAS

Este exemplar é a redação final da tese
defendida por Matilde Virginia
Ricardi Scaramucci

e aprovada pela Comissão Julgadora em
10, 03, 95

M. Cavalcanti
PROFA. DRA. MARILDA DO C. CAVALCANTI

**O PAPEL DO LÉXICO NA COMPREENSÃO EM LEITURA EM LÍNGUA
ESTRANGEIRA: FOCO NO PRODUTO E NO PROCESSO**

Matilde Virginia Ricardi Scaramucci

**Tese apresentada ao Departamento de Lingüística do Instituto de Estudos da
Linguagem da Universidade Estadual de Campinas, como requisito parcial para
a obtenção do grau de Doutor em Ciências.**

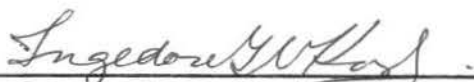
Campinas, janeiro de 1995



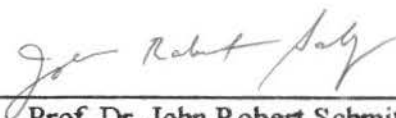
BANCA EXAMINADORA



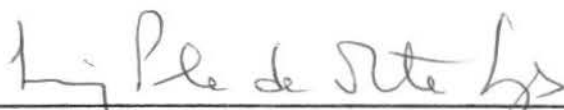
Profa. Dra. Marilda do Couto Cavalcanti
(Orientadora)



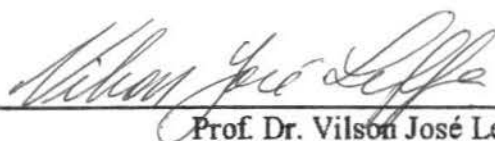
Profa. Dra. Ingedore Grundfeld Villaçã Koch



Prof. Dr. John Robert Schmitz



Prof. Dr. Luiz Paulo da Moita Lopes



Prof. Dr. Vilson José Leffa

AGRADECIMENTOS

Foram muitos os que me ajudaram a concluir este trabalho.
Meus sinceros agradecimentos

À Marilda, minha orientadora, pela paciência, dedicação, eficiência e principalmente incentivo, nos momentos de desânimo;

Ao Dr. Charles Alderson, pelo trabalho eficiente de co-orientação, especialmente na definição do desenho da pesquisa e elaboração dos instrumentos de avaliação;

Ao John Schmitz e à Ingedore Koch, pelas sugestões feitas por ocasião da qualificação;

Ao José Antonio, por me fazer ver os caminhos da carreira acadêmica e do uso do computador;

À Chiara e ao Caio, por me fazerem lembrar que existiam outras coisas importantes, além deste trabalho;

À minha mãe e ao meu pai (*in memoriam*), por terem sempre acreditado em mim;

Ao meus tios Elias (*in memoriam*) e Alzira (*in memoriam*), meus segundos pais, por terem, dentre outras coisas, me ensinado as primeiras letras;

Aos colegas do Departamento de Linguística Aplicada e do CEL, pela sobrecarga causada pelos cursos que deixei de dar;

À Profa. Dra. Regina Moran, pela boa vontade e eficiência com que acolheu meu problema estatístico;

À Maria Teresa Villalobos-Aguayo e ao Sérgio Siade, pela paciência na assessoria estatística;

À CAPES e ao FAEP/UNICAMP, pelos auxílios parciais a esta pesquisa, mais especificamente pela bolsa "sanduíche" na Inglaterra e pagamento de taxas à Universidade de Lancaster, respectivamente;

Aos funcionários do IEL, pela atenção e atendimento sempre eficiente;

Aos alunos-sujeitos, pela paciência e cooperação durante a exaustiva coleta de dados;

Aos meus alunos de Inglês Instrumental, pela motivação na escolha do tópico desta tese.

RESUMO

Este trabalho visa investigar, na perspectiva de um modelo de leitura interativo, o papel do vocabulário na compreensão em leitura em um contexto de língua estrangeira (LE doravante). Mais especificamente, seu objetivo é avaliar a competência lexical em inglês de alunos brasileiros em uma universidade estadual paulista e tentar relacionar essa competência à sua habilidade de leitura de textos gerais em inglês. Embora a literatura saliente a importância dessa contribuição, não há dados conclusivos referentes à natureza dessa relação, principalmente em LE. A evidência de um possível nível limiar de competência lexical é investigada, assim como as implicações desse limiar para o processo e o produto da compreensão. Considerando-se a complexidade do processo de leitura, ou seja, um processo em que há a interação de vários componentes, este estudo não poderia deixar de incluir dados sobre outros aspectos, tais como conhecimento prévio ou conceitual, proficiência em leitura em língua materna (LM doravante) e concepções do leitor sobre o processo de leitura. O problema é investigado sob dois focos diferentes, ou seja, de produto e de processo, envolvendo análises de ênfase quantitativa e qualitativa, que se complementam. Os dados de ênfase quantitativa, com foco na compreensão enquanto produto, considera 49 sujeitos e foram coletados através de questionários, testes de vocabulário, de leitura e de conhecimento conceitual. Os testes de vocabulário foram elaborados a partir de um conceito *rico* de vocabulário (Richards 1976). Esse conceito considera o vocabulário em sua extensão e profundidade, envolvendo vários componentes, que podem ser lingüísticos, psicolingüísticos e sociolingüísticos, incluindo também uma dimensão de uso. Os resultados desse estudo mostram uma relação forte entre competência lexical e compreensão em leitura, que pode ser observada através de correlações e comprovada através da análise fatorial. Os dados de ênfase qualitativa com foco no processo da compreensão foram obtidos através de protocolos de leitura de três sujeitos, com diferentes graus de proficiência em leitura, selecionados a partir do estudo de produto. Esses dados mostram que as pausas no processo de construção de sentido do texto dos três leitores são detonadas por problemas que, na superfície, podem ser considerados lexicais, ou de acesso a significados de palavras, mas que uma análise mais profunda mostram estarem relacionados também à dificuldades nos níveis sintático, morfológico, fonológico, semântico e pragmático-discursivo. A competência lexical mais desenvolvida do leitor intermediário avançado deste estudo lhe permite utilizar dos vários recursos (definidos como a interação entre fontes/níveis de conhecimento e as restrições do texto ou pistas) para o estabelecimento do contexto do texto e do contexto geral, na maior parte das vezes de forma automática, o que lhe possibilita construir o contexto de forma controlada, quando necessário, não apenas para a inferência de palavras, mas, principalmente, para a construção de um sentido para o texto, ou para processamentos de nível mais alto. A competência lexical menos desenvolvida dos outros dois leitores, por outro lado, caracterizada por dificuldades nos vários níveis acima identificados, leva-os a depender da construção de um contexto de forma controlada para o acesso às palavras. Como estas são muito numerosas, toda sua capacidade de processamento acaba sendo consumida nesse processo, e seu foco, na leitura, acaba sendo a inferência ou decodificação dessas palavras, e não a construção de um sentido para o texto. Os dados dos dois estudos mostram que o conhecimento conceitual, a leitura em LM assim como diferentes concepções de leitura, embora importantes, não parecem compensar, sem maiores problemas, lacunas na competência lexical, evidenciando, portanto, um nível limiar de competência lexical que é relativo, uma vez que parece variar de acordo com o texto e propósitos de leitura. Entretanto, embora uma competência lexical bem desenvolvida seja uma condição necessária para uma interpretação pragmática, ou para a negociação de sentidos que caracteriza uma leitura crítica, ela não é suficiente. Nesse sentido, fatores pragmáticos, relacionados à situação de leitura são fundamentais. Este estudo oferece: contribuições teóricas, relacionadas ao conceito de competência lexical; contribuições metodológicas, que dizem respeito principalmente ao seu desenho de pesquisa inovador e abrangente; e, finalmente contribuições práticas, vistas enquanto um conjunto de subsídios para o ensino e avaliação da leitura e do vocabulário em uma situação de LE.

SUMÁRIO

CAPÍTULO I – INTRODUÇÃO	1
1.1 – O contexto de pesquisa e justificativa da escolha do tópico	1
1.2 – O problema de pesquisa	3
1.3 – As questões de pesquisa	6
1.4 – A estrutura da tese	7
Notas	7
CAPÍTULO II – A COMPETÊNCIA LINGÜÍSTICA NAS PESQUISAS EM LEITURA EM LE	9
2.1 – As pesquisas sobre leitura em LE	9
2.2 – Os modelos/visões de leitura	11
2.2.1 – O modelo ascendente ou de decodificação: uma visão de leitura como extração de significado	11
2.2.2 – O modelo descendente ou psicolinguístico: uma visão de leitura como (re)criação de significado	13
2.2.3 – O modelo interativo: uma visão de leitura como construção de significado	18
2.3 – O componente lingüístico vs. cognitivo nas pesquisas em leitura em LE	25
2.3.1 – O componente cognitivo no ensino da leitura: um foco em estratégias	43
2.4 – O componente lexical	51
2.4.1 – O vocabulário: uma perspectiva histórica	51
2.4.2 – O vocabulário: a pesquisa sobre ensino e aquisição	55
2.4.3 – O vocabulário e a leitura	58
2.4.3.1 – A contribuição da leitura para a aquisição do vocabulário: uma abordagem indireta do vocabulário através da leitura	58
2.4.3.2 – O componente lexical nas pesquisas em leitura	61
2.4.3.3 – A competência/conhecimento lexical e sua avaliação	69
Notas	84
Citações originais	88
CAPÍTULO III – O DESENHO E A METODOLOGIA DE PESQUISA	94
3.1 – O desenho de pesquisa: a escolha da metodologia e sua justificativa	94
3.1.1 – A distinção quantitativo/qualitativo como paradigmas de pesquisa	96
3.1.2 – A distinção produto/processo na pesquisa em LA	99
3.1.3 – As dicotomias qualitativo/quantitativo e processo/produto neste trabalho	103
3.2 – Fase I: o estudo de produto com ênfase quantitativa	106
3.2.1 – Seleção e descrição dos sujeitos	106
3.2.2 – Procedimentos e instrumentos de pesquisa	107
3.2.2.1 – Questionário I	109
3.2.2.2 – Tarefa de leitura em inglês e questionários II e III	109
3.2.2.3 – Testes de vocabulário	110
3.2.2.4 – Teste de conhecimento conceitual	113
3.2.2.5 – Testes de leitura	113
3.2.2.5.1 – Cloze I	116
3.2.2.5.2 – Cloze II	117
3.2.2.5.3 – Testes de leitura (resumo/perguntas de respostas abertas)	118
3.2.2.6 – Escores do vestibular	119
3.2.3 – Técnicas estatísticas	120
3.2.3.1 – Análise de correspondência	120
3.2.3.2 – Análise de correlação	122
3.2.3.3 – Análise fatorial	122
3.3 – Fase II: o estudo de processo com ênfase qualitativa	123

3.3.1 – Seleção e descrição dos sujeitos	123
3.3.2 – Descrição dos procedimentos e instrumentos de pesquisa	123
3.3.3 – Coleta de dados	124
Notas	125
Citações originais	126
CAPÍTULO IV – O VOCABULÁRIO E A COMPREENSÃO ENQUANTO PRODUTO DA LEITURA: ANÁLISE E DISCUSSÃO DOS DADOS DE ÊNFASE QUANTITATIVA	128
4.1 – Preparação dos dados para a análise	129
4.1.1 – Perfil	129
4.1.2 – Vocabulário	129
4.1.3 – Conhecimento conceitual ou prévio	130
4.1.4 – Metacognição/concepções de leitura	131
4.1.5 – Leitura em LM e LE e proficiência em LE	131
4.2 – Análise dos dados	133
4.2.1 – Análise intracomponentes	134
4.2.1.1 – Perfil	134
4.2.1.2 – Vocabulário	137
4.2.1.2.1 – Vocabulário: foco no grupo	137
4.2.1.2.2 – Vocabulário: foco nos sujeitos	145
4.2.1.2.3 – Vocabulário: foco nas palavras	146
4.2.1.3 – Conhecimento conceitual ou prévio	151
4.2.1.4 – Metacognição/concepções de leitura	152
4.2.1.5 – Leitura em LM e LE e proficiência em LE	159
4.2.2 – Análise intercomponentes	161
4.2.2.1 – Análise de correlação	162
4.2.2.2 – Análise fatorial	174
4.3 – Discussão dos resultados	176
CAPÍTULO V – O VOCABULÁRIO E A COMPREENSÃO ENQUANTO PROCESSO DA LEITURA: ANÁLISE E DISCUSSÃO DOS DADOS DE ÊNFASE QUALITATIVA	190
5.1 – Perfil dos sujeitos	190
5.2 – Análise dos protocolos	194
5.3 – Discussão da análise dos protocolos	236
Notas	254
CAPÍTULO VI – CONCLUSÕES	255
6.1 – Resumo das conclusões	255
6.2 – Discussão comparativa	256
6.3 – Limitações	264
6.4 – Contribuições	266
6.4.1 Contribuições teóricas	267
6.4.2 Contribuições metodológicas	267
6.4.3 Contribuições práticas	268
Citações originais	275
APÊNDICE A – QUESTIONÁRIOS	276
A.1 – Questionário I	276
A.2 – Questionário II	276
A.3 – Tarefa de leitura	277
A.4 – Questionário I, II e III (parte A) preparados para a análise	278
A.5 – Questionário III	279
A.6 – Questionário III preparado para a análise	281
APÊNDICE B – INSTRUMENTOS/PROCEDIMENTOS PARA A AVALIAÇÃO DO VOCABULÁRIO	283

B.1 – Texto-base para testes de vocabulário	283
B.2 – Procedimento para seleção de palavras para os testes de vocabulário	283
B.3 – Instruções para testes de seleção de palavras	284
B.4 – Número de alunos relatando palavras desconhecidas	285
B.5 – Teste de vocabulário I (Familiaridade)	286
B.6 – Teste de vocabulário II (Frequência)	286
B.7 – Subteste de vocabulário III-A (Derivação/flexão)	287
B.8 – Subteste de vocabulário III-B (Significado/definição)	288
B.9 – Subteste de vocabulário III-C (Associação)	288
B.10 – Subteste de vocabulário III-D (Colocação)	289
B.11 – Subteste de vocabulário III-E (Contexto/classe gramatical)	289
B.12 – Teste de vocabulário IV (Inferência em contexto)	292
B.13 – Testes de vocabulário categorizados para a análise	293
B.14 – Testes de vocabulário recategorizados após primeira análise	294
APÊNDICE C – INSTRUMENTO PARA A AVALIAÇÃO DO	
CONHECIMENTO CONCEITUAL	295
C.1 – Teste de conhecimento conceitual	295
C.2 – Teste de conhecimento conceitual categorizado para análise	295
C.3 – Teste de conhecimento conceitual recategorizado após primeira análise	295
APÊNDICE D – INSTRUMENTOS PARA A AVALIAÇÃO DA LEITURA EM LM	
E LE E PROFICIÊNCIA EM LE	297
D.1 – Texto-base para Cloze I	297
D.2 – Cloze I	298
D.3 – Texto-base para Cloze II	299
D.4 – Cloze II	299
D.5 – Teste de leitura I	300
D.6 – Teste de leitura II	301
APÊNDICE E – TEXTOS-BASE PARA INTROSPECÇÃO	303
E.1 – Texto-base para familiarização com tarefa de introspecção	303
E.2 – Texto-base para tarefa de introspecção	303
E.3 – Texto-base em sua estrutura problema/solução	304
APÊNDICE F – OUTROS QUADROS/FIGURAS/TABELAS	306
APÊNDICE G – PROTOCOLOS	313
G.1 – Protocolo do leitor intermediário avançado	313
G.2 – Protocolo do leitor intermediário	317
G.3 – Protocolo do leitor iniciante	320
BIBLIOGRAFIA	327

LISTA DE QUADROS

Quadro 2.1 – Níveis de co-texto em torno da palavra desconhecida (Scott 1992)	30
Quadro 2.2 – Taxonomia de fontes de conhecimento (Haarstrup 1987)	31
Quadro 2.3 – Elementos do contexto (Henriques 1990)	32
Quadro 2.4 – Fontes de conhecimento mais usadas (Haarstrup 1989a)	65
Quadro 2.5 – Conhecimento lexical (Dale 1965)	73
Quadro 2.6 – Modelo de conhecimento da palavra (Haarstrup 1989a)	74
Quadro 2.7 – Conhecimento lexical (Cronbach 1942)	75
Quadro 2.8 – Conhecimento lexical (Read 1987)	76
Quadro 2.9 – Pressuposições sobre o conhecimento lexical do falante nativo (Richards 1976)	78
Quadro 2.10 – Conceito <i>rico</i> de vocabulário (Richards 1976)	79
Quadro 2.11 – Conhecimento da palavra (Nation 1984)	80
Quadro 4.1 – Estratégias/Concepções de leitura <i>local</i> e <i>global</i>	158
Quadro 5.1 – Estrutura problema/solução do texto-base	198
Quadro 5.2 – Itens lexicais chaves e cadeias coesivas	199
Quadro F.1 – Contextos vs. classes gramaticais	306
Quadro F.2 – Familiaridade vs. frequência	307

LISTA DE TABELAS

Tabela 2.1 – Questionário metacognitivo (Carrell 1989)	50
Tabela 3.1 – Instrumentos de coleta e seus objetivos	108
Tabela 3.2 – Cronograma da coleta de dados	108
Tabela 3.3 – Testes de vocabulário e seus focos de avaliação	111
Tabela 3.4 – Testes elaborados de acordo com um conceito <i>rico</i> de vocabulário	112
Tabela 3.5 – Itens (questões) dos testes de leitura I e II e seus objetivos	119
Tabela 4.1 – Vocabulário	130
Tabela 4.2 – Metacognição/Concepções de leitura	131
Tabela 4.3 – Valores dos itens dos testes de leitura I e II	132
Tabela 4.4 – Leitura	133
Tabela 4.5 – Proficiência em LE	133
Tabela 4.6 – Perfil do sujeitos: dados de identificação e <i>background</i> linguístico	135
Tabela 4.7 – Perfil dos sujeitos: leitura	136
Tabela 4.8 – Perfil dos sujeitos: avaliação da leitura	136
Tabela 4.9 – Familiaridade	138
Tabela 4.10 – Frequência	138
Tabela 4.11 – Teste de conhecimento <i>rico</i> de vocabulário	140
Tabela 4.12 – Inferência em contexto	141
Tabela 4.13 – Testes de vocabulário em ordem crescente de dificuldade	141
Tabela 4.14 – Familiaridade vs. frequência	143
Tabela 4.15 – Familiaridade vs. significado	143
Tabela 4.16 – Frequência (por palavra)	147
Tabela 4.17 – Palavras avaliadas, classificadas por falantes nativos/não nativos em um contínuo de frequência	148
Tabela 4.18 – Familiaridade (por palavra)	149
Tabela 4.19 – Conhecimento conceitual (por conceito)	151
Tabela 4.20 – Estratégias de confiança	155
Tabela 4.21 – Estratégias de correção ou reparo	155
Tabela 4.22 – Aspectos subjacentes às estratégias de eficiência e dificuldade	156
Tabela 4.23 – Estratégias de eficiência	156
Tabela 4.24 – Estratégias de dificuldade	158
Tabela 4.25 – Correlações entre variáveis do vocabulário	162
Tabela 4.26 – Correlações entre variáveis de LM	163
Tabela 4.27 – Correlações entre variáveis de leitura em LE	164
Tabela 4.28 – Correlações entre duas medidas de proficiência em LE	165
Tabela 4.29 – Correlações entre variáveis do conhecimento de vocabulário/inferência e leitura em LE	166
Tabela 4.30 – Correlações entre variáveis de leitura em LE, conhecimento conceitual, concepções de leitura em LE e leitura em LM	168
Tabela 4.31 – Sujeitos com escores baixos em conhecimento de vocabulário e escores altos em leitura	171
Tabela 4.32 – Sujeitos com escores altos em conhecimento de vocabulário e escores baixos em leitura	173
Tabela 5.1 – Perfil I dos sujeitos	191
Tabela 5.2 – Perfil II dos sujeitos	192
Tabela 5.3 – Pausas/problemas e soluções dos sujeitos	246
Tabela 5.4 – Pausas/problemas comuns aos sujeitos e soluções	247
Tabela B.1 – Número de alunos relatando palavras desconhecidas	285
Tabela F.1 – Todas as variáveis/escores do grupo	311
Tabela F.2 – Análise fatorial: padrão dos fatores	312

LISTA DE FIGURAS

Figura 2.1 – Um efeito limiar hipotético (limiar entre 1000-1800 palavras) (Scott 1990)	39
Figura 4.1 – Todos os testes de vocabulário	142
Figura 4.2 – Posição das palavras quanto à frequência	148
Figura 4.3 – Posição das palavras quanto à familiaridade	150
Figura 4.4 – Categorias das estratégias metacognitivas	153
Figura 4.5 – Estratégias metacognitivas	153
Figura 4.6 – Posição dos sujeitos quanto a concepções de leitura <i>local</i> e <i>global</i>	159
Figura 4.7 – Escore de leitura vs. escores de conhecimento de vocabulário	170
Figura 4.8 – Análise fatorial: Fator 1 × Fator 2	174
Figura 4.9 – Análise fatorial: Fator 2 × Fator 3	175
Figura F.1 – Familiaridade vs. significado	308
Figura F.2 – Posição dos sujeitos quanto à inferência em contexto	309
Figura F.3 – Posição dos sujeitos quanto a conhecimento <i>rico</i> de vocabulário	310

SÍMBOLOS E ABREVIACÕES NAS TRANSCRIÇÕES
(baseados em Cavalcanti 1983)

..	pausa curta (micropausa)
(...)	pausa
(... ...)	pausa longa (macropausa)
[]	adição ou explicação da pesquisadora
/ /	transcrições fonéticas de pronúncias incorretas
((LB))	lê em voz baixa
((LS))	lê silenciosamente
ITÁLICO	palavra ou expressão falada ou lida em inglês
LIA	leitor intermediário avançado
LIO	leitor intermediário
LIE	leitor iniciante
PE	pesquisadora

CAPÍTULO I

INTRODUÇÃO

A importância crescente da leitura em inglês no Brasil, resultado da centralidade da leitura na escola e do inglês como língua de difusão de pesquisa e informação, tem sido responsável por inúmeros trabalhos propondo alternativas, principalmente para um contexto de terceiro grau, onde se concentra grande parte desse ensino em nosso país. Apesar das valiosas contribuições, algumas questões ainda permanecem sem solução, exigindo considerações cuidadosas com relação à adequação dos modelos teóricos – que fundamentam as práticas adotadas – às necessidades e às dificuldades daqueles envolvidos em uma situação de LE. Uma dessas questões diz respeito ao papel do vocabulário na leitura e, conseqüentemente, no ensino da leitura em inglês em um contexto de LE.

Esta tese tem como objetivo oferecer subsídios de relevância teórica e prática que possam vir a preencher essa lacuna, investigando, na perspectiva de um modelo interativo, o papel da competência lexical de alunos universitários brasileiros na compreensão de textos gerais em inglês. Suas contribuições são no sentido de se buscar uma melhor compreensão e caracterização dessa competência e de sua relação com o processo de leitura em LE que permitam:

- indicação da ênfase a ser dada ao componente lexical em relação aos outros componentes e em que estágios da aprendizagem de leitura do aluno numa situação de LE, para que se possa ter um ensino mais adequado às suas necessidades específicas;
- sugestões de critérios para seleção de textos;
- indicações para a elaboração de instrumentos de avaliação de leitura e vocabulário que levem em conta sua complexidade.

Antes de explicitar as questões que orientam esta pesquisa, é necessária uma análise mais detalhada da situação que lhe serve de contexto.

1.1 – O contexto da pesquisa e a justificativa da escolha do tópico

O ensino de leitura no terceiro grau em uma universidade estadual paulista, conduzido através de uma abordagem instrumental, é o contexto que motivou esta pesquisa. A implementação da abordagem instrumental, nessa situação de ensino, teve início em 1980, a partir de uma reflexão que reconhecia a importância da leitura no ensino de LE e, também, de uma pressão prática, ou seja,

a necessidade de utilização de textos em inglês nos cursos de graduação.

Esse trabalho em sala de aula caracteriza-se, principalmente, por um foco em estratégias, em geral transferíveis da LM, visando ao desenvolvimento de uma proficiência que permita ao aluno tornar-se um leitor auto-suficiente, capaz de ler textos de dificuldade compatível com a demanda de seus cursos (1) (vide Kleiman & Terzi 1981; Terzi 1984; Moraes 1990). Nesse contexto, alunos de graduação de áreas diversas freqüentam as aulas de instrumental como disciplina obrigatória ou eletiva por dois semestres com 4 horas/aula semanais (2).

Acredita-se que a leitura de textos específicos da área de estudos de cada aluno seja possível a partir de uma proficiência em leitura desenvolvida com textos mais gerais. O vocabulário técnico, característico do discurso específico, apresenta uma dificuldade que pode ser superada através de uma transferência dos conceitos já aprendidos em LM para a situação de LE. Deficiências, especialmente de conteúdo específico dessas área de estudos, inviabilizaram um trabalho que, no início, era conduzido através de textos específicos. O trabalho se baseava no conceito de negociação do sentido, isto é, o professor contribui com conhecimento lingüístico e discursivo e o aluno com seu conhecimento de mundo e do discurso específico, isto é, uma "*distribuição desigual mas equitativa de conhecimentos*" (Widdowson 1981). Essa dificuldade deslocou o foco para textos mais gerais ou não específicos, extraídos de jornais e revistas, principalmente de vulgarização científica de temas variados (Busnardo & Moraes 1983).

O Inglês Instrumental tem sua origem no ESP — *English for Specific Purposes* ou *Inglês com Objetivos Específicos*, desenvolvido a partir do final dos anos 60 ou início dos anos 70 (Robinson 1980). Segundo Broughton (1980), essa abordagem surgiu na Inglaterra, pela necessidade de se ensinar uma segunda língua (L2 doravante) com finalidades específicas nas áreas de ciência, tecnologia e também no campo profissional e ocupacional. Para Roberts (1982), seus fundamentos estão em duas áreas de estudo, ou seja, na análise de necessidades do público-alvo e da linguagem apropriada a uma determinada área de interesse como, por exemplo, a linguagem da ciência. Nesse sentido, o ESP é parte da abordagem comunicativa, uma vez que dá prioridade às necessidades do público alvo, "*... selecionando conteúdos em função de objetivos bem precisos e operacionalizáveis de modo a responder aquilo que os destinatários esperam ...*" (Souza e Silva 1990:87). No Brasil, a necessidade maior de uso do inglês parece estar associada à leitura, o termo *instrumental* (3) passa, então, a ser usado como sinônimo de leitura, embora seus objetivos possam, em princípio, ser mais amplos.

A abordagem instrumental foi introduzida no Brasil no início dos anos 70, basicamente nas universidades federais e, desde então, tem sofrido modificações. Por volta de 1978, ganhou força através de um projeto nacional, que envolvia 23 universidades federais brasileiras e era coordenado pelo programa de mestrado em Lingüística Aplicada (LA doravante) ao Ensino de Línguas, da Pontifícia Universidade de São Paulo (vide Celani et alii 1988 para uma avaliação do projeto).

Segundo Maciel (1989:258), na fase inicial do projeto, havia uma grande ênfase no componente lingüístico, determinando "... a busca minuciosa do significado em cada palavra e a explicação exaustiva das construções gramaticais". Essa prática de sala de aula, portanto, parecia se fundamentar em um modelo de leitura como decodificação, com materiais instrucionais que privilegiavam estruturas lexicais, morfológicas e sintáticas. Depois desse período, em que o produto da compreensão era focalizado, passou-se a uma ênfase no processo de leitura, com exercícios que exploravam, principalmente, o conhecimento prévio do leitor, podendo refletir a adesão a um modelo descendente (*top-down*). Os trabalhos de Munby (1978) apresentam listas de habilidades de leitura que influenciaram a elaboração de materiais didáticos e de avaliação de leitura.

As abordagens instrumentais, em geral, parecem se fundamentar em um modelo de leitura descendente (vide Capítulo II, Seção 2.2.2 para uma discussão mais abrangente dessa questão). A atenção dada ao componente lingüístico se restringe ao ensino de estratégias de inferência lexical, através de um trabalho que focaliza sufixos e prefixos, reconhecimento de cognatos e de outros elementos mais globais de organização textual, como conectivos e referência, enquanto um ensino mais sistemático da gramática e do vocabulário é considerado ineficiente e conservador. Espera-se, no caso da abordagem instrumental, que a atividade do leitor seja ampliada para suprir também suas deficiências lingüísticas, ou seja, que ela seja decorrente ou um produto derivado da prática de leitura. Isso significa que o leitor, mesmo ainda em fase de aquisição, deverá compensar seu conhecimento incipiente de LE através da ativação de seus esquemas de conhecimento prévio. O desconhecimento da língua não deveria, em princípio, impedir que durante a leitura o leitor abstraísse do texto algum sentido — ou identificasse algumas das idéias centrais do texto lido (Braga e ali 1992).

1.2 - O problema de pesquisa

O uso de uma abordagem instrumental no contexto de ensino acima descrito, com alunos de áreas e níveis diferentes a cada semestre, nem sempre se mostrou eficiente. Enquanto, para alguns, ela permitiu avanços consideráveis em sua capacidade de leitura em LE, com reflexos até mesmo em LM, para outros, ela não acrescentou muito. Além de incapazes de ler de forma independente, esses alunos pareciam apresentar muitas das dificuldades iniciais, ainda no final do segundo semestre, levando, conseqüentemente, aluno e professor a um sentimento de extrema frustração. Quando indagados sobre suas dificuldades, a resposta era sempre a mesma ou seja, "*muitas palavras desconhecidas*".

A discussão dessa questão com colegas levou-me a constatar que a percepção dos alunos, com relação às lacunas de seu léxico e sua conseqüente dificuldade na leitura de textos em inglês,

não era apenas um fato isolado em minha prática, mas uma constante no ensino de outras línguas e em outros contextos de leitura instrumental.

Antes de qualquer tentativa de análise, minha primeira reação foi questionar a confiabilidade dos depoimentos que apontavam o vocabulário como problemático, uma vez que eram fornecidos pelos alunos, podendo ser resultado de uma prática equivocada. Entretanto, um exame mais detalhado desse contexto, através de avaliações, conversas com alunos e, principalmente, de uma observação contínua em sala de aula, revelou que as lacunas no conhecimento lingüístico dos alunos, mais especificamente no conhecimento lexical era, em geral, real, e eles tinham consciência de seus problemas, ao apontarem o vocabulário como sua maior dificuldade. Isso não significava, entretanto, que o vocabulário era sua única dificuldade, uma vez que também apresentavam problemas em outros os níveis, tais como habilidades discursivas e conteúdo específico da área de estudos.

Esse levantamento informal em sala de aula revelou ainda que o conhecimento lingüístico dos alunos que não apresentavam um bom rendimento em leitura era, quando comparado com os que apresentavam um bom rendimento, especialmente fraco, vago e impreciso. O fator distintivo entre os dois grupos poderia ser, portanto, o vocabulário.

Essa constatação levou-me a um questionamento sobre a abordagem instrumental, uma vez que, de acordo com seus proponentes, seria possível superar o vocabulário desconhecido através de uma compensação com outras fontes de conhecimento tais como conhecimento prévio, viabilizado através de um ensino centrado em estratégias.

Paralelamente a esse exame mais detalhado da situação, busquei, na literatura sobre leitura, alguns princípios norteadores com respeito à ênfase a ser dada ao conhecimento lexical vs. outros componentes em minha prática de sala de aula (vide Capítulo II, Seção 2.3 para uma revisão desses estudos). Apesar de extensa, essa literatura apresenta resultados que poderiam ser considerados, em sua maioria, conflitantes e, para mim, inconclusivos.

Com relação à contribuição do componente lingüístico (que inclui o conhecimento da gramática e do léxico), observam-se resultados que justificam posições que podem ser apresentadas em um contínuo que abrange desde a existência de um *nível limiar* (5) ou básico de conhecimento lingüístico (Clarke 1988) até a inexistência de qualquer relação significativa (Alderson & Richards 1977b), passando por posições intermediárias que reconhecem uma correlação forte entre compreensão e competência lingüística (Alderson, Bastien & Madrazo 1977a e Cziko 1980), ou entre compreensão e vocabulário (Cooper 1984 e Laufer 1987). Além dessas controvérsias teóricas, uma revisão dessa literatura mostra também limitações metodológicas relacionadas não apenas à metodologia de coleta dos dados, mas também ao próprio desenho de pesquisa.

A partir da análise da literatura que fundamenta o ensino de leitura, observo que grande parte dos estudos ainda são conduzidos com leitores em uma situação de LM. Sem deixar de

reconhecer a importância dessas contribuições, parece-me ingênuo ignorar suas limitações, uma vez que há diferenças nos dois processos de aquisição que não podem ser ignoradas, principalmente no que se refere ao grau de exposição e, conseqüentemente, de fluência na língua alvo. Enquanto o leitor em LM apresenta sistemas de conhecimento e habilidades completamente desenvolvidos, o leitor em L2, mais especialmente o leitor em LE, é um leitor em desenvolvimento, com lacunas e limitações em ambas as categorias (Eskey 1988). Essa constatação pode ser encontrada em alguns trabalhos recentes, como os desenvolvidos por Eskey (op. cit.) e Carrell (1988 1991) e que começam a questionar, de uma maneira mais sistemática, as perspectivas tradicionais de ensino de leitura, ao atribuir ao conhecimento lingüístico um papel muito mais determinante na compreensão, do que aquele previsto nas abordagens radicalmente descendentes. Nesses trabalhos, os autores reconhecem a superioridade de um modelo interativo para explicar a leitura em LE. A importância dessas observações é bastante grande, dado que esses autores são considerados expoentes de linhas teóricas globalistas ou descendentes.

De certa forma, também corroborando esse posicionamento, há estudos como os de Haynes (1983), dentre outros, que explicam, de uma maneira mais sistemática, a inferência em contexto, ressaltando os inúmeros problemas que o leitor de L2 ou LE tem que enfrentar ao tentar compensar a falta de competência lexical através de inferência e aquisição de palavras através do contexto (vide Capítulo II, Seção 2.4.3.2). No Brasil, as primeiras avaliações dos cursos instrumentais parecem sugerir a necessidade de uma ênfase maior no componente lingüístico (vide Celani e ali 1988; Oliveira 1990; Oliveira 1992; Braga e ali 1992).

Em resumo, aquilo que minha prática – assim como outras também centradas em estratégias – parece mostrar é que o nível de competência lingüística e mais especificamente, lexical, constitui a maior dificuldade na compreensão em leitura em LE. Embora alguns estudos reconheçam as restrições impostas por essa insuficiência, outros vêem a possibilidade de uma compensação desse conhecimento através de outras fontes, incluindo a LM. Mesmo consciente de que esses estudos têm metodologias e arcabouços teóricos diferentes que, em princípio, poderiam explicar os resultados contraditórios ou conflitantes, não posso deixar de salientar que a contribuição desses resultados tem que ser vista com reservas, uma vez que, além de apresentarem limitações teóricas e metodológicas, são conduzidos, em sua maioria, com leitores em LM. Além disso, embora se reconheça que a competência lingüística do falante de LE seja limitada, não se tem estudos que a caracterizem, mostrando em que sentido essa competência é limitada. Embora alguns trabalhos mais recentes reconheçam a necessidade de uma ênfase maior nas dimensões perceptiva e de decodificação no ensino, não se tem uma idéia clara da natureza da relação desse componente com a compreensão que possa fundamentar, de maneira mais sistemática, as abordagens de ensino.

1.3 – As questões de pesquisa

Tomando como ponto de partida o problema acima identificado, com suas limitações e aspectos controvertidos, o presente estudo investiga o papel da competência lexical (conhecimento do léxico assim como seu uso) na compreensão da leitura em inglês como LE. Mais especificamente, espera-se avaliar a competência lexical em inglês, de alunos universitários brasileiros, e relacionar esse conhecimento à sua habilidade de entender textos gerais em inglês. Embora focalizando a relação entre o vocabulário e a compreensão, não posso deixar de salientar outros componentes da leitura em LE, uma vez que, na perspectiva de um modelo interativo, ela é vista como uma inter-relação de fatores. Avaliar-se apenas a contribuição da competência lexical nessa interação seria, portanto, adotar uma visão reducionista e simplista de leitura, apagando sua complexidade. Assim, optou-se por incluir neste estudo a contribuição da proficiência em leitura em LM, do conhecimento prévio que o leitor traz ao texto, assim como de suas diferentes concepções de leitura. Cada um desses os fatores foi operacionalizado através de testes. Foram, portanto, desenvolvidos testes de leitura em LM e LE e um teste de conhecimento conceitual (conhecimento prévio). Para avaliar as concepções que os leitores têm de seu processo de leitura, foi usado um questionário metacognitivo, adaptado de Carrell 1989 (vide Capítulo II e III para uma discussão detalhada dessas questões).

A avaliação do vocabulário tornou necessária a explicitação e operacionalização de um conceito de competência/conhecimento lexical. Adotou-se, nesse sentido, um conceito de competência lexical que é mais abrangente, e que, além do conhecimento lexical, também considera a capacidade envolvida no seu uso. Esse conceito, denominado conceito *rico* de vocabulário (Richards 1976) é composto de um conjunto de componentes lingüísticos, psicolingüísticos e sociolingüísticos e, portanto, mais compatível com o modelo de leitura interativo que serve de arcabouço teórico para este trabalho.

A relação do léxico com a compreensão será buscada tanto do ponto de vista de produto como de processo (vide Capítulo III, Seção 3.1 para uma discussão mais detalhada dessa dicotomia). Um estudo com ênfase quantitativa, horizontal e foco no produto possibilitará uma caracterização da competência lexical dos sujeitos em LE e uma avaliação de seu papel vs. outros componentes na compreensão enquanto produto da leitura. Um estudo com ênfase qualitativa, vertical e foco no processo permitirá uma avaliação da capacidade de uso desse conhecimento ou sua interação com outras fontes de conhecimento na inferência de palavras e, mais especificamente, na construção de um sentido para o texto ou na busca da compreensão. Os dados de produto de 49 sujeitos de diferentes níveis de proficiência em leitura em LE foram obtidos através de testes de vocabulário e de leitura, num primeiro plano, e de testes de conhecimento conceitual e de um questionário metacognitivo, num segundo plano. Os dados de processo foram obtidos a partir de protocolos de pausas de três sujeitos de níveis diferentes de proficiência durante duas tarefas de introspecção, ou

seja, durante a leitura de um texto e identificação de palavras desconhecidas. Os três sujeitos foram selecionados a partir do primeiro estudo ou estudo de produto.

Através dessa investigação busco respostas para as seguintes questões:

1. Como se caracteriza a relação entre a competência lexical e a compreensão enquanto produto e processo da leitura? Ou seja, qual o papel do vocabulário na construção do sentido do texto e, conseqüentemente, na compreensão?
2. Qual a contribuição, para essa relação, sob o ponto de vista de processo e de produto, de fatores como proficiência em LE, conhecimento conceitual, leitura em LM assim como diferentes concepções de leitura?
3. Há evidências de um nível limiar de competência lexical para a compreensão enquanto processo e produto da leitura?

Essas questões mais gerais envolvem outras mais específicas, que serão explicitadas no início dos Capítulos IV e V, que discutem, respectivamente, o produto e processo da leitura.

1.4 – A estrutura da tese

Este trabalho está organizado em seis capítulos. O Capítulo I introduziu o problema de pesquisa e seu contexto, apresentando a justificativa da escolha do tópico e a contribuição esperada, assim como a estrutura da tese. O Capítulo II faz uma revisão crítica da bibliografia, analisando estudos que tratam da leitura e do vocabulário, definindo termos e conceitos necessários, salientando limitações teóricas, que servirão de ponto de partida para as questões investigadas. O Capítulo III introduz o desenho da pesquisa e sua justificativa, a partir das limitações metodológicas levantadas no capítulo anterior. A análise e discussão dos dados de ênfase quantitativa e a análise e discussão dos dados de ênfase qualitativa são apresentadas nos Capítulos IV e V, respectivamente. O Capítulo VI traz uma discussão comparativa dos resultados, que abrange os resultados dos dois estudos, assim como as conclusões, limitações e implicações teóricas, metodológicas e práticas desta tese. O Apêndice, subdividido em sete partes, inclui os testes para avaliar o vocabulário, a leitura e os outros componentes, o material usado na tarefa de introspecção e seu resultado, os seja, os protocolos dos três leitores, além de quadros, figuras e tabelas.

Notas

(1) Nos anos que se sucedem à coleta dos dados desta tese essa abordagem tem sido revista, passando a considerar também questões lingüísticas.

(2) Embora as disciplinas de Inglês Instrumental tenham em vista os alunos de graduação, num primeiro momento, também têm sido bastante procuradas por alunos de pós-graduação, que se utilizam das vagas remanescentes.

(3) O termo *instrumental* tem sido criticado. Para Celani (1981), "... é insatisfatório, por ser restrito e imediatista e, ainda por sugerir uma visão da língua como mero instrumento para se alcançar um fim". A autora propõe o termo *ESP*. Embora também discorde do termo *instrumental* para a abordagem usada nos cursos de leitura no Brasil, não concordo com a denominação *ESP*, uma vez que ela pressupõe a leitura de textos específicos da área de estudo do aluno. Como mostrei acima, nem sempre isso acontece. Na UNICAMP, por exemplo, o Inglês Instrumental é conduzido a partir de textos gerais.

CAPÍTULO II

A COMPETÊNCIA LINGÜÍSTICA NAS PESQUISAS EM LEITURA EM LE

No capítulo anterior, delimitei o problema desta investigação, apresentando sua justificativa, assim como, o contexto que o motivou. Neste capítulo, estruturado em quatro seções principais, busco construir, através de uma revisão crítica da literatura que focaliza o papel da competência lingüística e lexical na leitura, uma base teórica para esta investigação. As Seções 2.1 e 2.2 têm por objetivo, respectivamente, localizar a leitura em LE dentro da área mais geral da leitura em LM, assim como destacar, dentre os modelos/visões que fundamentam as pesquisas em leitura em LM e LE, aquele que constitui o arcabouço teórico desta investigação. A Seção 2.3 focaliza o componente lingüístico (a gramática e o vocabulário) vs. o componente cognitivo (habilidades/estratégias de uso do conhecimento prévio) na leitura em LE, apresentando algumas das diferentes hipóteses, identificadas na literatura em leitura, sobre o papel desses componentes na compreensão. O vocabulário em sua relação com a leitura é focalizado na Seção 2.4. Seu objetivo é apresentar evidências empíricas sobre seu papel na compreensão, destacando, dentre modelos de competência/conhecimento lexical, o de um conceito *rico* de vocabulário, que complementa a fundamentação deste estudo.

A segmentação em componente lingüístico, cognitivo e lexical, apesar de confusa, artificial e difícil de se justificar, mesmo sob o ponto de vista da pesquisa, está sendo adotada na organização da revisão apresentada neste capítulo, uma vez que é a segmentação seguida pelas pesquisas em leitura em geral. Uma discussão mais detalhada dessa questão, assim como uma proposta alternativa, serão apresentadas na Seção 2.4.3.3 deste capítulo.

2.1 – As pesquisas sobre leitura em LE

A área de pesquisa em leitura é bastante vasta. Já em 1984, Alderson salientava a impossibilidade de uma revisão, dado o volume de artigos desenvolvidos principalmente durante os últimos trinta anos.

Só no Brasil, principalmente nos últimos dez anos, a leitura tem sido foco de pesquisa e de reflexão em diversas áreas, dentre as quais destaca-se Educação (Silva 1988; Zilberman 1982; Bordini 1988); Teoria Literária (Lajolo 1994; Lajolo & Zilberman 1991); Análise do Discurso – linha francesa (Orlandi 1981, 1983); Coracini 1991; Tradução (Arrojo & Rajagopalan 1991); e

Linguística Aplicada (LA doravante) (Cavalcanti 1983, 1989; Moita Lopes 1986; Kleiman 1989a, 1989b). Nessas áreas, com exceção da LA, os estudos sobre leitura têm-se voltado mais para a reflexão teórica do que para a pesquisa empírica. Em LA, as pesquisas, publicadas em forma de teses de doutorado, artigos e livros, têm focalizado tanto a LM como a LE. Moita Lopes (1991, 1994) mostra, entretanto, que mais 50% das teses de doutorado, produzidas em LA por brasileiros, estão na área de leitura em inglês como LE.

O interesse pela pesquisa em leitura em LE no Brasil não pode ser considerado isoladamente, uma vez que parece fazer parte de uma tendência internacional que se iniciou por volta de 1980 trazendo, como consequência, um avanço considerável para a área. Apesar de vista como a mais importante das quatro habilidades (compreender, falar, ler e escrever) (1), principalmente em inglês como segunda língua (L2 doravante) ou ainda mais como LE, a leitura era, até essa época, negligenciada dentro da área de aquisição de L2. Essa situação, especula Carrell (1988), era resultado de uma tradição estruturalista, que privilegiava o papel da linguagem oral, em que a leitura em LE era vista como uma habilidade complementar ao ensino de habilidades orais.

Apesar desses últimos avanços, não se pode deixar de reconhecer que modelos abstratos, desenvolvidos a partir de resultados de pesquisas com leitores em LM, ainda têm sido, em grande parte, a base para explicar a leitura em L2 e LE (2), assim como para orientar seu ensino.

Embora as evidências até o momento tenham identificado pontos convergentes na natureza dos processos de leitura em LM e LE, não se pode deixar de lembrar que, quando focalizados sob o ponto de vista da aquisição/aprendizagem (3), apresentam diferenças que não podem ser ignoradas, principalmente no que se referem à fluência e grau de exposição à língua alvo. Para Eskey & Grabe (1988:227), dentre outros,

(...) tais modelos, em geral, são modelos do leitor 'ideal', completamente fluente, com sistemas de conhecimento e habilidades completamente desenvolvidos, enquanto o leitor de L2 é, quase por definição, um leitor em desenvolvimento, com lacunas e limitações em ambas as categorias (1c) (4).

O ensino da leitura em uma situação de LM já conta com uma base lingüística, que muitas vezes é inexistente ou insuficiente, quando o contexto é de L2 ou LE. Especialmente em situações de LE, as aulas de leitura (consistindo, na maior parte das vezes, de algumas poucas horas semanais) acabam sendo, para os aprendizes, as únicas oportunidades de exposição à língua alvo. A fluência na língua é, dessa forma, prejudicada.

Não posso deixar de salientar que esse posicionamento a respeito do processo de aquisição da leitura em LE não só é influenciado por resultados de pesquisas e observações em contextos específicos de ensino de leitura em LE, mas também, por um modelo/visão de leitura que considera,

dentre outras coisas, a importância de uma proficiência linguística na leitura. Sem dúvida, outras visões têm levado a posicionamentos diferentes com relação à natureza desse processo de aquisição. Por essa razão, apresento, na próxima seção, uma revisão dos modelos/visões de leitura que tiveram um maior impacto nos estudos sobre leitura em LM e LE, destacando o modelo de leitura que serve de arcabouço teórico para esta pesquisa. Essa revisão é fundamental, não apenas para justificar resultados, por vezes contraditórios, que poderão ser observados nas próximas seções, mas, principalmente, a escolha do modelo que fundamenta este trabalho.

2.2 – Os modelos/visões de leitura

Uma revisão dos já cem anos de pesquisa sobre leitura mostra que somente depois de 1955-1960 começaram a aparecer os primeiros modelos de leitura. Samuels & Kamil (1988) oferecem uma explicação para essa situação: até essa época, não havia uma tradição forte para a tentativa de conceitualização do conhecimento e da teoria sobre o processo de leitura. Ainda para esses autores, os modelos, em geral, baseiam-se em conhecimento limitado, influenciado pelas filosofias científicas e estudos proeminentes durante o contexto histórico em que foram desenvolvidos. Assim, até 1960, os modelos tinham uma forte influência behaviorista, buscando, principalmente, descrever o que era considerado como estímulo, ou seja, palavras impressas e respostas ao reconhecimento dessas palavras. Os modelos desenvolvidos depois dessa data, no entanto, são predominantemente cognitivistas, tentando mostrar a função, na leitura, de processos como memória e atenção. Além disso, os modelos dos anos 70 foram, em sua maioria, modelos lineares de processamento de informação, enquanto, os mais recentes, tendem a ser interativos.

Os modelos teóricos de leitura costumam ser classificados de acordo com dois critérios fundamentais: a direção do fluxo principal da informação (Rumelhart 1977; Stanovich 1980) e o grau de determinação do texto (Kleiman 1989a). Essa categorização é bastante ampla e genérica, resultando em três modelos básicos: o modelo de decodificação, o modelo psicolinguístico e o modelo interativo. As variações dentro de cada tipo privilegiam aspectos diversos da leitura. Diferentes visões ou concepções de leitura/compreensão são pressupostas nos diferentes modelos, com conseqüentes implicações, não apenas para o seu ensino, mas também para sua avaliação.

2.2.1 – O modelo ascendente ou de decodificação: uma visão de leitura como extração de significado

O modelo ascendente ou de decodificação é aquele em que o fluxo da informação é ascendente, isto é, se inicia com a percepção dos dados na página impressa, procedendo em uma

seqüência serial fixa, quase mecânica, sempre das unidades menores (reconhecimento de letras e palavras) para as maiores (frases, orações, ligações intersentenciais), até chegar ao significado.

Segundo Stanovich (op. cit.), uma grande limitação dos modelos ascendentes é a falta de realimentação, pois não há um mecanismo que possa controlar a influência dos estágios, que ocorrem posteriormente, sobre os que ocorrem anteriormente no sistema. Por causa dessa dificuldade, é difícil explicar os efeitos de contexto, e o papel do conhecimento prévio do tópico do texto, como variáveis no reconhecimento da palavra e na compreensão. Esses modelos, portanto, levam em conta os processos de nível baixo ou perceptivos, mas não conseguem explicar os processos de nível mais alto, ou cognitivos.

Além das denominações decodificação e ascendente (*bottom-up*), esse modelo tem recebido outras, tais como baseado nos dados (*data-driven*); de fora para dentro (*outside-in*); baseado no texto (*text-based*); processamento da informação (*information processing*); e, finalmente, posição de percepção direta (*direct perception position*).

O modelo de processamento serial de Gough (1976), assim como o modelo de processamento automático de LaBerge e Samuels (1976), são geralmente citados como exemplos de modelos de decodificação. Enquanto o primeiro poderia ser considerado um extremo da teoria, e, portanto, seu melhor representante, o segundo parece explicar o processo de leitura de uma maneira mais flexível e realista. Nesse modelo, os autores argumentam a favor de um processamento automático de palavras, que permitiria a liberação de espaço cognitivo para uma reflexão a respeito do significado do que está sendo lido (Grabe 1988).

A visão de leitura pressuposta nesse modelo é uma visão de extração de significado, em que o texto é visto como o foco da decodificação, ou objeto completamente determinado. Essa visão concebe o ato ilocucionário independente, origem de todas as pistas para se chegar ao significado, apostando na inviolabilidade da palavra e no sentido literal. Este, segundo essa tese, originar-se-ia dos significados de seus componentes, sendo suficiente, para o ouvinte, independentemente do contexto, chegar à intenção do falante (no caso, o autor do texto). Dessa forma, através da decodificação, entendida aqui como um processo ascendente, é possível recuperar a *mensagem* ou *intenção* do autor. O papel do leitor é de simples analisador de insumo gráfico, uma vez que suas contribuições para a compreensão são totalmente ignoradas. Baltra Montaner (1982), dentre outros, aponta uma correspondência entre uma visão ascendente e um conceito behaviorista-estruturalista de compreensão de linguagem.

Dascal (1986), discutindo a concepção de significado literal salienta que:

Poder-se-ia dizer que uma grande parte da história da filosofia da linguagem, neste século, consiste no descobrimento progressivo das dificuldades inerentes a esta concepção. Entre elas, as limitações do princípio da concepção de significados, a importância do contexto na

interpretação da frase, o dilema de traçar com precisão a linha divisória entre o dicionário e a enciclopédia, ou entre o analítico e o sintético, e a insuficiência crônica do 'significado literal' como meio de determinar as condições de verdade e as propriedades das frases (2c).

Sob o ponto de vista do ensino/avaliação de leitura em LM e LE, essa visão conduz a uma abordagem que enfatiza unicamente o texto ou sua decodificação, através do desenvolvimento de um conhecimento lingüístico sólido e efetivo, que permite o reconhecimento automático de palavras. Embora não se faça referência à função do contexto nessa abordagem, pode-se depreender que ele é, nesse caso, visto de maneira estreita, ou seja, um contexto que considera apenas os elementos do texto ou seu co-texto.

O texto, focalizado como um objeto determinado, faz da leitura um processo passivo de decodificação, não deixando espaço para a subjetividade do leitor. A leitura pós-alfabetização em LM no Brasil, tanto no primeiro como no segundo graus, tem-se pautado nessa visão (Cavalcanti 1992), assim como o ensino de leitura em LE. Não é difícil perceber a razão de sua popularidade, nem as suas implicações para o ensino. Conceber o texto como portador de um significado único, imutável, e garantir a possibilidade de recuperação desse significado proporciona segurança e poder ao professor/avaliador, facilitando seu trabalho. Essa visão permite ao professor impor sua interpretação como a única possível, podendo levá-lo "*... a ignorar as contribuições do aluno/leitor para a compreensão textual, a mistificar a palavra escrita e a desestimular questionamentos sobre o texto*" (Tavares 1993:29), com prejuízos sérios para esse aprendiz.

2.2.2 – O modelo descendente ou psicolingüístico: uma visão de leitura como (re)criação do significado

O modelo descendente ou psicolingüístico de leitura muda a ênfase do texto para o leitor. Por sua vez, a direção do fluxo principal da informação passa a ser descendente, iniciando-se no leitor e procedendo em direção ao texto, que é visto como objeto indeterminado e incompleto, cabendo, ao leitor, impor-lhe uma estrutura, (re)criar um significado (Kleiman 1989a). Nesse processo, entretanto, o leitor não parece se utilizar de todas as pistas textuais, mas apenas aquelas que lhe permitam confirmar suas hipóteses ou previsões, baseadas em seu conhecimento prévio. Quanto mais o leitor for capaz de fazer previsões, menor será sua necessidade de confirmação através do texto (Goodman 1973). Para esse autor, portanto, a leitura é vista como um jogo psicolingüístico de adivinhação (*a psycholinguistic guessing game*), e, a compreensão, fortemente determinada pelo universo do leitor. Este não é apenas um participante ativo de seu processo, fazendo previsões e processando informações, mas tudo em sua experiência ou conhecimento prévio exerce um papel significativo no processo: seu conhecimento lingüístico (*linguistic schemata*), do conteúdo ou

assunto (*content schemata*) e da estrutura retórica do texto (*formal schemata*). Como salienta Kato (1984:134),

(...) o conhecimento prévio do leitor passa a ter um papel igual ou até mesmo mais importante do que os dados do texto. A leitura de uma frase passa a ser vista não mais como uma operação precisa de análise e síntese, mas como um jogo de adivinhações.

Para Moita Lopes (1986:70), "... o modelo baseado no leitor quase que totalmente ignora o texto escrito ... o leitor interpreta o texto apenas pragmaticamente e ignora sua interpretação semântica" (3c) (5). A leitura, portanto, passa a ser um ato eminentemente cognitivo, contrastando com a visão anterior que a concebia como um ato perceptivo.

Além das denominações psicolingüístico e descendente (*top-down*), esse modelo tem recebido outras, tais como de dentro para fora (*inside-out*); guiado pelo conceito (*conceptually-driven*); baseado no leitor (*reader-based*); guiado pelo schema (*schema-driven*); baseado na predição (*prediction-based*); análise pela síntese (*analysis by synthesis*); e, finalmente, posição de teste de hipótese (*hypothesis-testing position*).

O modelo de testagem de hipóteses de Goodman (1973a, b; 1971, 1976), assim como o modelo de sistemas de comunicações de Ruddel (1976), podem ser considerados representantes dessa visão de leitura.

A principal contribuição da tendência descendente foi a de trazer, para o processo de leitura, uma dimensão nova, a do conhecimento prévio do leitor, mostrando que a leitura é muito mais do que apenas um processo passivo de extração da mensagem (6) do autor, através de uma decodificação linear ascendente de seus elementos lingüísticos, ou, em outras palavras, que todo significado está no texto. Sua principal limitação foi, entretanto, a de enfatizar excessivamente a contribuição dos processos cognitivos envolvidos na leitura, e, conseqüentemente, minimizar a dimensão de percepção e de decodificação desses processos. Seu impacto foi tão grande que, ao invés de complementar a visão ascendente ou de decodificação, passou a substituí-la. A visão de leitura subjacente, entretanto, como mostro a seguir, parece ter permanecido a mesma.

Uma outra crítica que se pode fazer ao modelo psicolingüístico é que, da mesma forma que no modelo anterior de decodificação, aspectos sociolingüísticos ou referentes às regras da língua em uso, e psicossociais, ou referentes às atitudes do leitor na interação, não são considerados.

Reconhecer o texto como um objeto indeterminado (na realidade, o termo mais adequado seria parcialmente indeterminado, pois é incompleto), entretanto, não implica que o leitor possa construir um significado ou impor-se totalmente a ele. Uma noção de construção de significado, como mostrarei na próxima seção, não pode ser confundida com um processo de preenchimento de lacunas que caracteriza esse modelo. A noção de (re)criação, que embora, à primeira vista, possa

significar uma liberdade total de interpretação, parece também implicar, como no modelo anterior, uma visão de extração de significado, uma vez que, através de seu conhecimento (prévio), o leitor tentará, na realidade, fazer sentido do que o autor tem a dizer. A expressão *jogo de adivinhações* não parece deixar dúvidas com relação à visão subjacente, pois adivinhação pressupõe uma única resposta correta (7). Ao invés de usar, para isso, um procedimento que se inicia nos elementos do texto (ou, o que Moita Lopes (op. cit.) denomina de *interpretação semântica*), como no modelo anterior, através de uma decodificação ascendente, o leitor vai usar seu conhecimento prévio, ou de mundo e do assunto do texto, através de um processamento descendente. Nessa interpretação, que Moita Lopes denomina de *pragmática*, entretanto, também serão levadas em conta as intenções do autor ou a sua *mensagem*. A descrição que Goodman (1988:12) faz da leitura parece confirmar mais uma vez essa visão, mostrando que "... o leitor reconstrói ... a mensagem que foi codificada pelo escritor através de uma apresentação gráfica" (4c).

Como resultado da idealização $LM = L2/LE$ acima mencionada, o modelo descendente causou uma verdadeira revolução na leitura em LM, e também em L2/LE. Essa influência se iniciou por volta de 1970 (Eskey 1973; Saville-Troike 1973) e, em 1979, Coady propõe um modelo para L2 baseado nessa tendência. Seu impacto pode ser observado em uma época em que a insatisfação com o método audiolingual no ensino em L2 era crescente e os professores estavam se tornando mais conscientes de que uma proficiência aural-oral não levava, automaticamente, a uma competência em leitura. Como consequência, a leitura como uma área independente começa a emergir (8).

Um exemplo claro da tendência descendente é a vasta lista de artigos apresentada em Eskey (1988:95), ou seja, "*praticamente tudo publicado em TESOL Quarterly nos anos de 1981 a 1984*". No Brasil, essa tendência parece se repetir ao se examinar *The ESPspecialist*, uma publicação da Pontifícia Universidade Católica de São Paulo. A grande maioria dos cursos de leitura instrumentais no Brasil parece se fundamentar nos pressupostos desse modelo.

As abordagens de ensino de leitura em L2/LM que se fundamentam em um modelo descendente, como por exemplo, as instrumentais, têm enfatizado o componente cognitivo ou o uso do conhecimento prévio através de um trabalho que focaliza estratégias, como mostro na Seção 2.3. Dentre elas, destacam-se, especialmente, as estratégias de inferência lexical (vide Seção 2.4.3.2 deste capítulo). Essas abordagens, contrariamente àquelas baseadas no modelo de leitura ascendente, salientam a importância do contexto tanto para a inferência de palavras como para a extração da intenção do autor, em detrimento do desenvolvimento de um conhecimento lingüístico mais efetivo e da decodificação automática. A noção de contexto, entretanto, não seria a mesma do modelo anterior, uma vez que, neste caso, as pistas textuais ou o co-texto são apenas usados quando necessários. Pode-se depreender, portanto, uma noção de contexto praticamente limitada ao conhecimento prévio que o leitor traz para o texto.

Embora o conhecimento prévio seja enfatizado no modelo descendente, observa-se, nas

abordagens de ensino de leitura que pressupõem essa visão, que os tratamentos dados ao componente lingüístico, apesar de parte desse conhecimento, parecem divergir. Em algumas delas, nota-se uma tendência a ignorá-lo completamente. Em outras, a abordá-lo apenas superficialmente. Apesar das variações, a tendência geral é de considerá-lo como algo secundário, que pode ser compensado via ativação de esquemas de conhecimento de mundo, do assunto do texto, independente do nível de proficiência lingüística apresentada pelo aluno/leitor. Essa interpretação, que parece "confundir" os aspectos perceptivos do texto ou de decodificação, com o conhecimento lingüístico necessário para essa decodificação, entretanto, poderia ser justificada, não apenas pela falta de uma explicitação maior do papel desse componente no modelo, mas, principalmente, pelas condições ou o momento histórico em que essa abordagem foi introduzida no ensino de leitura. Não se pode deixar de salientar que o modelo descendente (foco no leitor) se contrapõe ao modelo ascendente (foco no texto), que enfatiza o lingüístico e a decodificação. A tendência mais natural, nessa situação, sem dúvida alguma, seria buscar um outro extremo, seguindo o movimento pendular caracteriza as tendências de pesquisa e o desenvolvimento na ciência em geral (9).

Uma análise mais cuidadosa, contudo, poderia mostrar que, a falta de ênfase na dimensão perceptiva ou de decodificação, não implica, necessariamente, na exclusão do conhecimento lingüístico. Uma interpretação mais consistente com o modelo é, portanto, de um conhecimento lingüístico pressuposto. Convém salientar que o modelo descendente foi elaborado para explicar a leitura em LM e, nesse contexto, o componente lingüístico geralmente está presente a priori, e não costuma ser problemático. Não há, portanto, razão para enfatizá-lo nesse ensino. O fato de o leitor não ter necessidade de se utilizar de todas as pistas textuais não significa que possa ignorar esse conhecimento, uma vez que sem ele, as pistas não seriam pistas, mas apenas símbolos sem sentido em uma página. Embora um modelo descendente possa ser utilizado para explicar o processo de leitura em LE, parece problemático tomá-lo enquanto arcabouço teórico para fundamentar um processo de aquisição de leitura em LE, porque, nesse caso, as necessidades dos aprendizes são diferentes. Um foco em estratégias, ou no componente cognitivo via ativação de esquemas de conhecimento prévio, não parece suficiente para compensar as insuficiências lingüísticas e levar à compreensão. Em outras palavras, um ensino de leitura descendente não parece instrumentalizar o leitor para uma leitura independente. Pelo menos é o que parecem mostrar os resultados de algumas das pesquisas relatadas no Capítulo I, assim como as observações baseadas no contexto de leitura deste trabalho.

Uma hipótese que poderia ser aventada para explicar o sucesso de abordagens descendentes, principalmente em contextos de LE, fazendo com que a crítica acima, ou o fato de serem baseadas em um modelo de LM, passasse despercebida ou fosse colocada em um segundo plano, estaria relacionada a uma questão ideológica (10). A aprendizagem de inglês em situações de LE, mais precisamente, em países em desenvolvimento ou do Terceiro Mundo, parece associada à questões

mais complexas de "submissão" e "colonialismo". A língua não pode ser considerada um instrumento "neutro" de transmissão de informação e de acesso a novas tecnologias; ela é, também, um instrumento de transmissão dos valores de uma cultura, que nesse caso, é vista como dominante. O ensino da LE, nesse contexto, em geral orientado por uma motivação integradora, tem exacerbado esses sentimentos e incentivado rejeições à essa língua e a esses valores. Não resta dúvida de que, uma abordagem que se propõe a ensinar a leitura, sem que o aprendiz tenha que aprender a língua alvo, apenas fazendo uso dos recursos cognitivos de que ele já dispõe em LM, tem um apelo muito grande. A motivação, nesse caso, seria instrumental, contrastando com a motivação mais integradora das abordagens anteriores. O fato de essa abordagem se concentrar apenas na leitura global, ou fazer da extração da idéia principal do texto o único objetivo da leitura, também poderia estar contribuindo para sua popularidade. O acesso à bibliografia e às publicações em língua inglesa seriam possível, nesse caso, sem que o leitor tivesse que se submeter aos efeitos de uma ideologia dominadora. Por outro lado, sob o ponto de vista de seus proponentes ou difusores, a proposta também pode ser interessante, pois seria mais chance de aplicar e difundir seus resultados de pesquisa em LM.

Essa possibilidade, entretanto, não deixa de ser ilusória e equivocada, porque apenas dá ao leitor a impressão de que ele está em controle da situação, uma vez que a visão de leitura subjacente ainda é de recuperação de significado, em que o autor é visto como um mero decodificador da mensagem do autor. Uma alternativa, que me parece mais interessante para países em desenvolvimento que se sentem ameaçados por ideologias dominadoras, seria exatamente a oposta, ou seja, adotar abordagens que criem condições para a formação de leitores críticos, capazes de discernir possíveis tendências de domínio e tirar proveito do que lhe é interessante. E para que uma leitura crítica seja possível em LE é necessário, a meu ver, uma competência lingüística que possibilite a negociação de significados. Na minha opinião, essa posição é mais consistente também com a maneira como o conhecimento é construído, isto é, não apagando ou ignorando o que vem antes mas construindo a partir dele.

Uma visão um tanto estreita, e até mesmo simplista da leitura se confirma ao observar-se, nessa abordagem, a tendência de estabelecer a idéia global do texto como o único nível de compreensão, deixando de considerar o propósito ou objetivo da leitura. Essa crítica não tem por objetivo subestimar a contribuição de uma leitura global, mas apenas colocá-la em uma dimensão mais realista e produtiva, em que é concebida como um começo promissor de um processo que leva às hipóteses internas sobre o texto (Galves & Busnardo 1983; Galves 1983). O fato de as abordagens descendentes se concentrarem na idéia central não parece, entretanto, uma escolha deliberada, mas o único objetivo de leitura possível, dadas as restrições apresentadas pelos leitores, principalmente no aspecto lingüístico.

2.2.3 – O modelo interativo: uma visão de leitura como construção do significado

Embora também proposto a partir do trabalho teórico e de pesquisa em LM, o modelo interativo, entretanto, é mais abrangente e, portanto, mais condizente com as necessidades do leitor em seu processo de aquisição de leitura em LE. Por esse motivo, foi escolhido como arcabouço teórico deste estudo. Nesse modelo, a leitura é vista como um processo cognitivo e, ao mesmo tempo, perceptivo, envolvendo uma combinação de processos ascendentes e descendentes.

Antes de buscar uma caracterização desses modelos, entretanto, não posso deixar de salientar que o termo *interativo* tem sido usado com sentidos diferentes dentro da literatura, referindo-se, principalmente, a dois conjuntos independentes mas relacionados de pesquisa, como salientam Eskey (1988) e Grabe (1988).

Grabe (op. cit.), em um artigo que relaciona esse modelo à leitura em LE faz, uma distinção entre os conceitos *processo interativo* e *modelo interativo* (11). De acordo com esse autor, a visão de leitura como um processo interativo é discutida por Widdowson (1979), como um processo de combinação da informação textual (ascendente) com a informação que o leitor traz para o texto (descendente), ou da interação entre a mente do leitor e os elementos do texto. Durante a leitura ocorre a ativação dos vários tipos de conhecimento na mente do leitor, que, por sua vez, como resultado da informação nova fornecida pelo texto, são refinados e ampliados. A leitura caracterizar-se-ia, assim, como um diálogo de negociação do sentido entre o leitor e o texto.

Embora essa visão tenha trazido contribuições muito importantes, introduzindo aspectos pragmáticos ou interacionais não considerados nos outros modelos, concordo com esses autores quando afirmam que a dominância do conceito de leitura, como um processo interativo em LE, eclipsou algumas questões importantes. O que esse conceito assume é, mais uma vez uma semelhança entre os processos de leitura em LM e LE e a possibilidade de transferência automática de habilidades da LM para a L2, deixando de lado, principalmente, a diferença entre a proficiência lingüística em LE e LM, ou a limitação no vocabulário/sintaxe dos leitores em LE. Uma outra implicação dessa visão é a idéia de que os textos são informação, quando, na realidade, são apenas marcas em uma página, que os leitores devem converter em linguagem/informação.

Complementando essa visão, os autores chamam a atenção para uma outra versão do termo interativo, relacionada ao seu uso em modelo interativo (Rumelhart 1977; Stanovich 1980). Para Grabe (op. cit.:58),

Essa visão de leitura não deveria ser considerada uma versão alternativa da leitura como um processo interativo. A questão não é a relação do leitor com o texto mas as relações de processamento entre as várias habilidades na leitura (5c).

Essa mesma visão é compartilhada por Eskey (op. cit.:96) quando afirma que

(...) interativo refere-se à interação entre informação obtida através de uma decodificação ascendente e informação obtida através de uma análise descendente, ambas dependentes de certos tipos de conhecimento prévio e certos tipos de habilidades de processamento de informação (6c).

Um modelo interativo caracteriza-se, portanto, por apresentar uma bidirecionalidade de fluxo da informação, isto é, do texto para o leitor (ascendente) e também do leitor para o texto (descendente). Nesse caso, a construção do sentido durante o processo de leitura é feita através de uma interação leitor/texto, ou leitor/pistas do texto. As informações não se encontram no texto, que apenas oferece pistas indexicais para o leitor (Cicourel 1974). Para que essas pistas se transformem em informações, é necessário que o leitor use seus vários níveis de conhecimento, inclusive o lingüístico. Dentro dessa visão, o termo *estratégia*, que tinha sua aplicação restrita a procedimentos descendentes, passa a descrever também procedimentos de decodificação e interpretação no âmbito sintático e semântico (Kato 1983). A leitura bem sucedida é muito mais do que decodificação pura e simples, embora a decodificação (denominada habilidade de nível baixo ou *lower-level skill*) também seja um processo cognitivo, envolvendo habilidades de processamento ascendentes e descendentes, e a compreensão bem sucedida não pode ser alcançada sem eles.

As duas versões do termo interativo, que acabo de apresentar, não são excludentes, uma vez que os modelos interativos incorporam as implicações de leitura como um processo interativo. Portanto, neste trabalho, o termo refere-se aos dois processos, ou seja, de interação entre o leitor/autor através do texto, e, ao mesmo tempo, entre o texto e os vários níveis de conhecimento do leitor.

A maior contribuição do modelo interativo foi introduzir uma visão de leitura e de compreensão diferente daquela pressuposta nos modelos anteriores, isto é, uma visão de construção de significado. Assumir uma posição unilateral de que o texto é portador de um significado único, impondo-se totalmente à capacidade interpretativa do leitor ou, por outro lado, ignorar o texto ou concebê-lo como objeto totalmente indeterminado não parece consistente com a complexidade e dinamismo característicos do processo de leitura.

Entretanto, a consideração do leitor e de seu conjunto de conhecimentos/competências parecem introduzir dúvidas com relação à possibilidade de se ensinar e, principalmente, de se avaliar essa leitura. A sólida base lingüística, que dava ao professor e avaliador segurança e controle sobre esse processo, começa, de ser forma, a ser abalada ou questionada.

Numa perspectiva, portanto, de leitura como processo dinâmico e interativo do leitor com o autor através do texto, seria possível recuperar a leitura? O que seria, nessa visão, uma leitura

adequada? (12) Quais seriam as implicações dessa posição para a avaliação? Ao responder essas perguntas, adoto o termo leitura "adequada" de acordo com Paes de Barros e Rojo (1984:58):

(...) uma leitura 'adequada' (...) é uma leitura 'negociada': nem só e todo significado pretendido pelo autor será recuperado, nem qualquer significado imposto pelo leitor, mas a intersecção entre o conjunto cognitivo do leitor (...) [construído socialmente, abrangendo conhecimento de mundo, de discurso, de estratégias para o processamento desse discurso, conhecimento de contexto, valores, crenças, atitudes] e o conjunto conceitual e estrutural nuclear proposto pelo autor no texto.

Nesse contexto de interação, a legibilidade de um texto, outrora definida a partir de elementos textuais, passa, então, a ser considerada, não em função apenas de suas marcas formais explícitas, mas a partir dessa relação de interação (13). Sob essa perspectiva, um mesmo texto pode permitir não apenas várias leituras para diferentes leitores, como também para um mesmo leitor, em momentos diferentes, uma vez que ela é vista como uma atividade "... eminentemente discursiva (e nessa medida também lingüística) expressão de uma subjetividade complexa, (individual, social, ideológica)..." (Coracini et alii 1988) (14).

Entretanto, admitir-se uma pluralidade de leituras não significa que qualquer leitura seja possível. Concorro com Braga e Busnardo (1984), seguindo Widdowson (1979) quando salientam que não se pode ignorar que todo o texto (alguns mais, outros menos) impõem restrições à liberdade interpretativa do sujeito.

Um texto pode conter apenas sugestões para a obtenção do significado, mas essas sugestões existem e restringem as interpretações possíveis. A liberdade do sujeito é limitada pelas restrições que existem a nível da língua, do discurso e do conhecimento de mundo (Braga e Busnardo op. cit.:28)

O reconhecimento de um núcleo "mais ou menos estável ou convergente" (Cavalcanti 1992) torna possível seu ensino/avaliação. Entretanto, como salienta Cavalcanti (op. cit.:224),

Trata-se de uma visão de convergência que difere daquela visão tradicional, que era estabelecida artificialmente por um único sujeito (professor ou pesquisador, por exemplo), que partia do pressuposto de que o texto só podia ser lido de uma maneira. Agora, a convergência precisa ser redefinida como fruto de um processo de intersubjetividade de diferentes leitores e como algo distinto da pluralidade. Essa redefinição deve levar em conta que o texto é um objeto indeterminado, mas com significado em potencial.

A visão de leitura como construção ou negociação, pressuposta no modelo interativo, não pode, entretanto, ser confundida com uma outra, que ora tem se sobressaído na literatura. Apesar de também conceber a leitura como um processo construtivo e o texto como um objeto totalmente indeterminado, difere da que acabo de caracterizar, uma vez que não considera o papel do componente lingüístico, ou seja, não incorpora qualquer discussão sobre a direção do fluxo da informação, ou mais especificamente, de processamentos ascendentes ou descendentes. Essa visão, construída à luz da teoria desconstrucionista de Derrida,

(...) aceita que qualquer relação entre sujeito e objeto seja inevitavelmente mediada por uma interpretação, produto, por sua vez, das circunstâncias históricas, geográficas, psicológicas, sociais, institucionais — que constituem esse sujeito. Os significados seriam, assim, inevitavelmente atribuídos, impostos aos objetos e não descobertos neles ou a partir deles. Conseqüentemente, toda leitura envolveria uma produção — e não uma descoberta — de significados, criados a partir das circunstâncias que constituem o leitor e a situação da leitura. Em poucas palavras, o significado não se esconde nem se embute no texto à espera de que um leitor decifre, compreenda ou resgate; o significado é produzido pelo leitor a partir de suas circunstâncias e das convenções que organizam e delimitam suas instituições, inclusive a linguagem (Arrojo & Rajagopalan 1991:46).

Embora essa posição e a outra acima caracterizada tenham um ponto em comum, uma vez que ambas consideram a leitura mais do que uma "descoberta", ou que significado não se "embute ou se esconde no texto", sendo construído ou gerado pelas contingências da própria situação em que a leitura é produzida, divergem no que diz respeito à natureza das restrições impostas a essa construção. No caso da segunda visão, elas não dizem respeito aos elementos lingüísticos ou ao texto, mas "... às convenções que organizam e delimitam suas instituições", o que pode ser denominado "comunidade interpretativa". Nesse contexto, a própria linguagem é vista como convenção, portanto, também sujeita a essas restrições. Dentro dessa perspectiva, portanto, poder-se-ia afirmar que "... nenhuma leitura é definitiva, e nesse sentido, cada leitura, sem excluir a do autor, é um acidente do texto" (Paz 1974). Esse posicionamento, devo salientar, torna impossíveis a avaliação e também o ensino da leitura.

Vários são os modelos interativos propostos na literatura (vide McClelland & Rumelhart 1981; Taylor & Taylor 1983; LaBerge & Samuels 1976; Perfetti 1985, 1986a, 1986b; Stanovich 1980; Koch & Travaglia 1989; Cavalcanti 1983, 1989; Moita Lopes 1986, dentre outros) e uma revisão abrangente foge do escopo desta seção. A seleção de alguns modelos tem como justificativa sua relevância para este trabalho, permitindo uma caracterização mais completa de seu arcabouço teórico.

O modelo de Stanovich (op. cit.) é também denominado de modelo interativo-compensatório. Embora elaborado para a leitura em LM, tem sido usado para explicar o processo de leitura em LE. Esse modelo tem, como ponto de partida, um conceito chave na discussão dos componentes de leitura e também neste estudo, ou seja, o conceito de compensação, que será retomado na Seção 2.3. Para Stanovich, um processo em qualquer nível pode compensar deficiências em outro nível, o que poderia significar o não reconhecimento de um nível limiar de conhecimento, seja ele lingüístico ou esquemático. Samuels & Kamil (1988:32) discutem o conceito de compensação proposto por Stanovich:

(...) se houver uma deficiência em um estágio inicial da análise da palavra impressa, esta poderá ser compensada através de estruturas de conhecimento de ordem mais alta. O leitor iniciante, que é impreciso e lento no reconhecimento de palavras, mas tem conhecimento do tópico do texto, pode usar um processamento descendente para compensar essa deficiência. Por outro lado, para o leitor que é um hábil decodificador, mas não apresenta conhecimentos suficientes sobre o tópico do texto, torna-se mais fácil, simplesmente, reconhecer palavras através de um processamento ascendente (7c).

Um aspecto um tanto negligenciado por todos os modelos até agora apresentados, inclusive os interativos, são os fatores pragmáticos e interacionais (entendidos aqui como o contexto situacional, os interlocutores em si, suas crenças e intenções comunicativas, a função comunicativa do texto). Uma visão restrita de interação, ou seja, aquela que considera apenas os vários níveis de conhecimento do leitor e das restrições do texto, sem levar em conta a interação enquanto um processo de comunicação do leitor e do autor via texto, tem sido culpada por essa lacuna. Como exemplos de modelos que consideram esse componente na interação ou seja, o contexto situacional ou de leitura, cito os modelos de Koch & Travaglia (1989), Cavalcanti (1983, 1989) e Moita Lopes (1986). Basicamente os três concebem a leitura como um processo multidisciplinar, interativo e construtivo.

Koch e Travaglia (op. cit.) apresentam a coerência vista como uma teoria do sentido ou interpretabilidade do texto, que inclui vários elementos, a saber: lingüísticos (seu conhecimento e uso); conhecimento de mundo (bem como o grau em que esse conhecimento é compartilhado pelo autor e leitor do texto) e fatores pragmáticos e interacionais (entendidos aqui como o contexto situacional, os interlocutores, suas crenças e intenções comunicativas, a função comunicativa do texto). Para esses autores "... tudo que afeta (auxilia, possibilita ou dificulta, impede) essa interpretação do texto tem a ver com o estabelecimento da coerência". Uma correlação estrita entre os fenômenos da coerência e compreensão é colocada pelos autores da seguinte forma:

(...) ocorrendo, na interação texto-usuários, a construção de um sentido (ou de uma continuidade de sentidos na conversação ou em textos mais longos), haverá compreensão. Do contrário, por mais organizado que esteja o texto do ponto de vista estritamente lingüístico, a compreensão não se dará e, para o receptor, ele se apresentará como destituído de coerência.

A coerência, definida nesses termos é semântica e pragmática, uma vez que considera os atos ilocucionários que as proposições realizam; não é apenas lógica, resultante de fatores cognitivos, mas também influenciada por fatores interpessoais, como a posição do falante ou do leitor na situação de comunicação, assim como as intenções comunicativas dos interlocutores. Desse modo, portanto, não é uma característica do texto nem do leitor, mas se encontra "... no processo que coloca texto e usuários em relação, numa situação dada ..." (Koch & Travaglia op. cit.) ou na interação entre o texto e leitor, a partir da situação comunicativa.

A proposta de Cavalcanti (op. cit.) para o processo de leitura também se baseia em fontes multidisciplinares. A leitura é, nesse caso, também dependente do jogo de inter-relações entre conhecimento prévio (estruturas de conhecimento e competência comunicativa); sistemas de valores (crenças ideológicas, religiosas, filosóficas e pontos de vista científicos); conhecimento acumulado (resultado dinâmico da interação entre as estruturas de conhecimento e sistemas de valores do leitor e as restrições do texto (organização, ponto de vista, intenções); e ainda de restrições do contexto de leitura ou variáveis de desempenho (foco de atenção imediata, as expectativas, o propósito e a motivação). Especificamente, a autora

(...) investiga um aspecto dos processos de construção de discurso, isto é, os problemas que surgem na criação de significados na interação com o texto. Essa criação de significados vai além da decodificação de palavras e orações em direção à atribuição de força ilocucionária, isto é, refere-se à interpretação pragmática (Cavalcanti 1989:14).

A pragmática da interação leitor-texto é definida como a

(...) como a busca de equilíbrio entre a saliência-textual e a relevância-leitor. É aqui explicada em termos de itens lexicais chaves para a saliência textual e itens contextualmente relevantes para relevância-leitor (Cavalcanti op. cit.:70).

Embora os mesmos componentes estejam presentes nos modelos de Koch & Travaglia e Cavalcanti, convém ressaltar que, enquanto no primeiro, a compreensão é abordada sob o ponto de vista da coerência, no segundo, é vista sob o ponto de vista da relevância. A justificativa apresentada pela autora para a centralidade da relevância (Sperber & Wilson 1981) é que ela capta o ponto de

vista do leitor/ouvinte na interpretação pragmática, geralmente ignorado na literatura de base pragmática (15), que tende a considerar apenas o falante ou leitor, ou sua intencionalidade, na compreensão do enunciado (16). Apesar de focos diferentes, pode-se dizer que as duas propostas são coincidentes, ao considerar o leitor na interação (17).

Uma diferença, entretanto, que torna a proposta de Cavalcanti particularmente relevante para este trabalho é o fato de ter sido desenvolvida dentro do panorama de pesquisa em LE, além de mostrar a criação de significados a partir de itens lexicais chaves ou competência lexical. O conceito de itens lexicais, proposto pela autora, vai além da coesão lexical (Halliday & Hasan 1976) e das relações semânticas (Collerson 1974). Uma distinção é feita entre itens lexicais chaves e itens contextualmente relevantes. Itens lexicais chaves são elementos salientes em relação aos outros elementos no texto, como ilhas de confiança textual. São expressões indexicais recorrentes, cuja função é preservar a unidade do texto, através da coesão topical, servindo de base para a criação da coerência. Dessa forma, o conceito é semântico e pragmático. É semântico por basear-se na saliência assinalada pela coesão léxico-gramatical e de estrutura de informação no texto. É pragmático, por referir-se às expressões indexicais, cuja saliência é negociada em termos da relevância-leitor em direção à criação de coerência. Itens lexicais chaves constituem a espinha dorsal para o estabelecimento de conteúdo proposicional e força ilocucionária, formando fios condutores semânticos, ativando estruturas de conhecimento e sistemas de valores do leitor.

Itens contextualmente relevantes, por sua vez, referem-se às pistas de superfície que o leitor escolhe focalizar em sua interação com o texto. Sua escolha não é uma decisão do leitor, mas depende das variáveis de desempenho, dentre outros fatores. São eles os responsáveis pelo aspecto idiossincrático da leitura. Quando um elo é estabelecido entre itens lexicais chaves e itens contextualmente relevantes, há concordância intersubjetiva na leitura.

O modelo interativo de esquemas proposto por Moita Lopes (op. cit.) assemelha-se ao de Cavalcanti em vários sentidos. Além de ambos terem sido elaborados especificamente para explicar a leitura em LE, de considerarem uma competência discursiva e comunicativa, além da competência lingüística, ambos contêm uma dimensão pragmática, que consiste na negociação do sentido durante a interação e não em abstração, ou seja, em frases isolada. Além do conhecimento lingüístico (níveis sintático, lexical, semântico), o modelo também incorpora um nível esquemático (conhecimento prévio) e um nível pragmático, que é alcançado através de procedimentos interpretativos que fazem parte de um *conceito de capacidade* proposto por Widdowson (1983). Esse conceito é importante (e, particularmente para este trabalho) na medida em que inclui a capacidade de realizar a língua em uso, ou de utilizar os procedimentos de interpretação necessários para a negociação do significado entre os interlocutores. Segundo Moita Lopes, a noção de procedimentos de interpretação (determinadas pelo propósito de leitura, tipo de texto, nível de compreensão, que poderiam corresponder ao que Cavalcanti denomina molduras interativas), parece incluir uma característica

típica da interação social humana, enfatizando, portanto, a relevância do contexto interacional no processo de construção da compreensão.

O significado, nessa visão, portanto, não estaria nem no texto (co-texto ou contexto do texto) ou em sua sintaxe, nem no leitor ou no contexto geral, mas na inter-relação entre esses aspectos. Essa seria, segundo seu autor, uma visão de significado como processo, contrastando com a visão tradicional de significado enquanto produto. Esse conceito corresponderia ao aspecto pragmático da interação proposto por Cavalcanti.

Neste trabalho, portanto, o termo interativo está sendo usado para referir-se à interação entre os vários níveis de conhecimento do leitor assim como à interação entre o leitor e autor via texto. O componente linguístico, assim como todo conhecimento prévio em geral, é visto como um elemento para construção do significado, visando uma interpretação pragmática, e não mais como o único elemento de significado. O conceito de conhecimento *rico* de vocabulário, apresentado na Seção 2.4.3.2 deste capítulo complementará este arcabouço teórico.

Tendo caracterizado o modelo/visão de leitura que fundamenta este trabalho, passo a seguir a uma revisão das evidências sobre o papel do componente linguístico vs. cognitivo nas pesquisas em leitura em LE.

2.3 – O componente linguístico vs. cognitivo nas pesquisas em leitura em LE

Introduzindo esta seção, focalizo algumas questões terminológicas que permearam a discussão na seção anterior, e que serão retomadas nas seguintes. Na literatura sobre leitura, o componente linguístico (18) tem incluído os níveis lexical, sintático, morfológico e semântico, enquanto o cognitivo tem-se referido às habilidades/estratégias para uso do conhecimento prévio, que inclui, por sua vez, todos os outros tipos de conhecimento, excluindo-se o linguístico.

A segmentação entre componente linguístico e cognitivo, adotada na literatura, e usada na revisão apresentada nesta seção, pretende levar em conta os dois tipos de processamento que ocorrem na leitura, ou seja, ascendente e descendente, através de uma associação conhecimento linguístico/decodificação vs. processamento ascendente e recursos cognitivos/outros tipos de conhecimento vs. processamento descendente. Apesar de usada em modelos interativos, essa divisão é, a meu ver, confusa, e ainda parece refletir tendências de processamento linear, separando conhecimento linguístico dos outros tipos de conhecimento prévio, que são incorporados ao componente cognitivo, assim como conhecimento linguístico da capacidade de uso desse conhecimento (também incorporada ao componente cognitivo). As dificuldades envolvidas na avaliação em separado dos dois componentes, assim como o reconhecimento de que mesmo a decodificação, antes percebida como processamento ascendente, também envolve conhecimentos

e habilidades descendentes, poderiam ser vistas como evidências contrárias a essa segmentação.

Dessa forma, uma caracterização da competência necessária à leitura, mais condizente com o que ocorre em um modelo interativo, seria de um componente composto de duas dimensões: uma de conhecimento prévio que o leitor traz para o texto, que incluiria todos os tipos de conhecimento, isto é, lingüístico, esquemático, de mundo, e uma outra dimensão de capacidade de uso desses conhecimentos. As duas dimensões seriam necessárias, mas nenhuma delas suficiente isoladamente.

Algumas considerações adicionais sobre o conhecimento prévio fazem-se necessárias. O conceito de conhecimento prévio é usado em várias disciplinas (vide Levinson 1983; Brown & Yule 1983; Giddins 1976 para uma revisão do conceito nas áreas de Pragmática, Análise do Discurso e Sociologia). Nesse sentido, ele é abrangente, referindo-se, a aspectos como crenças, valores, ideologias, assim como aos vários tipos de conhecimento que o leitor traz ao texto, ou seja, lingüístico, discursivo, enciclopédico, de conteúdo ou do assunto, de mundo, cultural.

Uma crítica a um conceito de conhecimento prévio não diferenciado é apresentada por Fairclough (1985:754). Para o autor, que parte do ponto de vista da análise crítica do discurso (19) a aceitação desse conceito "... coloca a análise do discurso na posição de reproduzir de 'maneira não crítica' certos efeitos ideológicos..." (8c), reduzindo diversos aspectos da interação, tais como crenças, valores, ideologias, assim como conhecimento propriamente dito, em conhecimento. Ele prossegue, fazendo uma distinção entre os conceitos "conhecimento" e "ideologia":

'Conhecimento' implica em fatos a serem conhecidos, codificados em proposições, que estão objetivamente e transparentemente relacionados a elas. Mas 'ideologia' (...) envolve a representação 'do mundo' vista na perspectiva de um interesse em particular, de forma que a relação entre proposição e fato não é transparente, mas mediada pela atividade representacional. Portanto, ideologia não pode ser reduzida a 'conhecimento' sem distorção (Fairclough op. cit.:754) (9c).

Essa concepção de conhecimento prévio é, na opinião do autor, uma concepção cognitiva e psicológica, na medida em que reduz o aspecto sociológico ao cognitivo através da metáfora da competência. Os fatores sociais não aparecem, apenas a "competência social" das pessoas. Esse conhecimento é atribuído a cada participante, em cada interação, como se fosse um conjunto de atributos da pessoa, ou seja, o que ela sabe. Interações, nessa visão, resultam de pessoas vistas como autônomas, de vontade própria, que se juntam e mobilizam suas bases de conhecimento para fazer sentido do discurso. *"O sujeito 'competente' das concepções cognitivas de interação é o sujeito autônomo da ideologia"* (Fairclough op. cit.:755) (10c).

Observa-se, na literatura sobre leitura, entretanto, um estreitamento do conceito de conhecimento prévio. Este, muito freqüentemente, é utilizado em referência ao conhecimento de mundo, ou cultural, ou ainda conceitual ou de conteúdo, do assunto do texto (*content schemata*),

excluindo assim, o conhecimento lingüístico, que geralmente é considerado separadamente, como mostrei acima. É com esse sentido menos abrangente, portanto, apenas de conhecimento conceitual que ele é operacionalizado, através de um teste, no estudo de produto deste trabalho (vide Capítulo III). Entretanto, devo salientar, que essa visão restrita de conhecimento prévio não é aquela proposta nos modelos de Cavalcanti e Moita Lopes, que servem de arcabouço teórico para este estudo.

Muitos são os estudos que avaliam a importância do conhecimento prévio para a leitura. Essa relação tem sido estudada amplamente pelos proponentes da teoria de esquemas (Bartlett 1932; Minsky 1975; Schank & Abelson 1977; Rumelhart & Ortony 1977, dentre muitos outros). Há pesquisas que consideram a importância desse conhecimento também para uma situação de LE (Carrell 1983; Lee 1986; Hock 1987; Johnson 1982; Hudson 1982; Adams 1982; Hammadou 1991). Nota-se, nesses estudos, que a operacionalização do conhecimento prévio varia desde a simples classificação prévia pelo pesquisador de um assunto como familiar/não familiar, até uma avaliação mais abrangente do assunto.

Antes de rever alguns desses trabalhos, entretanto, considero importante salientar uma diferença aparente entre as situações de L2 e de LM, no que se refere ao uso desse conhecimento. Enquanto em LM, o conhecimento do assunto já pressupõe um conhecimento de vocabulário, em L2 ou LE, isso parece não acontecer. É possível conhecer-se um assunto (em LM) e não saber o significado das palavras envolvidas (em LE). A ativação desse conhecimento parece depender, em grande parte, de conhecimento de vocabulário e portanto, não tão simples como tem sido sugerida. Os estudos de Carrell (op. cit.) e Adams (op. cit.) poderiam oferecer alguma evidência, embora tênue, para a minha observação. Carrell observa que os leitores não-nativos de seu estudo não mostraram efeitos significativos de conhecimento prévio, contrariamente ao que ocorre com leitores nativos. Entretanto, Lee (op. cit.), através de uma réplica desse estudo, chega a resultados que apontam para uma direção oposta, mostrando que as evidências de Carrell são influenciadas por uma questão metodológica, uma vez que as memorizações foram feitas na língua alvo, limitando a habilidade dos sujeitos em demonstrar sua compreensão. Adams, entretanto, mostra que os efeitos do conhecimento prévio (através de informações dadas sobre o tópico antes da leitura de um texto) foram maiores em LM (inglês, no caso) do que em L2 (francês). Nesse estudo, as informações sobre o tópico do texto haviam sido dadas em inglês, mesmo quando os leitores estavam lendo o texto em francês.

O artigo de Hammadou (op. cit.) apresenta os resultados de dois estudos paralelos, um em francês e outro em italiano como L2, que tem como perguntas de pesquisa, basicamente, a mesma questão dos outros artigos acima. Entretanto, Hammadou também acrescenta um componente de processo, uma vez que está interessado nas diferenças qualitativas na inferência em L2 de acordo com familiaridade no tópico e proficiência lingüística. A abordagem usada na operacionalização do conhecimento prévio foi tentar maximizar uma possível escala de diferenças de familiaridade,

selecionando três tópicos de uma lista de treze apresentados, a um grupo de sujeitos iniciantes e avançados. Depois de selecionados três tópicos, foi solicitado a um segundo grupo de sujeitos que dessem um número de 1 a 3 para cada um dos tópicos (auto-avaliação) ou que classificassem esses tópicos de acordo com seu conhecimento. A familiaridade seria resultante do conhecimento prévio. Esse segundo grupo de sujeitos também foi solicitado a ler três textos em L2, e a recontar, por escrito, cada uma deles em inglês como LM. Os resultados mostram que os tópicos familiares não foram mais bem compreendidos que os não familiares, contrariando resultados anteriores apresentados na literatura. Além disso, apontam a proficiência lingüística como o fator mais significativo nas duas linguas e, a familiaridade, como não significativo. Os resultados também mostram que, embora os alunos mais avançados tenham compreendido mais, as diferenças entre os protocolos, baseados nos tópicos mais familiares e nos não familiares, não foi relevante. Por outro lado, embora os leitores menos proficientes mostrassem sinais de compreensão diminuída, esta não parece ter sido causada por menor familiaridade. Entretanto, uma análise qualitativa dos protocolos parece revelar que houve uma quantidade muito maior de inferências que se desviavam do texto original nos protocolos dos leitores iniciantes, o que pode ser interpretado como uma tentativa de construção de significado, a partir do conhecimento prévio. Uma medida direta dos processos internos de compreensão e inferência, por exemplo, através de uma técnica mentalista, poderia ter sido útil nesse sentido. Uma das conclusões do autor é que

Encontrar um método direto para identificar conhecimento prévio claramente permanece um problema. Neste estudo, as contradições entre as avaliações da familiaridade do tópico e desempenho real na tarefa de recontar levantam a questão se a simples auto-avaliação do sujeito com relação à familiaridade do tópico A em relação a B era realmente precisa. Uma razão para a falta de significância da variável familiaridade, portanto, podem ser imprecisões em tais auto-avaliações (Hammadou op. cit.:31) (11c).

Um outro problema, além da dificuldade relacionada ao método, diz respeito à avaliação separada de dois aspectos, a proficiência lingüística e o conhecimento prévio, que, na realidade, se apresentam inter-relacionados. Nesses estudos focalizados, os dois aspectos aparecem como competidores, embora estejam, na realidade, em uma relação de complementaridade. Um reconhecimento do papel do conhecimento prévio na proficiência lingüística, através do contexto, fica evidente a partir da afirmação de Hammadou (op. cit.:28):

O tema conhecimento prévio e seu papel na compreensão é repetido em muitas discussões sobre proficiência lingüística. A maioria das teorias atuais de proficiência lingüística em LE incorporam o conhecimento prévio no domínio do 'contexto', tanto influenciando a proficiência ou

como uma parte de seu construto (...) Teorias lingüísticas atuais reconhecem que leitores mais proficientes dependem menos do conhecimento prévio e mais do conhecimento dos elementos da língua. Proficiência é definida, em parte, como uma medida de dependência do contexto (...) (Hammadou op. cit.:28) (12c).

A referência de Hammadou ao contexto, assim como sua relevância na compreensão, levam-me a uma consideração mais detalhada desse conceito. Embora essa discussão possa, em princípio, parecer redundante, uma vez que um posicionamento com relação à essa noção é praticamente determinado pela adesão a um modelo ou visão de leitura, considero importante retomá-la, dado que se constitui um dos focos de investigação neste trabalho.

Para Goodwin & Duranti (1992), que abordam a questão sob um ponto de vista abrangente, incluindo outras áreas além da leitura, uma definição formal, precisa e técnica de contexto não parece ser possível na situação atual, apesar das intravisiões analíticas importantes que ela poderia trazer. Os autores reconhecem que o termo é confuso, apresentando significados diferentes, dentro dos paradigmas de pesquisa, e, até mesmo, dentro de tradições particulares. Dessa forma, parece ser definido mais pela prática em cada caso, pelo uso do conceito em problemas analíticos particulares, do que por uma definição formal.

Na área de leitura, a noção de contexto estreita que se pode depreender das visões de leitura descendente e ascendente aparece refletida no próprio direcionamento da pesquisa. Um foco em elementos do texto (contexto do texto) ou em elementos do conhecimento prévio (contexto do leitor) tem sido responsável por uma polarização de estudos, que consideram ora o papel ou a contribuição de um conhecimento lingüístico ou lexical para o reconhecimento de palavras e compreensão, ora a inferência de palavras e compreensão através do contexto. As duas áreas de pesquisa são vistas como independentes, como se um vocabulário sólido eliminasse a necessidade de contexto (visão ascendente) e como se a inferência de palavras através do contexto fosse possível sem conhecimento de vocabulário (visão descendente) (20).

Algumas caracterizações explícitas da noção de contexto tem sido propostas na literatura, algumas delas a partir de estudos sobre inferência lexical. Nesse caso, o contexto tem-se referido ao conjunto de fontes de informação usadas pelo leitor durante a inferência de palavras (vide Clarke & Nation 1980; Nation & Coady 1988; Stenberg & Powell 1983; Ames 1966; Van Parreren & Schouten-van Parreren 1981; Haarstrup 1989; Scott 1990; Henriques 1990; alguns desses estudos são retomados na seção deste capítulo referente à inferência lexical). Embora a discussão seja, nesse caso, centrada na função do contexto para a inferência de palavras, é importante ressaltar sua função mais ampla dentro da compreensão, ou seja, em processamentos de nível mais alto. Esses dois usos do contexto por leitores de nível diferentes de proficiência são discutidos em várias partes deste trabalho, constituindo o foco da investigação apresentada no Capítulo V.

Van Parreren & Schouten-van Parreren (op. cit.), em um trabalho que analisa inferências

lexicais mal-sucedidas, apresentam um conceito de contexto que inclui os seguintes níveis, organizados de forma hierárquica: sintático (estrutura da oração), semântico (contexto imediato ou mais amplo da palavra assim como, quando necessário, o conhecimento de mundo), lexical (forma da palavra) e estilístico (apreciar o sentido exato da palavra, principalmente de um cognato). O leitor eficiente já entra no nível mais apropriado, ou seja, é capaz de identificar qual o nível de contexto necessário para cada situação, sem passar por todos os níveis da hierarquia. Como pode-se observar, o conceito de Van Parreren & Schouten-van Parreren (op. cit.) parece limitar-se ao nível de co-texto ou às restrições do texto, embora façam referência ao conhecimento de mundo, necessário por vezes ao reconhecimento da palavra. Não fica muito claro o que os autores querem dizer com "efeito estilístico de um cognato".

Uma outra contribuição para essa caracterização de contexto ou de fontes de informação é fornecida por um estudo conduzido com brasileiros. Para Scott (1992), o contexto se utiliza dos elementos do co-texto e também o conhecimento prévio (*background*), que é organizado em esquemas (vide Rumelhart & Ortony 1977). Esse conhecimento inclui conhecimento de mundo e da língua, além de como o texto funciona até o momento, o que seria correspondente ao que Cavalcanti (1983, 1989) denomina conhecimento acumulado. Como mostra Kato (1982),

(...) uma versão assimilada do sentido do co-texto é presente de alguma forma na cabeça do leitor. O leitor não pode deixar de interpretar o co-texto da maneira como ele o vê, e depois desse ponto, o item está no co-texto e na memória.

Scott (op. cit.:195) postula uma taxonomia de níveis de co-texto:

Quadro 2.1 – Níveis de co-texto em torno da palavra desconhecida (Scott 1992)

-
1. A estória até o momento
 2. O capítulo ou seção
 3. O parágrafo
 4. A sentença
 5. Algumas palavras de cada lado
 6. As palavras imediatamente vizinhas
-

Uma taxonomia de fontes de informação, ainda mais abrangente, é apresentada por Haarstrup (1987, 1989), em estudo com leitores dinamarqueses de dois níveis de proficiência (alta e baixa) lendo inglês como LE (vide Quadro 2.2 abaixo). Sua conclusão é que leitores mais proficientes usam uma combinação de mais de uma fonte de informação num processamento paralelo.

Quadro 2.2 – Taxonomia de fontes de conhecimento (Haarstrup 1987)

Contextual	Intralingüístico	Interlingüístico
I. O texto	I A palavra testada	I. L1 (dinamarquês)
1. Uma única palavra do contexto imediato	1. Fonologia/ortografia	1. Fonologia/ortografia
2. O contexto imediato	2. Morfologia	2. Morfologia
3. Uma parte específica do contexto além da sentença da palavra testada	a. prefixo	3. Léxico
4. Uso global do texto	b. sufixo	4. Colocações
	c. radical	5. Semântica
	3. Léxico	
	4. Classe da palavra	
	5. Colocações	
	6. Semântica	
II. Conhecimento de mundo	II. A sintaxe da sentença	II. LE (latim, alemão, francês, etc.)
		1. Reflexões gerais
		a. Reflexões sobre a origem da palavra
		b. Palavra testada pronunciada em LE
		2. Morfologia
		3. Léxico
		4. Semântica

Embora o conceito de Haarstrup (op. cit.) e de Scott (op. cit.) sejam semelhantes, na medida em que incluem o co-texto, conhecimento lingüístico e de mundo, considerando, portanto, um contexto do texto e um contexto geral, o de Haarstrup ainda inclui fontes de conhecimento interlingüísticas (da LM e de outras línguas) e uma consideração mais detalhada da situação de inferência, referindo-se a fatores sócio/psicológicos como atitudes e motivações do sujeito, dentre outros. Embora a inclusão de pistas interlingüísticas (LM e outras) seja um aspecto bastante importante desse modelo, pode-se observar alguns aspectos redundantes ou de superposição. A segmentação em fontes contextuais e fontes intralingüísticas inclui, como fontes de conhecimento, os elementos do co-texto (a meu ver, mais bem caracterizados como *restrições* ou *pistas* do texto) além de considerar conhecimento prévio como contexto. Além disso, embora o conhecimento lexical pressuposto pela autora seja particularmente interessante, na medida em que envolve vários aspectos (fonologia, morfologia, dentre outros), estes são vistos enquanto "propriedades apenas das palavras", (21) deixando de considerar que seu uso está relacionado às restrições do contexto. Em outras palavras, o que quero dizer é que a utilização, por exemplo, de conhecimentos sobre a colocação de uma palavra ou suas associações sintagmáticas não podem ser considerados em abstração mas sim dentro de um co-texto, da mesma forma que conhecimento semântico ou lexical (significado) da palavra não pode ser visto isoladamente de outras palavras, envolvendo também conhecimento de mundo. Dessa forma, não haveria necessidade de se separar entre pistas contextuais e pistas intralingüísticas. Uma visão de conhecimento, que não incorpora uma dimensão de uso, ou de como esse conhecimento é acessado na inferência de palavras, assim como uma visão de contexto definida ora pelos elementos do co-texto, ora pelo conhecimento prévio, parecem estar refletidas nessa

proposta.

O estudo de Henriques (1990) analisa o contexto na perspectiva da sala de aula de leitura. Seu objetivo é verificar, a partir da lista de catorze elementos no Quadro 2.3 abaixo, aqueles que seus alunos consideravam mais úteis na inferência de palavras, depois de terem sido expostos a quinze horas de aula para o desenvolvimento dessas habilidades.

Quadro 2.3 – Elementos do contexto (Henriques 1990)

1. O contexto da mesma sentença
 2. O texto todo
 3. Cognatos
 4. Conhecimento prévio do tópico
 5. Contexto das sentenças precedentes e seguintes
 6. Conhecimento de português
 7. Listas tipográficas
 8. Explicações ou descrições oferecidas pelo texto
 9. Repetição da palavra desconhecida em contextos diferentes
 10. Exemplos dados no texto
 11. Antônimos fornecidos pelo texto
 12. Conhecimento gramatical (morfologia da palavra e sintaxe)
 13. Sinônimos fornecidos pelo texto
 14. Os sons das palavras
-

Os alunos tinham, por tarefa, julgá-los de acordo com uma escala do tipo Likert de 5 pontos, desde *sempre usado* até *nunca usado*. A preferência foi para o contexto sentencial imediato, ajuda do texto até o momento ou conhecimento acumulado, conhecimento prévio do tópico e da LM, no caso, português. Uma outra conclusão, que parece corroborar o estudo de Haarstrup é que os alunos de Henriques também usaram um processamento paralelo, ou foram capazes de usar mais de um tipo de elemento.

Apesar de mais detalhados, os aspectos de Henriques coincidem com os de Haarstrup, envolvendo aspectos de co-texto e conhecimento de mundo, do contexto interlingüístico e intralingüístico. A mesma redundância acima salientada ou uma superposição de elementos, também é observada neste caso. Muitos dos elementos têm, realidade, um mesmo tipo ou tipos de conhecimento subjacentes, o que torna difícil para o leitor selecionar o mais útil ou eficiente. Por exemplo, o reconhecimento do contexto da oração e das orações precedentes e seguintes, dos cognatos, das pistas tipográficas, exemplos, antônimos, sinônimos, sons, dependem de mais de um tipo de conhecimento lingüístico ou lexical, assim como também dos elementos do co-texto; a identificação de cognatos está relacionada ao conhecimento lingüístico nas duas línguas, portanto, conhecimento de português e conhecimentos da língua alvo, além das restrições do co-texto.

As restrições que faço a respeito desses modelos poderiam ser justificadas, uma vez que todos eles limitam-se a processos de inferência de palavras, não considerando a construção do sentido ou da compreensão. Nesse sentido, a proposta de Goodwin & Duranti (1992) é

particularmente relevante. Para esses autores, o contexto é usado quando se quer interpretar algum tipo de fala. Entretanto, nem sempre esse é o caso. As vezes, a fala pode ser usada para servir de contexto a algum outro tipo de comportamento. Portanto, a sugestão de Goodwin e Duranti (op. cit.) é tratar como *evento focal* o que está sendo contextualizado e como *pano de fundo* o que serve de contexto. Cabe ao analista delinear esses papéis, dependendo de seus objetivos.

O evento focal, contudo, não parece suficiente para uma interpretação adequada. Há necessidade de uma análise de outros aspectos, tais como cenário cultural, situação de fala, pressuposições compartilhadas, dentro dos quais o evento está inserido. Contexto, dessa maneira, é visto como uma moldura que envolve o evento examinado.

Para uma análise do contexto visto sob essa perspectiva é necessário, na opinião desses autores:

- Tomar como ponto de partida a perspectiva dos participantes cujo comportamento está sendo analisado, e não a perspectiva do analista;
- Descrever o conhecimento sócio-histórico que o participante traz para a situação;
- Levar em conta a capacidade dos seres humanos de reformular dinamicamente o contexto, que oferece organização para suas ações dentro da própria interação. Assim, o contexto está sendo constantemente modificado.

Transpondo as observações de Duranti & Goodwin, adicionadas às de Haarstrup (op. cit.), para a situação de leitura deste trabalho, é necessário, portanto, considerar:

- A situação ou contexto social onde ocorre leitura, envolvendo tipo de texto, língua, nível de compreensão, propósito da leitura;
- O conhecimento sócio-histórico, cultural, crenças, valores, assim como outros tipos de conhecimento prévio e habilidades que o leitor traz para a situação de leitura. Estes dois primeiros itens, portanto, constituem o contexto geral;
- Informações ou restrições morfológicas, sintáticas e discursivas do texto, ou o contexto local ou dentro do texto.
- Reformulação dinâmica do contexto, ou seja, a inferência de uma palavra altera o contexto, passando a fazer parte do conhecimento que o leitor traz ao texto (conhecimento acumulado de Cavalcanti). Assim, o contexto é constantemente reformulado a partir do seu uso.

Para finalizar esta discussão, e, de certa forma, reiterar o conceito de contexto deste estudo, não poderia deixar de fazer uma observação referente ao seu uso no ensino de leitura. As abordagens que têm por base visões descendentes, mais especificamente as abordagens instrumentais, têm explorado extensivamente a noção de contexto através de estratégias de inferência lexical (vide seções 2.4.3.1 e 2.4.3.2 deste capítulo). Observa-se, nesse sentido, a tendência de se utilizar de contextos denominados *ricos* ou *grávidos* (*pregnants*), que são contextos redundantes, especialmente escolhidos por fornecerem ao leitor poucas restrições, ou várias possibilidades de uso, ou pistas, para

a inferência do significado das palavras, podendo ser acessados de várias maneiras. Não entrando por ora no mérito dessa estratégia, ou melhor, deixando para mais adiante a crítica que se pode fazer em relação à eficiência de seu uso, não posso deixar de salientar, entretanto, que a denominação contexto *rico* é vista sob a ótica do leitor ideal, que tem os vários níveis de conhecimento necessários para acessar as várias pistas que estão no texto. Sem esses conhecimentos, as informações do texto não podem ser consideradas pistas ou contexto. Este não é, portanto, independente do leitor ou independente do texto, mas criado a partir da relação entre os elementos do conhecimento do leitor e pistas do texto. Ele não está nem na "cabeça do leitor" nem no texto, mas resulta de uma interação entre esses dois conjuntos de conhecimento. Uma melhor caracterização desse contexto, portanto, seria *potencialmente rico* ou *potencialmente redundante*, o que salientaria a necessidade de se analisar o contexto na perspectiva do leitor e não do analista ou do leitor ideal, proficiente. O comentário de Nation & Coady (1988:) sobre um estudo de Liu & Nation (1985) parece ilustrar o contexto visto na perspectiva do analista ou ainda do leitor ideal:

O importante corolário é que se os aprendizes citados foram capazes de adivinhar a maioria das palavras, então, as pistas necessárias estão ali, para outros aprendizes, talvez menos dotados, também usarem. Não é um objetivo irrealista esperar que os aprendizes possam inferir 60 e 80 por cento das palavras desconhecidas em um texto, se a densidade de palavras desconhecidas não for muito alta (13c).

Tendo esclarecido as noções de conhecimento prévio e contexto na literatura, assim como convencionalizado seu uso neste trabalho, apresento a revisão da literatura que considera a contribuição do componente linguístico e do componente cognitivo para a leitura.

Um panorama bastante claro das várias tendências e hipóteses existentes na área sobre a natureza da leitura em LE é apresentado em um artigo de Alderson (1984) intitulado *Leitura em língua estrangeira: proficiência em leitura ou proficiência linguística? (Reading in a foreign language: a reading problem or a language problem?)*. Esse artigo pode ser considerado um marco no desenvolvimento das pesquisas em leitura em LE, incluindo, numa mesma discussão, questões que eram consideradas separadamente. É particularmente relevante para esta revisão, uma vez que os posicionamentos com relação à natureza da leitura em LE são também posicionamentos a respeito da contribuição do conhecimento linguístico e do cognitivo. A partir dessa época, vários estudos tiveram como foco essa questão (vide Carrell 1991; vide também Leffa 1984; Maciel 1989; Barbosa 1990) para uma discussão desse aspecto em um contexto de ensino/aprendizagem de inglês no Brasil) (22).

Basicamente são três as hipóteses levantadas:

1. Problemas de leitura em LE são resultado de uma proficiência em leitura insuficiente em LM.

Essa hipótese parece se fundamentar no conceito de universais de leitura de um modelo descendente, que prevê a transferência de uma proficiência em leitura (*reading abilities*) (23) da LM para a L2 ou LE. O componente cognitivo é enfatizado em detrimento do componente lingüístico.

2. Problemas de leitura em LE são resultado de um conhecimento inadequado da língua alvo, ou seja, de lacunas na proficiência lingüística. Essa hipótese coloca uma ênfase bastante grande nessa proficiência, mostrando, talvez, uma visão de leitura como decodificação, de acordo com um modelo ascendente. Nessa hipótese não é clara a contribuição de outros componentes, principalmente de habilidades de leitura.
3. Problemas de leitura em LE são resultantes da falta de habilidades de LM na leitura em LE, pela falta de conhecimento da LE. Bons leitores em LM serão também bons leitores em LE, se tiverem um nível limiar de proficiência lingüística. Essa hipótese reconhece a importância do conhecimento lingüístico e de proficiência/habilidades de leitura, sendo mais consistente, portanto, com uma visão de leitura interativa.

Uma revisão das evidências em favor da hipótese 1, ou seja, aquela que considera a leitura deficiente em LE apenas um problema de leitura ou de proficiência ou fluência, mostra limitações principalmente pelo fato de serem provenientes de estudos com bilíngües, e não com leitores de L2 ou LE, além de não terem sido realizados para provar essas hipóteses (vide Barrera-Vasquez 1953; Modiano 1966; Barik & Swain 1975; Cowan & Sarmad 1976; vide também Cummins 1979, para as limitações desses estudos).

As evidências para a hipótese 2 são mais fortes do que para a hipótese 1, como mostram vários estudos (vide Alderson, Bastien e Madrazo 1977; Aron 1980). As evidências ainda mais fortes, entretanto, são para a hipótese 3, apresentadas por Cziko (1980), Cooper (1984), Clarke (1978), Hock (1987) e Carrell (1991).

Em um estudo comparativo entre as estratégias usadas por leitores de níveis diferentes de proficiência lingüística em inglês e aquelas usadas por leitores nativos, Cziko (op. cit.) conclui que as estratégias estão relacionadas ao nível de competência lingüística do leitor. Enquanto leitores com competência lingüística que se assemelha a de um falante nativo ou avançada usam uma estratégia interativa, baseando-se, ao mesmo tempo, em informações gráficas e contextuais, leitores com uma competência lingüística limitada são incapazes de se utilizar de estratégias que possuem como leitores proficientes em LM, como por exemplo, a habilidade de se utilizar de restrições textuais (*discourse constraints*), usando uma estratégia unidirecional (ou ascendente), que se utiliza unicamente de informações gráficas. Infelizmente, como salienta Alderson (1984), apesar da forte correlação apresentada, não se pode concluir que a competência lingüística insuficiente é a causa desse comportamento, uma vez que Cziko não oferece informações sobre a competência de leitura em LM de seus sujeitos nem a possibilidade de se comparar a habilidade de leitores de níveis de competência

lingüística diferentes com sua habilidade de leitura nas duas línguas. Uma outra limitação para os resultados de Cziko deve-se ao fato de ter usado a análise de lapsos orais (*miscue analysis*) como único método para a coleta de dados. Esse método tem sido criticado, por partir de um pressuposto não testado, ou seja, de que a leitura silenciosa e a leitura oral são equivalentes, quando, na realidade, as estratégias envolvidas podem ser muito diferentes.

O estudo de Hock (op. cit.) tem como objetivo investigar a contribuição da proficiência lingüística e do conhecimento conceitual para a leitura de alunos universitários de três áreas diferentes (Medicina, Economia e Direito) na Malásia. Para avaliar a proficiência lingüística, foi usada uma bateria de testes (EPTB) desenvolvida por Alan Davies para uso em universidades inglesas. O conhecimento conceitual foi operacionalizado de duas maneiras: através de uma medida da familiaridade dos sujeitos com sua área e com o tópico em discussão. Nesses testes, os sujeitos tinham que avaliar a importância dos fatos específicos ao tópico. Foram elaborados, portanto, testes de conhecimento conceitual para cada uma das áreas. O teste de leitura propriamente dito consistiu em testes do tipo cloze, nas três áreas, elaborado seguindo o critério de apagamento fixo, de 10 palavras, ou seja, com lacunas a cada 10 palavras, corrigida pelo método de palavra exata (vide Capítulo III para uma discussão do método cloze). Os sujeitos das três áreas foram submetidos aos testes de proficiência, a todos os testes de conhecimento conceitual e de leitura (três áreas). Assim, cada sujeito leu textos familiares (de sua própria área) e não familiares (de áreas diferentes).

Através da análise de regressão linear pode-se concluir, com relação aos textos familiares que

(...) embora a compreensão de um texto possa ser prevista, em grande parte, pela quantidade de exposição que se tem sobre o assunto e o domínio da linguagem do texto, é realmente este último que determina quanto o leitor em LE pode extrair e interpretar do significado do texto (Hock op. cit.:12) (14c).

Com relação aos textos não familiares, *"... é a competência lingüística da língua alvo que prediz o sucesso na compreensão ao invés do conhecimento prévio sobre o assunto do texto"* (Hock op. cit.: 14) (15c).

Cooper (op. cit.), analisando as dificuldades lingüísticas de leitores de níveis diferentes, lendo em inglês como LE, conclui que leitores com vocabulário deficiente (especialmente vocabulário subtécnico e marcadores sentenciais) encontram-se em uma posição de grande desvantagem, que se caracteriza pela incapacidade em estabelecer as relações semânticas entre as palavras e a inability em usar pistas lingüísticas para inferir o significado dessas palavras. Esses leitores fracos descritos por Cooper assemelham-se àqueles descritos por Hosenfeld (1977b), ou seja, presos aos significados das palavras e incapazes de inferir.

A conclusão de Clarke (op. cit.) é bastante relacionada com a de Cziko e Cooper. Entretanto, além de enfatizar a importância de uma competência lingüística básica para a compreensão, reconhece a existência de um nível limiar na compreensão, que é colocado exatamente por essa competência lingüística insuficiente. Comparando as estratégias de leitores bons e fracos lendo inglês como LE e espanhol como LM, o autor conclui que um limitado controle sobre a língua bloqueia o sistema do bom leitor, fazendo-o reverter a estratégias menos eficientes, ou seja estratégias que se apóiam em informações textuais, inibindo, dessa forma, a habilidade do leitor em ativar, por exemplo, esquemas de nível mais alto.

Antes de salientar algumas das limitações do estudo de Clarke, exploro mais cuidadosamente a noção de *nível limiar*. Embora ela já tenha sido mencionada no Capítulo I e na Seção 2.2 deste capítulo, considero necessário retomá-la nesta seção, dada a complexidade do conceito e sua aplicação neste trabalho.

A hipótese do *nível limiar* ou efeito limiar parece ter sido colocada originalmente por Cummins (1979) para se referir a um nível mínimo de competência lingüística, necessário para que bilíngües possam usufruir dos benefícios do bilingüismo. Segundo Alderson (op. cit.), a noção baseia-se em um levantamento de estudos aparentemente contraditórios sobre os efeitos do bilingüismo. Através da re-interpretação e comparação dos dados, Cummins mostra que na realidade, os estudos que haviam mostrado efeitos positivos estavam associados a bilíngües que apresentavam competências equilibradas nas duas línguas, em contextos de adição, onde os bilíngües, em geral, apresentam altos níveis de competência em L2. Para Cummins (op. cit.:29), portanto, "... alunos que conseguiram alcançar um limiar podem ter um desempenho bem diferente em tarefas cognitivas e acadêmicas quando comparados com aqueles que não conseguiram alcançar o limiar" (16c).

Desde então, a noção tem sido usada em contextos diferentes. Principalmente duas variações principais podem ser identificadas. Uma delas refere-se ao trabalho proposto pelo Conselho da Europa, com respeito ao desenvolvimento de um modelo para a especificação de objetivos para o ensino de LE. Levando-se em consideração os objetivos de um grupo de aprendizes adultos em potencial, ou seja, a sobrevivência em LE como turistas, ou em contatos com estrangeiros em seu próprio país, foram definidas especificações do que seria minimamente necessário, em termos de conhecimento e habilidades lingüísticas, para esses aprendizes. Essas especificações, que receberam o nome de *Threshold Level* (van Ek 1976) corresponderiam, portanto, a um nível básico de proficiência para um propósito específico (24).

Um outro uso da noção tem sido associado aos estudos sobre a *Hipótese do curto-circuito* (*Short-circuit hypothesis*) proposta por Clarke (op. cit.) e Cziko (op. cit.), referindo-se, nesse caso, a um teto de conhecimento lingüístico (*linguistic ceiling*) ou um nível mínimo de conhecimento lingüístico necessário para a compreensão. Segundo essa hipótese, portanto, leitores em LE só seriam

capazes de ler em LE da mesma forma que em LM, depois de terem alcançado um nível limiar de competência lingüística, uma vez que, para autores como Coady (1979), por exemplo, a L2 depende de estratégias de processamento que envolvem conhecimento substancial da língua alvo.

Os níveis limiares, entretanto, não têm sido vistos apenas como limiares lingüísticos em relação à proficiência necessária para a compreensão em leitura. Também podem ser vistos em relação à inferência (conhecimento lingüístico ou lexical mínimo ou ainda apenas extensão do vocabulário) para inferir palavras.

Quando Johns (1978:157) afirma que

(...) parece haver uma 'efeito limiar' quando mais do que aproximadamente 50 em cada 1,000 palavras são desconhecidas. Por causa desse limiar, a percepção das estruturas gerais pode ser efetivamente bloqueada, o que, por sua vez, significa que não há contexto suficiente para permitir a inferência bem sucedida (...) (17c).

refere-se um nível lexical para a inferência de palavras.

O mesmo pode-se dizer de Clarke & Nation (1980:212):

Nossa experiência mostra que aprendizes com um vocabulário de aproximadamente 3,000 são capazes de inferir, em média, 60-70% das palavras desconhecidas do texto. Alguns aprendizes regularmente conseguem por volta de 80% (18c).

ou ainda do estudo de Haarstrup (1989) e de Scott (1990), que serão apresentados na Seção 2.4.3.2 deste capítulo.

Algumas observações sobre o conceito de nível limiar e seu uso podem ser feitas a partir desses estudos. Observa-se, em primeiro lugar, a tendência de se considerar um limiar lingüístico quando a compreensão está envolvida, e um limiar lexical, no caso da inferência, embora as duas modalidades, em princípio, poderiam aplicar-se tanto à compreensão como à inferência. Também pode-se notar que o limiar, nesse caso, refere-se à extensão total do léxico do leitor e à sua influência na inferência de palavras em um texto, o que implicaria, portanto, em um levantamento das palavras conhecidas do aprendiz em termos absolutos. Esse levantamento, em geral difícil de ser operacionalizado, limitado por questões metodológicas, é obtido por aproximações, através de testes do tipo listas de palavras (vide Seção 2.4.3.3 para uma discussão sobre instrumentos para a avaliação do vocabulário). Uma outra maneira de caracterização de um nível limiar seria a partir de um levantamento das palavras conhecidas do leitor em um texto, que são avaliadas pela influência que exercem na inferência e/ou na compreensão. Essa seria, portanto, uma estimativa em termos relativos, que poderia servir de base para generalizações sobre a extensão total do léxico do leitor.

Entretanto, devo ressaltar que não é com esse sentido que os termos *limiar absoluto* e *limiar relativo* são empregados na literatura. Um nível limiar absoluto é aquele invariável, enquanto o relativo pode variar, dependendo das demandas da tarefa, nível ou estágio de desenvolvimento cognitivo e também nível de conhecimento prévio ou conceitual do aprendiz. A meu ver, os termos *estático* e *dinâmico* seriam mais esclarecedores, nesse caso, além de não se confundirem com os termos absoluto e relativo, que proponho acima.

Segundo Scott (op. cit.) pelo menos duas interpretações da noção de nível limiar podem ser feitas. Para facilitar a explicação, o autor usa um gráfico, reproduzido a seguir, mostrando uma situação hipotética ou idealizada:

Taxa de inferência lexical

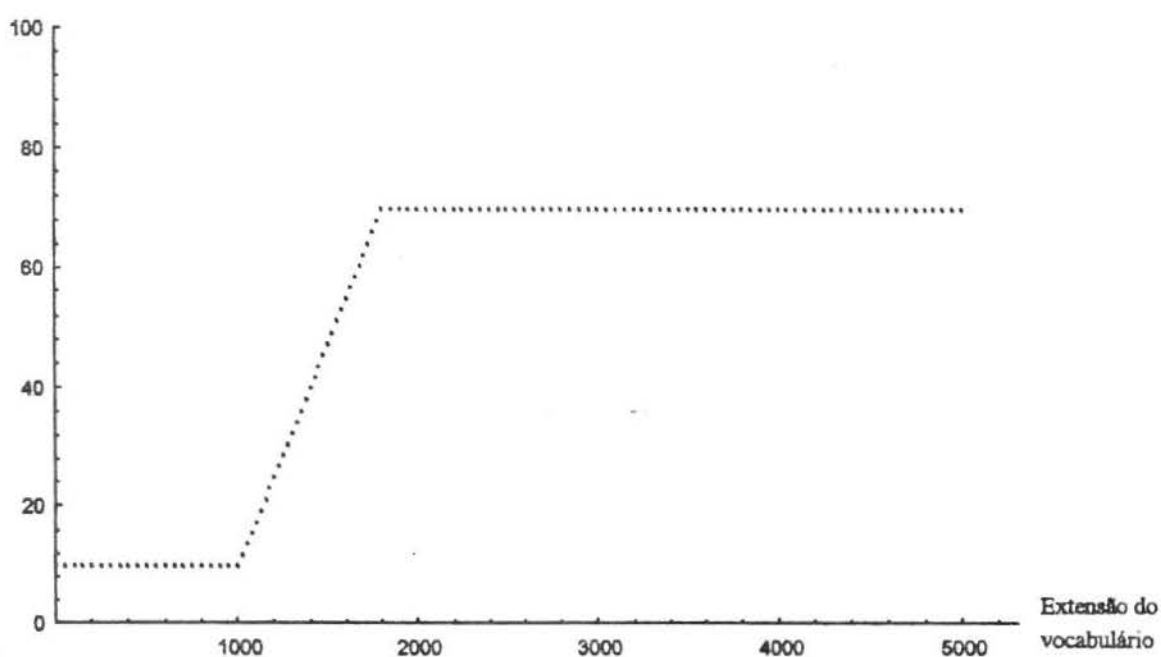


Figura 2.1 – Um efeito limiar hipotético (limiar entre 1000-1800 palavras) (Scott 1990)

Cada ponto na linha representa o desempenho de um sujeito. Em uma situação real, os pontos formam um nuvem. A linha é apenas uma idealização ou uma "tendência central". Uma correlação perfeita pressupõe valores das duas variáveis (no caso, inferência de palavras e extensão do vocabulário) subindo juntos. Entretanto, para o autor, o crescimento na linha não pode ser linear, mas irregular.

Deve haver uma 'saliência' ou 'salto' ('bump') na linha, exatamente como se houvesse um limiar de porta, e alguma diferença entre 'dentro' e 'fora' para quem está em um limiar ... não haverá apenas uma diferença clara entre os sujeitos de cada lado do ponto rotulado 'limiar', mas também uma alteração reconectível na linha do gráfico. A posição dessa própria 'saliência' indica onde uma mudança qualitativa ocorre (Scott op. cit.:116-117) (19c).

Concordo com Scott quando salienta que essa interpretação de nível limiar pode não ser a mesma dada por outros autores. A existência de uma "saliência" ou "salto" pode não ser uma condição necessária para que se possa falar em um nível limiar.

Tal uso meramente significa que os sujeitos que estão na parte esquerda do eixo x (horizontal) tem um desempenho que não é tão bom como os sujeitos mais à direita, e que a diferença é digna de nota. Isso é o que afirma Haarstrup, por exemplo (Scott op. cit.:117) (20c).

Tendo esclarecido a noção de nível limiar, importante para a discussão a seguir, retomo os resultados de Clarke (op. cit.).

Enquanto esse estudo permite, de certa forma, explorar a relação entre a proficiência em leitura em LM e LE no mesmo sujeito (impossível através do estudo de Cziko) (op. cit.), não permite examinar a relação entre a competência de leitura em duas línguas diferentes, e níveis de proficiência na LE, uma vez que todos os sujeitos apresentam, aproximadamente, o mesmo nível de proficiência em inglês. Além disso, algumas limitações são identificadas. Em primeiro lugar, os leitores considerados bons e fracos são definidos puramente com base em seu desempenho em testes do tipo *cloze*. Esses resultados são questionáveis, se for levado em conta que, testes desse tipo, têm sido mais bem caracterizados enquanto medidas de habilidades lingüísticas de nível baixo do que de nível alto (Alderson, op. cit.), apresentando, por isso, uma correlação bastante alta com gramática e vocabulário. Além disso, a única informação com relação ao processo de leitura, nesse estudo, havia sido obtida através da análise de lapsos orais, que, como já salientei, é um instrumento controvertido.

Apesar dessas limitações, os resultados de Clarke são confirmados por Cziko e também por Cummins (op. cit.), que defende a existência de um nível limiar de competência lingüística anterior ao aparecimento de qualquer benefício que possa advir com relação ao bilingüismo. Entretanto, enquanto para Clarke e Cziko esse limiar é absoluto, para Cummins, assim como Alderson e outros, esse nível limiar é relativo, podendo variar dependendo do tipo de tarefa, tipo de texto e desenvolvimento cognitivo e conceitual (ou prévio) do leitor. Dessa forma, o limiar lingüístico seria *"... parte de uma rede maior de informação necessária para o leitor em segunda língua ... necessária mas não suficiente"* (Devine 1988:267) (21c). Estudos em LM confirmam este aspecto, mostrando que a leitura não está totalmente isenta de problemas, mesmo quando os leitores possuem um sistema de conhecimento lingüístico totalmente desenvolvido. Entretanto, algumas variações desse posicionamento podem ser identificadas. Por exemplo, para Hudson (1988), com base em um estudo de fatores determinantes da leitura em LE, embora uma competência lingüística deficiente possa vir a colocar um nível limiar na compreensão, esse efeito pode ser superado, ou compensado, pela habilidade do leitor em ativar o esquema apropriado. Em vista disso, Hudson (op. cit.:85) argumenta a favor também de um nível limiar relativo, que pode ser lingüístico e psicolingüístico:

É necessário verificar quais fatores do segundo componente estão envolvidos na transferência de comportamentos que caracterizam uma leitura eficiente, e como eles interagem com níveis diferentes de um limiar lingüístico do primeiro componente para formar, na realidade, um limiar lingüístico e psicolingüístico (22c).

Assim, fatores do segundo componente ou proficiência/habilidades em leitura estariam afetando o grau com que o nível limiar lingüístico afeta a compreensão, do mesmo modo que um limiar lingüístico estaria afetando o grau com que as habilidades eficientes de leitura são usadas. Isso quer dizer que o limiar pode estar no primeiro componente, isto é, na proficiência lingüística ou no segundo componente, ou seja, na proficiência em leitura:

Apesar de os resultados dos estudos de Cziko e Clarke indicarem que uma competência lingüística básica está envolvida na compreensão em leitura, o que se observa é que bons leitores em LM mantiveram uma vantagem sobre os leitores fracos em LM com mesmo nível de proficiência na leitura em LE. Isso pode ser indicativo de que alguma habilidade não lingüística está em operação. Isto é, parece provável que habilidades de leitura, ou seja, fatores do segundo componente, podem afetar o grau em que o limite de proficiência restringe a compreensão, da mesma forma que um limite de proficiência pode restringir o grau com que habilidades de leitura eficientes são usadas. Nesta simbiose, cada componente pode afetar o grau com que o outro componente afeta a compreensão (Hudson op. cit:84) (23c).

O fato de leitores fracos serem incapazes de se utilizarem de pistas discursivas é, para Hudson, um sintoma de baixa proficiência, juntamente com a inabilidade em produzir um esquema apropriado (por não possuir o esquema, não conseguir ativar o esquema ou não ser suficientemente flexível para mudar de esquema quando o primeiro se prova inadequado). A impossibilidade de se avaliar precisamente o que está ocorrendo parece advir do caráter interativo desses componentes.

A posição de Hudson é, portanto, intermediária entre a de Clarke e Cziko e a de Stanovich (1980), introduzida na Seção 2.2.3 deste capítulo. Stanovich (op. cit.) não parece reconhecer qualquer tipo de nível limiar, uma vez que, para ele, um processo em qualquer nível pode compensar deficiências em outro nível. Stanovich usa, como argumento, o que Samuels & Kamil (1988) consideram uma "anomalia" em alguns estudos, ou seja, porque leitores fracos se utilizam, algumas vezes, de contextos, mais freqüentemente do que leitores bons (contrariando as evidências apresentadas acima por Cziko, Clarke e Cooper):

Sob certas condições maus leitores mostram, quando comparados com leitores bons, maior sensibilidade às restrições contextuais ... A razão pela qual bons leitores são menos sensíveis aos efeitos do contexto é que suas fontes de conhecimento raramente são fracas (...) (24c).

Uma maior necessidade de uso do contexto por leitores fracos também é reconhecida por Van Dijk & Kintsh (1983:23-24).

São os leitores fracos, não os bons, que mais usam o contexto como elemento facilitador para o reconhecimento de palavras. Além disso, não é verdade que bons leitores não levam a decodificação a sério; eles se concentram em cada palavra de contúdo (...) Tem sido mostrado várias vezes que o melhor discriminador entre o desempenho de bons e maus leitores são simples tarefas de identificação de letras e palavras. O que acontece, realmente, de errado com maus leitores, é que eles reconhecem palavras isoladas de maneira imprecisa e lenta demais, além de compensar pela falta de habilidades de decodificação com contexto, adivinhação dependente ou testagem de hipóteses (...) Bons leitores, com suas habilidades superiores de decodificação, podem decodificar letras e palavras rapidamente de maneira ascendente, e, portanto, não precisam recorrer à estratégias de adivinhação — o que realmente está em questão são a velocidade e a precisão de operações de reconhecimento de palavras, independentemente ou fora de contexto (25c).

Entretanto, embora esses autores salientem o uso de contexto por leitores menos proficientes, não reconhecem eficiência nesse uso e, na minha opinião, não poderia ser usado como argumento para justificar uma compensação.

Dessa forma, com base nessas evidências e, também, em algumas observações sem base empírica, da situação foco de investigação desta tese, poderia afirmar que a compensação através do contexto não parece ocorrer da maneira sugerida por Stanovich (op. cit.), ou seja, de forma tão automática e sem prejuízos para a compreensão. Observações a respeito da leitura em LM também sugerem que, essa compensação, pode não ser tão simples quanto parece à primeira vista. Um conceito de compensação proposto por Stanovich, entretanto, em que a compensação pode operar em qualquer direção difere daquele que pode ser inferido a partir de uma visão de leitura descendente, em que a compensação é feita do conhecimento prévio para o lingüístico.

Embora, em princípio, minha posição seja consistente com a de Hudson, quando afirma que o limiar não poderia ser absoluto mas relativo, variando com a tarefa, tipo de texto e outras habilidades/conhecimentos, não concordo com a afirmação de que o limiar também poderia ser psicolingüístico. Sem um conhecimento mínimo de vocabulário ou de conhecimento lingüístico torna-se difícil ao leitor ativar qualquer esquema, ou se utilizar de habilidades ou da proficiência desenvolvidas enquanto leitor em LM. É necessário salientar, entretanto, que minha posição não exclui um conceito de compensação e uma inter-relação entre os vários componentes, uma vez que sua exclusão seria inconsistente com a visão de interação que adoto como arcabouço teórico desta pesquisa. Entretanto, uma compensação, estaria, a meu ver, condicionada a um nível limiar. Em outras palavras, para que a compensação possa ocorrer, é necessário um nível limiar de

conhecimento lexical, visto, portanto, como um fator determinante para que a interação entre os outros componentes possa ocorrer.

Complementando esta seção, que apresentou e avaliou os resultados de estudos a respeito do papel do componente lingüístico e cognitivo na leitura, discuto, a seguir, o componente lingüístico visto sob o ponto de vista do ensino da leitura.

2.3.1 – O componente cognitivo no ensino da leitura: um foco em estratégias

Apesar das evidências mais fortes a favor da hipótese 3, isto é, de que problemas de leitura em LE são resultantes da dificuldade de utilização das habilidades de LM na leitura em LE, pela falta de conhecimento da LE, observa-se, no ensino da leitura, reiterando o que afirmei antes, uma tendência a se enfatizar o uso do conhecimento prévio, através de abordagens que focalizam estratégias (25). Essa tendência parece fundamentada em uma visão de leitura descendente, ou ainda de leitura interativa-compensatória (definida enquanto compensação entre os vários níveis de conhecimento do leitor). O desenvolvimento de estratégias, ou uma conscientização dos recursos em LM, permitiria o leitor superar suas lacunas lingüísticas.

O termo *estratégia* tem sido definido e entendido de maneiras diferentes em áreas como Psicologia, Economia, Política, Lingüística e LA. Essa falta de acordo terminológico evidencia-se até mesmo nas várias sub-áreas da LA, incluindo a leitura (vide Moura 1992; Schmitz no prelo; Stevick 1990, para discussões sobre terminologia). Nesse sentido, ela tem-se referido ao componente cognitivo, ou mais especificamente, a procedimentos, processos, ações, comportamentos, adquirindo, principalmente, a conotação de procedimentos superficiais para solução de problemas.

No ensino de leitura, as estratégias parecem incluir ou se confundir com as chamadas habilidades de leitura (*reading skills* ou *sub-skills*), incluindo a inferência de palavras através do contexto, o reconhecimento de cognatos e famílias de palavras, a leitura rápida para construir a idéia central do texto (*skimming*), leitura rápida para localização de informações específicas (*scanning*), predição ou levantamento de hipótese, a identificação e uso de palavras de referência, a separação entre idéia principal e detalhes de apoio, a ativação e uso de conhecimento prévio apropriado, o reconhecimento da estrutura do texto (Carrell 1989), dentre outras.

Para Barnett (1988:150), estratégias de leitura são "... *operações mentais envolvidas quando leitores abordam o texto efetivamente e fazem sentido do que lêem*" (26c). Observa-se, a partir dessa definição, que a autora apenas refere-se à estratégias enquanto procedimentos produtivos. Entretanto, essa não é uma definição de consenso. Para alguns autores, estratégias são abordagens para executar uma tarefa de aprendizagem, seja ela produtiva ou improdutiva. Entretanto, quando o foco é o ensino, o termo estratégia parece referir-se apenas àquelas consideradas

produtivas.

A pesquisa sobre estratégias de leitura é bastante vasta e sua revisão foge do escopo deste trabalho (vide Olshavsky 1976-1977; Baker & Brown 1984; Oxford & Crookall 1989; Kern 1989; Ramos 1988, dentre outros). No início, os estudos sobre estratégias usavam medidas indiretas, tais como observação da recitação oral, análise de lapsos orais (*miscue analysis*) e até mesmo de resultados de testes, ou medidas de produto. Refiro-me a essas medidas como indiretas porque, a partir delas, o pesquisador tentava inferir os procedimentos usados no processo de leitura. Entretanto, com o advento de técnicas mentalistas, por exemplo, as medidas passaram a ser mais diretas. A leitura oral, nesse caso, não é considerada apenas uma recitação, mas uma externalização de um processo que pode ser subvocalizado (Cohen 1987). Muitos estudos foram conduzidos usando-se essas técnicas, tanto em LM (Olshavsky op. cit.; Baker & Brown op. cit., dentre outros) como em LE (Hosenfeld 1984).

O foco dessa pesquisa sobre estratégias, usando-se técnicas mentalistas era, portanto, uma tentativa de identificação de estratégias "eficientes" ou "ineficientes" de forma que se pudesse ensinar os leitores a usar aquelas eficientes. Entretanto, pode-se dizer que essa abordagem simplifica de certo modo a questão, uma vez que as estratégias não são inherentemente eficientes ou ineficientes, mas que

(...) a seleção de uma estratégia, sob um certo número de circunstâncias (tais como o texto, a situação de leitura e a interação com outras estratégias) podem criar uma situação de fracasso na leitura mesmo, quando estratégias consideradas eficientes estão em uso (Cohen op. cit.:60) (27c).

Além disso, nesses estudos, o número de sujeitos é geralmente pequeno, e dada a natureza individualizada da aplicação que os leitores fazem das estratégias, torna-se difícil a generalização de seus resultados. Essa pesquisa é, ainda, limitada ao conhecimento de estratégia (cognição), deixando de investigar a consciência, ou o entendimento que os leitores têm das estratégias (metacognição). Cognição implica em "possuir" as habilidades; metacognição refere-se à consciência e controle consciente sobre essas habilidades. Esta é uma distinção semelhante, senão a mesma, feita por Brown (1980) entre *saber, saber como saber e saber sobre o saber*.

Essa limitação também é observada no ensino de leitura instrumental no Brasil, seguindo a tendência geral da literatura internacional, como afirmei anteriormente. Baseada em observações desses cursos, gostaria de fazer minhas as palavras de Tavares (1993:33) quando afirma que, em geral, as estratégias são ensinadas

(...) como elementos isolados (...) O aluno passa a conhecer estratégias (...) mas não é ensinado a combiná-las e integrá-las ao seu processo de leitura. Por vezes, não se esclarece para o aluno as relações entre

os diversos tipos de problemas de compreensão que ele pode ter e os tipos de estratégias apropriadas à possível solução destes problemas (...) [Assim] o foco destes cursos parece ser mais o conhecimento de estratégias de leitura do que o conhecimento de como, quando e por quê usá-las, que envolve metacognição.

Com relação à consciência metacognitiva das estratégias, mais recentemente, dois aspectos são reconhecidos, ou seja, uma consciência da cognição e controle da cognição (Baker & Brown op. cit.; Flavell 1981). O primeiro deles, ou conhecimento da cognição, é definido como o conhecimento que o leitor tem de seus recursos cognitivos e a compatibilidade entre o leitor e a situação de leitura. De acordo com Carrell (1989:122),

(...) se um leitor está consciente do que necessário para um desempenho efetivo, então é possível tomar as providências necessárias para conseguir suprir as exigências da situação de leitura de maneira mais efetiva. Se, no entanto, os leitor não está consciente de suas próprias limitações como leitor, ou da complexidade da tarefa envolvida, então, dificilmente, poder-se-ão esperar ações preventivas para antecipar os problemas ou recuperar-se deles (28c).

Relacionado a esse aspecto de metacognição está a concepção que o leitor faz do seu processo de leitura. Estudos com crianças aprendendo a ler em inglês como LM (Harste & Burke 1977; Rhodes 1979; Delawater 1975; Dank 1976) oferecem evidências para se afirmar que os leitores têm modelos internalizados do processo de leitura, que, nesse caso, são vistos como "... um sistema de pressuposições através dos quais o leitor irá processar informação textual disponível [ou] ... schemata que governa o comportamento" (Harste & Burke op. cit.:32) (29c).

Com relação a esses modelos, essa pesquisa consegue identificar pelo menos três tendências: leitores centrados no significado (*meaning-centered readers*), leitores centrados na palavra (*word-centered readers*) e leitores centrados no som (*sound-centered readers*).

O estudo conduzido por Devine (1988) apresenta resultados nesse sentido. Seguindo a linha da pesquisa em LM, que mostra a tendência de leitores menos proficientes em focalizar a leitura como um processo de decodificação, ao invés de considerá-lo um processo de construção do significado, esse estudo tem por objetivo investigar os modelos de leitores de L2 e sua influência no comportamento apresentado durante a leitura. Para isso, a autora conduz um estudo de caso com dois leitores que, embora apresentando orientações diferentes (isto é, um é centrado no significado e o outro no som), mostram o mesmo nível de proficiência (iniciantes) em L2 (inglês). Sua conclusão é que leitores centrados no significado, ou que apresentam estratégias *globais* (*top-down*), demonstram compreensão boa ou excelente em uma tarefa de recontar, feita depois da leitura oral, enquanto leitores centrados no som, ou na palavra, ou que apresentam estratégias *locais* (*bottom-up*),

apresentam compreensão deficiente ou bastante deficiente. Entretanto, uma outra contribuição do estudo de Devine (op. cit.:130), mais importante para esta revisão, é que

(...) primeiro, as orientações teóricas que o leitor tem de sua leitura podem determinar, de certa maneira, o grau com que uma proficiência lingüística limitada restringe sua competência de leitura; e, segundo, que os modelos que os leitores têm podem ser de importância crítica, uma vez que lhes permite conseguir um equilíbrio entre um processamento ascendente ou guiado pelos dados do texto, e um processamento descendente, guiado pelo conhecimento, necessários para a interpretação de um texto (30c).

Entretanto, não se pode deixar de salientar que este estudo também tem suas limitações, relacionadas principalmente à natureza complexa da leitura e deficiências na proficiência em L2. Devine (op. cit.:130) reconhece que

É geralmente difícil para os leitores (e aprendizes de língua em geral) articular modelos de processos mentais; essas operações mentais podem, simplesmente, não estarem disponíveis para a introspecção. Para leitores em L2, como estes deste estudo, a dificuldade fica acrescida pela falta de fluência na L2. Além disso, como se observa freqüentemente, processos cognitivos complexos, muito provavelmente, envolvem um número de estratégias de interação. Leitores, por exemplo, podem utilizar uma estratégia para a leitura de um romance, uma outra para a leitura de material científico ou técnico, e uma combinação de estratégias para a leitura de um jornal. Propor um único modelo, mesmo que complexo, pode subestimar essa complexidade (31c).

Uma outro aspecto dessa pesquisa em estratégias, embora em LM, é a falta de correspondência entre o que os sujeitos "dizem fazer" e o que "realmente fazem". Essa constatação levou alguns autores a concluir que não se deve acreditar no que os alunos/leitores "afirmam fazer" durante a leitura. O leitor, muitas vezes, indica que a estratégia é eficiente mas não a usa. Paradoxalmente, no entanto, também é possível o leitor não indicar ou salientar uma estratégia como importante, mas usá-la em sua leitura.

Na literatura sobre leitura, costuma-se distinguir dois tipos de estratégias, ou seja, *cognitivas* e *metacognitivas*, seguindo-se a segmentação entre cognição e metacognição. Estratégias cognitivas são aquelas automáticas, inconscientes, que possibilitam a leitura rápida e eficiente. Estratégias metacognitivas, por outro lado, referem-se a procedimentos conscientes, que permitem a desautomatização e controle das estratégias cognitivas para auto-regulamento da compreensão. Kleiman (1985) cita como exemplos de estratégias cognitivas as diversas estratégias de segmentação sintática e de recuperação anafórica (Kato 1984). Especificamente com relação ao léxico, também

são estratégias cognitivas o reconhecimento global das palavras e de pareamento de elementos cognatos entre duas línguas. Como exemplos de estratégias metacognitivas, Kleiman (op. cit.) cita os esforços conscientes do leitor para resolver equívocos ou inconsistências detetadas após o processamento automático, ou após uma interpretação que se mostra inconsistente com o material em processamento. Saber que tipo de estratégias usar, ou as ações a tomar, para se construir um significado do texto, e quando usar essas estratégias, para conseguir vários objetivos, são atividades metacognitivas. A utilização do contexto para inferir palavras seria também desse tipo. Em suma, poderiam ser definidas como

(...) comportamentos usados para centrar, ordenar, planejar e avaliar a aprendizagem. Essas estratégias, que vão 'além do cognitivo' são usadas para fornecer 'controle executivo' sobre o processo de aprendizagem (Oxford & Crookall 1989:405) (32c).

Entretanto, pode-se observar, na literatura, que um processamento automático, inconsciente e fluente (estratégia cognitiva) não tem sido, na maior parte das vezes, caracterizado pelo uso de estratégias. Geralmente estas têm sido associadas apenas às situações de problema, quando há pausas no processamento (estratégias metacognitivas). A definição de Cohen (1987:61) ilustra essa tendência. Para ele, estratégias de leitura

(...) referem-se aos processos mentais que os leitores escolhem usar, conscientemente, em suas tarefas de leitura. Tais estratégias podem contribuir para a compreensão bem sucedida ou desviar-se dela. Em princípio, o que distingue as estratégias de outros processos é o elemento de escolha envolvido em sua seleção. Na realidade, a atenção que os leitores dão à sua escolha de estratégias pode ser colocada em um contínuo, que compreende desde atenção total até falta de atenção. Se os leitores tiverem que indicar quais as estratégias usadas, é provável que saibam descrever até aquelas que usam pouco, uma vez que estas estão, por definição, dentro do domínio da consciência. Eles não seriam capazes, entretanto, de descrever certos processos de leitura inconscientes (33c).

Essa visão mais estreita de estratégia, apenas enquanto procedimentos conscientes em situações de problema ou procedimentos compensatórios é, a meu ver, problemática, uma vez que parece sugerir que o conhecimento, seja ele lingüístico ou esquemático, dispensa um componente de uso, ou que esse componente é algo separado, apenas acionado quando há rupturas no procedimento automático. Essa visão parece estimular ainda mais, portanto, no ensino, uma tensão ou competição sobre a importância de um componente sobre o outro.

As estratégias metacognitivas costumam ser divididas em dois grupos, que variam de um

autor para outro. Barnett (1988) as classifica em *baseadas no texto* e *baseadas na palavra*. *Baseadas no texto* são aquelas que usam o texto como um todo, ou grande parte ou partes dele. Elas incluem o uso de conhecimento prévio, predição, uso do título e ilustrações, *skimming*, *scanning*, dentre outras. Elas relacionam palavra, ações e idéias, ou usam o contexto como um todo. Block (1986) prefere denominá-las estratégias de *compreensão geral*, em oposição àquelas de *lingüística local*; Barnard (1980) refere-se a elas como estratégias *local* e *global*, ou ainda Olshavsky (1976) como *relacionadas à oração (clause-related)* e *relacionadas à palavra (word-related)*.

Associações são comumente feitas entre estratégias do tipo *local* e leitores menos proficientes, e entre estratégias do tipo *global* e leitores mais proficientes. Um uso fluente e flexível de estratégias parece caracterizar o bom leitor, enquanto um uso rígido e inapropriado tende a caracterizar maus leitores (Kern 1989).

Apesar da tendência da literatura em conceber o ensino de estratégias como mais eficiente do que o ensino centrado na construção de uma base lingüística, apresentando várias listas de estratégias consideradas produtivas ou eficientes (Hosenfeld 1977b), Barnett (op. cit.) salientava que até 1988, não havia muitos estudos que realmente avaliassem sua eficiência. Hamp-Lyons (1985) mostra que o ensino de estratégias pode ter um efeito positivo na compreensão de alunos intermediários. Os trabalhos de Barnett (1988), Kern (op. cit.), Carrell (1989) são também tentativas nesse sentido.

O trabalho de Barnett (1988), focalizando leitores maduros de nível universitário de francês como LE, investiga as relações entre compreensão, uso de estratégias e percepção do uso de estratégias. Seus resultados indicam uma correlação significativa entre os três aspectos. Dessa forma, não apenas uso de estratégia (lembrar o contexto durante a leitura, por exemplo) mas, também, a consciência sobre o uso de tais estratégias demonstrou ser importante. Os alunos que efetivamente consideram e lembram-se do contexto à medida que lêem ou que usam essa estratégia, apresentam uma melhor compreensão quando comparados com alunos que usam essa estratégia menos freqüentemente ou não tão bem. Da mesma forma, os alunos que afirmam usar essas estratégias consideradas mais produtivas (percepção de uso) realmente lêem melhor através do contexto do que alunos que não têm consciência de uso dessas estratégias.

Para Carrell (op. cit.), entretanto, uma grande limitação do estudo de Barnett está no questionário usado para eliciar as percepções dos sujeitos sobre uso das estratégias. Ele é um questionário de múltipla escolha, desenvolvido e corrigido em termos de respostas corretas pré-determinadas, isto é, julgamentos pré-determinados pelo pesquisador sobre a eficiência de diferentes estratégias de leitura. Além disso, a autora combina em um único item várias estratégias diferentes, forçando o leitor a escolher apenas uma, quando na realidade, várias respostas podem estar corretas.

O estudo de Carrell é uma tentativa de investigar, como o estudo anterior, não apenas uso de estratégias mas também consciência do uso, através de um desenho experimental de pesquisa e

um número maior de sujeitos. Sua pesquisa focaliza, portanto, a consciência metacognitiva de leitores em L2 sobre estratégias de leitura, ou resultado de sub-habilidades, assim como sua utilidade em LM e L2, além da relação entre essa consciência metacognitiva e sua compreensão de textos em LM e L2. Seu estudo envolve dois grupos de sujeitos: 45 falantes nativos de espanhol estudando inglês como L2 (grupo de proficiência mais baixa) e 75 falantes nativos de inglês aprendendo espanhol como LE (grupo mais avançado).

Os resultados com relação à L2 mostram que o grupo mais avançado tende a ser mais *global* com relação às suas percepções de estratégias eficientes e que causam dificuldade, enquanto o grupo de proficiência inferior é mais dependente de habilidades de decodificação ascendente, que Carrell reconhece como reflexos metacognitivos do *efeito de curto-circuito* ou do limiar de proficiência lingüística (*language proficiency ceiling*) de Clark (1988).

Para esse estudo, foi desenvolvido um questionário de 44 itens em duas versões, ou seja, inglês e traduzido para o espanhol (vide Apêndice A, Seção A.5). Ele se subdivide em duas partes: parte A, introdutória, de identificação do sujeito e parte B, de fatores metacognitivos propriamente dita. A parte B do questionário consiste em 36 afirmações sobre leitura silenciosa em inglês. As respostas são dadas através de uma escala Likert de 1 a 5, onde 1 indica *acordo total* com a afirmação e 5 *desacordo total*. Os itens são subdivididos em quatro grupos ou categorias de metacognição, conforme mostra a Tabela 2.1 abaixo. A análise dos dados foi conduzida em cada um dos quatro grupos ou categorias, usando-se regressões simples. Em alguns casos, quanto mais o sujeito concordava com a afirmação metacognitiva, melhor sua proficiência em leitura (relação positiva). Em outros casos, quanto mais o sujeito discordava da afirmação metacognitiva, melhor sua proficiência em leitura (relação negativa).

Uma grande vantagem do questionário de Carrell (op. cit.) é não prejudicar a eficiência das estratégias, uma vez que deixa esse julgamento para os sujeitos. Em um segundo momento investiga a relação entre esses julgamentos e a eficiência em leitura. Além disso, o questionário elicit os julgamentos dos leitores sobre suas estratégias de leitura sem forçá-los a escolher entre várias estratégias diferentes. Entretanto, uma limitação ou uma palavra de alerta na interpretação dos resultados desse estudo, assim como dos outros acima apresentados, está em considerar um modelo único de leitura para situações diversas, que não leva em conta, por exemplo, a leitura com propósitos diferentes ou textos de dificuldade variada, limitação essa também levantada por Devine (op. cit.) quando discute modelos internalizados do leitor.

O estudo de Kern (op. cit.) tem, por objetivo, verificar possíveis diferenças no desempenho dos sujeitos resultante do ensino de estratégias de inferência lexical. Nesse caso, da mesma forma que no estudo de Hamp-Lyons (op. cit.) acima citado, os sujeitos são todos de nível intermediário, embora esse autor subdivida o grupo em três níveis de proficiência em leitura. Seus resultados mostram que embora os alunos mais fracos fossem os que mais vantagem tiraram desse ensino, esse

ganho não foi significativo. O alunos que denomina *iniciantes*, na realidade, apenas o são com relação ao grupo. O fato de serem de nível intermediário parece fazer uma grande diferença. Investigações que consideram sujeitos com outros níveis de proficiência seriam importantes para mostrar a eficiência de um ensino centrado em estratégias. Uma outra observação que se pode fazer a respeito do estudo de Kern é que seus sujeitos, além de estarem expostos a um ensino com foco em estratégias, também estavam expostos à questões linguísticas tais como gramática, habilidades de produção de texto, dentre outras. Um outro aspecto que deveria ter sido investigado, também nesse caso, seria a proficiência de leitura em LM dos leitores, uma vez que o autor explica a diferença de desempenho dos grupos médio e forte como sendo devida à transferência de proficiência em leitura em LM, sem entretanto, apresentar evidências para essa explicação.

Tabela 2.1 – Questionário metacognitivo (Carrell 1989)

Categorias	Número de estratégias
1. Confiança	6 (itens 1-6)
2. Correção ou reparo	5 (itens 7-11)
3. Eficiência	17 (itens 12-20; 29-36)
som/letra	3
palavra/significado	5
idéia central	2
conhecimento prévio	2
detalhes de conteúdo	2
organização do texto	2
sintaxe da oração	1
4. Dificuldade	8 (itens 21-28)
som/letra	2
palavra/significado	1
idéia central do texto	1
conhecimento prévio	1
organização do texto	1
sintaxe da oração	1

Tendo apresentado uma revisão de estudos que avaliam o papel do componente linguístico e cognitivo na leitura, focalizo, na próxima seção, a relação de um aspecto desse componente linguístico, ou seja, do vocabulário, com a compreensão em leitura.

2.4 – O componente lexical

Uma divisão comum, feita pela lingüística em geral, e adotada no ensino de línguas e leitura/ensino de leitura, é entre o vocabulário (ou o aspecto lexical/semântico) e a gramática (sintaxe, morfologia). A mesma tensão que se observa com relação ao componente lingüístico e cognitivo pode ser identificada, também com referência a esses dois aspectos do componente lingüístico, que têm sido abordados como componentes distintos e seus papéis avaliados separadamente (vide Macnamara 1967; Ulijn 1978; Ulijn & Kempem 1976; Cooper 1984, dentre outros). Muitas vezes uma contribuição maior do vocabulário tem sido sugerida. Outras vezes a da sintaxe. Também não faltam autores (Barnett 1986; Rumelhart 1977) salientando a necessidade dos dois tipos de habilidades ou níveis de análise, dada a interação existente entre eles.

Antes de apresentar as evidências empíricas dessa literatura, entretanto, considero importante fazer uma retrospectiva da importância do vocabulário na LA em geral e especificamente, na área de aquisição/ensino de línguas.

2.4.1 – O vocabulário: uma perspectiva histórica

Esta seção apresenta o vocabulário sob uma perspectiva histórica e o situa dentro da área mais geral da aquisição da linguagem. Seu objetivo, e também da próxima seção, que aborda aspectos da pesquisa sobre o ensino/aquisição do vocabulário, é servir de pano de fundo para as questões mais específicas que envolvem o vocabulário e a leitura, foco deste trabalho, apresentadas na Seção 2.4.3.2 deste capítulo.

Grande parte dos artigos sobre vocabulário (vide Meara 1980, para uma revisão da literatura sobre aquisição de vocabulário), escritos por volta dos anos 80, invariavelmente, iniciam-se com a mesma observação, ou seja, que o vocabulário é o aspecto do ensino/aprendizagem de L2 mais negligenciado, não só pelos recentes desenvolvimentos na pesquisa, como pela LA em geral, especialmente durante 1960-1980 (ou de 1945-1970, para outros autores). Essa negligência, que Levenston (1979) considera uma discriminação, causa ainda mais estranheza, uma vez que, na opinião quase unânime dos próprios aprendizes, o vocabulário é o aspecto que mais dificulta essa aquisição, principalmente depois de ultrapassados seus estágios iniciais (Meara op. cit.; Laufer 1986; Carter 1987b; Carter & McCarthy 1988).

Na opinião de Laufer (op. cit.), é muito difícil justificar essa negligência, porque é senso comum que sem vocabulário adequado não se consegue competência ou desempenho adequado em uma língua. Alguns estudos (Meara 1984; Politzer 1978) até mesmo mostram que erros lexicais são mais sérios, e causam mais rupturas, do que erros fonológicos ou gramaticais. Especulando sobre as

possíveis causas do que ela denomina o *paradoxo da negligência*, a autora salienta que a aquisição de linguagem tem sido influenciada pelas teorias lingüísticas e pelo interesse de metodólogos. Uma análise das visões dominantes, em cada dessas áreas, mostra que as condições realmente não foram propícias para que o vocabulário recebesse a atenção merecida. Os lingüistas

(...) preferiram estudar a gramática e a fonologia porque são sistemas fechados e, portanto, se prestam muito mais à abstração e generalização do que o vocabulário, que não é um sistema fechado mas um conjunto aberto ... e como uma boa teoria é aquela que dá conta do maior número de eventos da maneira mais simples possível, significa que a teoria da gramática é mais poderosa do que a teoria do léxico (Laufer op. cit.:70) (34c).

Levenston (1979:148), por sua vez, salienta que os psicolingüistas

(...) têm reagido contra uma tendência anterior de lingüistas (orientados pela teoria da aprendizagem) de se concentrar na aprendizagem de vocabulário, vista como uma aprendizagem de associações, ao invés da aquisição da gramática, que necessita de regras (35c).

Marton (1977:34), ao referir-se aos metodólogos, afirma que

Com relação à interação entre metodólogos e pesquisadores de aquisição da linguagem, até recentemente, o interesse principal de ambos parecia estar nos estágios iniciais da aprendizagem da língua. Acreditava-se que, nesses estágios, era mais importante concentrar-se na gramática (36c).

Com exceção do trabalho conduzido pelo chamado *Movimento de controle do vocabulário* (vide próxima seção), que se iniciou em 1930, e que tinha por meta, a seleção de um vocabulário básico a ser usado nos estágios iniciais da aprendizagem (26), a visão dominante nos anos 40, 50 e 60, então influenciada pelos lingüistas estruturalistas americanos, era relegar o vocabulário a um segundo plano. O maior problema na aprendizagem de uma língua são seus sons e estrutura gramatical, salientava Fries (1945). Dar ao vocabulário um espaço próprio dentro da sala de aula poderia sugerir aos aprendizes a falsa idéia de que a aprendizagem de uma língua resume-se em aprender palavras. A afirmação de Gleason (1961:7) mostra claramente essa tendência, parecendo sugerir a falta de "conteúdo" ou "expressão" no vocabulário:

Na aprendizagem de uma segunda língua, você vai achar o vocabulário comparativamente fácil, apesar de ser o que os alunos mais temem. A

parte mais difícil é apreender o conteúdo e expressão de novas estruturas (37c).

Como a análise contrastiva era, nessa época, uma tendência dominante, havia também a preocupação em se comparar o vocabulário da língua alvo com o vocabulário da LM. O estruturalismo e a análise contrastiva, juntamente com a psicologia comportamental, deram origem ao método áudio-lingual, que seguia a visão acima apresentada, ou seja, muita estrutura e pouco vocabulário.

O advento da lingüística transformacional de Chomsky, nos anos 60, apenas serviu para reforçar a idéia de que *"... o vocabulário era de certa forma periférico, uma irritante irregularidade em uma gramática ordenada"* (Carter & McCarthy op. cit.:41) (38c).

Foi somente no início dos anos 70, que essa visão começou a ser desafiada. Bright & McGregor (1970), juntamente com Wilkins (1972) foram alguns dos primeiros a manifestarem descontentamento com o papel do vocabulário no ensino. A opinião de Wilkins (op. cit.:111) é que se *"... sem gramática muito pouco se pode comunicar, sem vocabulário nada se pode comunicar"* (39c).

Convém lembrar que até essa época, as preocupações, no que se referem ao ensino de vocabulário, concentravam-se na seleção de itens baseados no critério de frequência. Começou-se a sentir a necessidade de se ensinar estratégias para inferência, dada a constatação da impossibilidade de se ensinar todas as palavras necessárias. Assim, inicia-se uma nova visão de ensino de vocabulário, isto é, não mais como listas de palavras, mas como uma habilidade lingüística, de responsabilidade do aprendiz. Lord (1974), seguindo a visão de Vygotsky (1978), mostra que *"... a relação entre pensamento e palavras é dinâmica: é um processo, não uma coisa"* (Carter & McCarthy op. cit.:43) (40c), representando, portanto, uma mudança de uma abordagem que enfatizava um controle ou seleção de vocabulário, para uma outra, que se concentrava no significado. Uma visão não substitui a outra, no entanto, mas passam a se complementar.

Com o desenvolvimento do movimento nocional-funcional, comunicativo e de propósitos específicos, a preocupação com o léxico passou a ser vista como resultado das necessidades do aprendiz. Para Carter (1987b), o vocabulário parece recuperar, a partir desse momento, seu lugar de prestígio no ensino de línguas. Um ensino centrado no aprendiz passa a considerar o vocabulário enquanto recurso estratégico para suas necessidades, dentro de seus objetivos comunicativos.

Laufer (op. cit.:71) reconhece o início de uma mudança para uma atitude mais positiva frente ao vocabulário e uma atenção maior para questões lexicais nas pesquisas de aquisição. As causas dessa mudança, segundo a autora, estariam nos resultados de pesquisa em semântica, numa visão mais equilibrada da aprendizagem da linguagem e no interesse em aprendizes mais avançados. Para a autora,

(...) o ímpeto maior a uma pesquisa futura sobre a aquisição de vocabulário, provavelmente, se origina de alguns princípios do ensino comunicativo. Não pode haver fluência sem uma base sólida de vocabulário; nenhum insumo abrangente e relevante com um léxico pobre (...) A abordagem de ensino orientada para o aprendiz e o trabalho sobre motivação deram origem à análise das necessidades do aprendiz tanto pelo professor como pelo próprio aprendiz. Tal análise, necessariamente, mostra a necessidade de uma melhor aprendizagem de vocabulário (41c).

Carter (op. cit.) também salienta o aumento de interesse no ensino de vocabulário desde o final dos anos 70, especialmente na Grã-Bretanha, onde projetos de pesquisa sobre o léxico já eram desenvolvidos no final de 1960 (Sinclair et alii 1970), principalmente sobre semântica lexical (Lyons 1968, 1977; Leech 1974), e também onde raízes mais profundas sempre existiram no movimento de controle de vocabulário e lexicografia em EFL (Cowie 1984).

A década de 80 mostra um interesse crescente pelo vocabulário, confirmado pelos inúmeros trabalhos sobre o assunto. Bastante importante e ilustrativa das tendências dos anos 80 foi a publicação de uma coletânea de artigos do *Regional English Language Centre (RELC)* de Singapura, intitulado *Guidelines for Vocabulary Teaching*. Os tópicos abordados mostram uma combinação abordagens e de interesses variados, tais como colocações, procedimentos *cloze*, contexto, dentre outros. Além dessas contribuições, também merece ser destacado um artigo-levantamento (*survey article*) de Meara (1980) assim como sua abrangente bibliografia anotada sobre ensino de vocabulário. A preocupação crescente com o que é aprender vocabulário e não mais apenas com o que aprender coloca questões tais como no que consiste uma competência lexical, base necessária para se estabelecer objetivos para o ensino e avaliação do vocabulário

O engajamento do aprendiz no seu processo de aprendizagem também é a base do conhecido livro de Gairns & Redman (1986) para o ensino de vocabulário. Para aprendizes que já têm um conhecimento básico, Chanell (1981) sugere maneiras de incorporar intravisiões da *teoria de campos e análise componencial*, adaptando grades e escalas para ilustrar as características semânticas e suas colocações. Além disso, essa visão mostra uma tendência a individualização e auto-gerenciamento do processo de aprendizagem (Rivers 1983). Também importantes foram as publicações de Wallace (1982) e Allen (1983). Na opinião do primeiro, o vocabulário deve ser ensinado. A tarefa envolvida em saber uma palavra é, nesse caso, baseada num conceito rico de vocabulário que envolve vários aspectos: "... saber uma palavra de forma que possa ser reconhecida, lembrada quando desejada, relacionada a um objeto ou um conceito, corretamente usada, pronunciada e soletrada, "colocada" adequadamente, usada no nível certo de formalidade e com consciência de suas conotações e associações" (Wallace op. cit.:27) (42c). Esse conceito, baseado em Richards (1976), amplia o escopo do vocabulário em LA, podendo ser considerado uma

base promissora para o desenvolvimento do ensino/avaliação/pesquisa de vocabulário, como mostro na Seção 2.4.3.3.

Continua nessa época o interesse pelo vocabulário generalizável e estratégico, para usos geral e acadêmico. Essa pesquisa inclui o trabalho de Carter (1982, 1986, 1987a) em vocabulário básico (*core*) e a pesquisa em ESP de Hutchinson and Waters (1981). Para esses autores, a terminologia técnica não é o material-insumo mais útil para alunos de ESP. Mais necessário para esses alunos "*... não é um corpo de linguagem técnica mas a habilidade de mobilizar os recursos do inglês geral na solução de problemas técnicos. O vocabulário procedimental (procedural) ou mais geral ajuda o aprendiz chegar até o vocabulário específico ou técnico* (Carter & McCarthy 1988:50) (43c).

A partir do final dos anos 80 e 90, as tendências da pesquisa parecem concentrar-se nos seguintes aspectos:

- Qualidade da aprendizagem de vocabulário, em especial redes associativas e campos semânticos, como técnicas de expansão de vocabulário (Crow 1986; Crow e Quiley 1985);
- Relações com a leitura e compreensão, principalmente as relações entre expansão de vocabulário e aumento na compreensão. Essa questão envolve aspectos relacionados à natureza e extensão do vocabulário, principalmente para a leitura de textos não simplificados; e
- Qualidade do conhecimento de vocabulário.

Complementando esta seção, focalizo, a seguir, o vocabulário sob o ponto de vista da pesquisa sobre ensino e aquisição.

2.4.2 – O vocabulário: a pesquisa sobre ensino e aquisição

Meara (1980), resumindo a pesquisa na área aquisição de vocabulário até os anos 80, mostra que, apesar da tendência de mudança com relação ao vocabulário, que se fez sentir nos últimos anos dessa década (Meara 1983; Twomey 1979; Dale & Razik 1963), mencionada na seção anterior,

(...) a pesquisa (...) não tem sido teórica e sistemática (...) não há teorias claras de aquisição de vocabulário (...) e, de uma maneira geral (...) tem evitado as questões teóricas sérias que surgem uma vez que se distancia desse nível bem básico (...) Os trabalhos são em escala pequena, e nenhum deles, até agora, tem contribuído muito para a construção de um corpo de conhecimento coerente e substancial (Meara op. cit.: 221-222) (44c).

Grande parte desses estudos têm sido na área de ensino de vocabulário, e não na

aprendizagem de vocabulário, o que causa estranheza, uma vez que o ensino dever-se-ia basear no conhecimento que se tem sobre a aprendizagem. Para Meara (1982:29),

O ensino efetivo do vocabulário, especialmente nos estágios iniciais, é, em um grau considerável, dependente de respostas satisfatórias para a seguinte questão: como se caracteriza o léxico mental de um falante de L2 e como difere do de um falante nativo monolíngüe (45c).

As áreas majoritárias dessa pesquisa são seleção e controle de vocabulário (*extensão do vocabulário (vocabulary size) e contagem de frequência (frequency counts)*) (vide Wilkins 1972; Bright & McGregor 1970; Mackey 1965, para uma revisão crítica dessa pesquisa). O trabalho mais bem desenvolvido na área de controle de vocabulário têm sido "...tentativas de justificar a seleção de itens de vocabulário para inclusão em cursos e exames com base na contagem de frequência e medidas objetivas similares"(Meara 1980:222-223) (46c).

Essas contagens de frequência (General Service List ou West 1953; Thorndike-Lorge List (1944); Kucera & Francis 1967, dentre outras) entretanto, fora seu uso em testes de vocabulário, foi muitas vezes criticada, pois serviu de base para as conhecidas fórmulas de legibilidade, que acabou orientando a simplificação textual, levando, como salienta Kato (1985), a uma "imbecilização" dos textos escolares nos EUA (vide Bruce & Hubin 1988, para uma revisão crítica dessa pesquisa).

Uma outra área dentro de aquisição de vocabulário que merece ser salientada é a do uso de técnicas mnemônicas para o ensino. Dentre os vários métodos propostos, principalmente nos Estados Unidos, destaca-se o da *palavra chave (key word)*. Esse método consiste, num primeiro estágio do aprendizado, na apresentação de palavras na LM, denominadas palavras chaves, que são foneticamente similares às palavras na língua alvo; num segundo estágio, essas palavras são associadas com a tradução em LM, da palavra original na língua-alvo, através de uma imagem visual que chama a atenção (Meara 1980; vide também Atkinson & Raugh 1975; Atkinson 1975; Raugh & Atkinson 1974; Pressley 1977).

Apesar dos resultados surpreendentes à primeira vista, desafiando a crença acima salientada, isto é, a necessidade de uma seleção de itens a serem aprendidos, dada a impossibilidade de se aprender um vocabulário mais efetivo, eles devem, como salienta Meara (op. cit.), ser interpretados com um certo cuidado. Muitos problemas parecem estar envolvidos, sendo o mais sério, no meu ponto de vista, o fato de o método ser baseado em um modelo de aprendizagem que considera a *associação de pares entre duas línguas (paired associate learning)* (Crothers & Suppes 1967; Kellog & Howe 1971; Nation 1983) que não se mostra suficientemente rico para dar conta da complexidade envolvida na aquisição do vocabulário. Nesse sentido, concordo com Meara (op. cit.:225) quando afirma que

Aprender vocabulário não é apenas uma questão de adquirir formas equivalentes em tradução (...) e ignorar completamente os padrões complexos de relações de significado que caracterizam um léxico totalmente formado, em oposição a uma mera lista (...) alguma estruturação precisa ocorrer, mesmo quando a menor lista de palavras é aprendida (47c).

Apesar de suas contribuições, as duas áreas de pesquisa, ou seja, controle do vocabulário e técnicas de memorização, ainda apresentam muitas limitações. A principal delas é, ao mostrar uma preocupação excessiva com questões relativas à seleção (quais palavras merecem ser ensinadas) e apresentação (qual seria a melhor forma de apresentação), ou seja, uma concentração apenas nos aspectos periféricos ou de gerenciamento do processo de aprendizagem, deixar de focalizar os aspectos centrais da aquisição de vocabulário, isto é, como as palavras são aprendidas, ou as diferenças entre memorização de curto e longo prazo, ou ainda entre recuperação de um significado e procedimentos de memorização adotados, e entre retenção de significados e o aprendizado da palavra para uso produtivo e receptivo (Meara op. cit.), seja na leitura ou na escrita.

Uma outra crítica que se pode fazer a essa pesquisa é que, embora ela reconheça as dificuldades envolvidas na aprendizagem de um vocabulário razoável em LE ou L2 num curto período de tempo, não apresenta nenhuma evidência teórica clara que justifique uma limitação no número de palavras a serem aprendidas pelo falante não-nativo. Além disso, apesar de alguns estudos isolados em L2, e do interesse ou relevância para uma segunda língua, eles têm-se limitado a investigações em LM.

Além dessas duas áreas, a pesquisa em ensino e aquisição de vocabulário também salienta um trabalho experimental para investigar como bilíngües retêm as palavras em seus dicionários mentais, e como esse léxico ou léxicos estão estruturados. O objetivo desse trabalho está em oferecer modelos, embora ainda incompletos, que sirvam de parâmetros, para se julgar o nível do comportamento esperado de falantes de uma LE menos fluentes. Uma de suas evidências fortes é a interferência que existe em condições de interlíngua, sempre da língua mais fraca para a mais forte, e ela está principalmente relacionada à forma fonética da palavra, apoiando a noção de que os dois léxicos do bilíngüe são, na realidade, apenas um.

Nessa área de pesquisa sobre o léxico dos bilíngües, o método que mais se destaca, por ser mais acessível e mais facilmente entendido, parece ser o uso de associações. Ele consiste em apresentar um número de palavras aos sujeitos participantes, que são solicitados a responder com a primeira associação que lhes vem à mente. Os sujeitos desse tipo de experimento, em geral, tendem a apresentar os mesmos padrões de resposta (*estereótipo de associação* ou *associational stereotypy*), que se classificam em duas categorias: *paradigmáticas* (respostas pertencem à mesma grande classe de forma da palavra original, ou seja, para a palavra *cachorro*, as associações seriam *gato*, *osso*, *rabo*, por exemplo) ou *sintagmáticas* (respostas combinam com a palavra estímulo para formar uma

frase ou um conjunto sintático). No caso da palavra cachorro, as associações seriam *morde, late*, dentre outros). Os resultados dessa pesquisa indicam:

- Respostas na língua mais fraca tendem a ser menos estereotipadas;
- Em medidas de fluência, como rapidez e número de associações, o falante bilingüe apresenta mais problemas na língua mais fraca, que é menos organizada e menos acessível;
- Respostas foneticamente associadas (*clang associates*) constituem uma grande proporção das respostas produzidas por aprendizes menos proficientes, sugerindo que, pelo menos em certos estágios do aprendizado, as propriedades fonéticas das palavras da LE podem ser mais salientes do que suas formas semânticas.
- Não há evidência para afirmar que existem diferenças paradigmáticas/sintagmáticas entre aprendizes de LE.

A próxima seção considera o vocabulário em sua relação com a leitura.

2.4.3 – O vocabulário e a leitura

A revisão que acabo de apresentar explorou algumas questões relativas ao ensino e aprendizagem do vocabulário em geral. Grande parte dessa pesquisa focalizou o ensino explícito ou direto de vocabulário. Entretanto, esse ensino também pode ser indireto, ou seja, através do contexto e da leitura.

Nesta seção, subdividida em duas partes, focalizo o vocabulário em sua relação com a leitura. Esta relação pode ser vista sob duas perspectivas, isto é, de um ensino indireto, em que a leitura é vista como contexto para a aquisição do vocabulário, que será abordado na Seção 2.4.3.1; ou ainda sob um outro prisma, isto é, da contribuição do vocabulário para a leitura, que constitui o foco propriamente dito deste trabalho e será abordada na Seção 2.4.3.2. Tanto uma como outra situação são de relevância para o leitor não-nativo, que apresenta lacunas em sua competência lexical. As contribuições e as limitações dos desenvolvimentos dessa pesquisa servirão de ponto de partida para os aspectos investigados neste trabalho.

2.4.3.1 – A contribuição da leitura para a aquisição do vocabulário: uma abordagem indireta do vocabulário através da leitura

Além das abordagens de ensino direto do vocabulário, através de métodos e técnicas que focalizam exercícios específicos para sua expansão, tais como *palavra chave*, estratégias de inferência de palavras através do contexto, listas de palavras, dentre outras, conduzidas em sala de

aula, mencionadas na seção anterior, há também, na literatura, abordagens de ensino indireto de vocabulário. Nesse tipo de ensino, a atenção do aprendiz é geralmente focalizada no assunto ou tema do texto, e não no vocabulário. Embora, nessas situações, sejam usadas estratégias de inferência lexical, há uma distinção entre a construção do significado através de pistas contextuais e o real aprendizado e retenção desses significados. Enquanto estudos sobre inferência de palavras concentram-se no processo e produto da leitura, dando atenção aos tipos de pistas disponíveis no contexto, grau de sucesso e insucesso do aprendiz assim como o efeito da prática no uso dessas pistas (vide Seção 2.4.3.2 a seguir), estudos sobre a aprendizagem de palavras através do contexto, embora algumas vezes também considerem a presença de pistas, têm seu interesse mais voltado para o significado da palavra, depois que o aprendiz a encontra em contexto.

As abordagens de ensino indireto de vocabulário encontram apoio em teorias como a *Hipótese do insumo* de Krashen (*Input hypothesis*) (Krashen 1985, 1988, 1989). Seus seguidores reconhecem ser possível aprender um número bastante grande de palavras, mesmo quando a atenção do aprendiz não esteja diretamente voltada para o vocabulário. Entretanto, para que a aprendizagem ocorra, algumas condições são consideradas necessárias: a quantidade de palavras desconhecidas deve ser baixa, o aprendiz precisa ter interesse no assunto. Este, por sua vez, precisa conter alguns itens que estão apenas levemente acima do nível atual de rendimento do aprendiz que, por sua vez, devem ser entendidos através do contexto em que ocorrem, incluindo o contexto lingüístico e não lingüístico.

Ambos os métodos de ensino de vocabulário, ou seja, o ensino direto e o indireto, têm sido investigados empiricamente, e as duas abordagens, contrapostas e comparadas, dividindo a opinião dos pesquisadores, como se pode observar a partir das observações a seguir.

Para Jenkins, Stein & Wysocki (1984:769),

(...) a aprendizagem através do contexto ainda é uma explicação precária; ainda faltam evidências de que indivíduos realmente aprendem significados de palavras através de experiências com os contextos (48c).

Realmente, estudos em LE, que consideram a aprendizagem em contexto não têm mostrado o resultado apresentado em contextos de LM (Anderson & Shiffrin 1980). O estudo de Parry (1988) também apresenta evidências de que, mesmo palavras laboriosamente inferidas através do contexto, não são aprendidas (vide também Twadell 1980; Nagy & Herman 1987).

Entretanto, essa opinião não é consensual. Evidência a favor da possibilidade de aprendizagem de palavras em contexto é apresentada por Elley & Mangubhai 1981; Seibert 1930; Pickering 1982; Politzer & Politzer 1972; Henning 1973; Van Parreren & Schouten-van Parreren 1985, dentre outros; vide também Stahl & Fairbanks 1986, para uma meta-análise de vários estudos).

Nagy & Anderson (1984:304) concluem que

(...) mesmo o ensino de vocabulário mais direto e sistemático não dá conta de uma proporção significativa de todas as palavras que as crianças realmente aprendem, nem consegue cobrir mais do que uma modesta proporção das palavras que elas encontrarão nos materiais de leitura escolares (49c).

O estudo de Jenkins et alii (1984) corrobora essa conclusão. O autor conduziu um estudo em que palavras de baixa frequência eram apresentadas em contextos considerados *ricos* e informativos, através de textos lidos durante vários dias. As palavras foram aprendidas através do contexto, e quanto mais as palavras eram apresentadas, mais eram aprendidas. Entretanto, a aprendizagem não foi tão grande quanto se esperava. As palavras que faziam parte de uma lista com sinônimos e de uma oração em contexto tiveram um efeito maior do que aquelas apresentadas em textos.

Algumas hipóteses poderiam ser aventadas para explicar resultados aparentemente contraditórios. Uma delas poderia ser atribuída às variações dos próprios métodos de ensino, como também às limitações metodológicas dos estudos que se propõem a investigar o papel do ensino direto ou indireto do vocabulário, que, muitas vezes deixam de registrar a aprendizagem principalmente, quando esta não ocorre instantaneamente. Uma outra explicação estaria na falta de definições ou de conceitos comuns, tais como *o que é saber uma palavra, o que é aprender uma palavra* ou *quais as diferenças entre memorização de curto e longo prazo*, ou seja, na falta de explicitação dos critérios ou parâmetros usados no julgamento de palavras conhecidas, que permitam a comparação dos resultados dessas pesquisas.

Entretanto, a meu ver, as duas abordagens não são excludentes mas se complementam, uma vez que sua eficiência só pode ser julgada a partir do propósito que se destina. E, nesse sentido, os propósitos são diferentes. Nos estágios iniciais de aprendizagem de uma língua, um foco em técnicas de ensino direto pode ser eficiente, enquanto nos estágios mais avançados, o ensino passaria a ser conduzido através da leitura e contexto, ou seja, através de um ensino indireto. Isso não significa que o ensino direto seja baseado em palavras isoladas ou listas. A presença do contexto é fundamental nos dois tipos de abordagens. Tendo adquirido um vocabulário mais básico, o aluno estaria apto a expandi-lo, incorporando palavras de baixa frequência, através de estratégias de inferência e uso de contexto, ou seja, numa abordagem indireta. Como coloca Krashen (1984), seria impossível ensinar vocabulário através do contexto e leitura sem as condições necessárias. A meu ver, essas condições parecem estar presentes mais frequentemente em contextos de LM, menos em contextos de L2, e muito menos ou quase nunca em contextos de LE. Elas não implicam apenas no engajamento do aprendiz em um grande volume de leitura extensiva, mas, principalmente, no desenvolvimento de

de uma base lingüística e habilidades de inferência, ou de uso desse léxico através do contexto. A utilização de uma abordagem indireta em estágios iniciais, ou de um ensino não intervencionista de vocabulário, combinado com uma situação em que a qualidade e quantidade de insumo recebido são geralmente bastante limitados, tem sido responsável por grande parte da frustração e incapacidade dos alunos, principalmente frente aos textos de leitura.

Tendo apresentado, nas duas últimas seções, uma retrospectiva da importância do vocabulário, assim como uma revisão de pesquisas sobre seu ensino/aprendizagem, passo a considerar, na próxima seção, sua relação com a leitura, foco propriamente dito desta investigação.

2.4.3.2 – O componente lexical nas pesquisas em leitura

Complementando a Seção 2.2 deste capítulo, que considerou a relação entre o componente lingüístico e a leitura, apresento, a seguir, uma revisão da literatura que focaliza a relação da leitura com o vocabulário, ou mais especificamente, que avalia o papel desse conhecimento para a compreensão em leitura.

Uma forte relação entre esses dois aspectos tem sido mostrada inúmeras vezes na literatura, através de métodos diferentes, que incluem análises fatoriais (Davies 1944, 1972; Spearitt 1972) ou correlações entre vocabulário e medidas de compreensão (Farr 1969, dentre outros) ou ainda na pesquisa sobre legibilidade (Chall 1958; Harrison 1980).

Muitas hipóteses têm sido aventadas para explicar essa associação. Anderson & Freebody (1983) salientam três: a instrumentalista, a da aptidão geral e a do conhecimento geral. A instrumentalista sugere que o conhecimento de vocabulário tem um efeito direto na compreensão, recomendando, portanto, uma abordagem de ensino intervencionista ou direta de vocabulário. As outras duas hipóteses parecem associar a relação do vocabulário e da leitura a um terceiro fator, que pode ser inteligência ou conhecimento de mundo.

Entretanto, apesar dessas evidências, também podem-se observar estudos que não mostram de maneira consistente a força dessa relação, apresentando resultados apontando para uma direção oposta (vide Alderson & Richards 1977b).

As investigações que abordam a relação entre conhecimento lexical e leitura distinguem-se pela ênfase dada às questões pesquisadas, isto é, conhecimento lexical, inferência em contexto e compreensão. Entretanto, a maioria delas focaliza a inferência lexical como aspecto central, onde sua relação com o conhecimento lexical é explorada. Estudos dessa natureza, em geral, podem avaliar ou não a existência de um nível lexical limiar que, nesse caso, refere-se à inferência (vide Scott 1984, 1990; Haarstrup 1987, 1989).

A inferência lexical (27) é normalmente vista na literatura como dependente de um

conjunto de estratégias utilizando pistas oferecidas pelo contexto. O uso do dicionário também seria uma estratégia para se chegar ao significado de palavras, mas que não se utiliza do contexto. Esse tipo de inferência não deve ser confundido com um outro, central na construção do sentido do texto, que se relaciona

(...) à solução das decorrências, pressuposições e implicaturas, usando suposições transitórias e coisas do tipo, procedimentos esses necessários para toda a 'leitura entrelinhas' que se faz todo o tempo (Scott 1984) (50c).

Como mostrei na Seção 2.1 deste capítulo, foi apenas no início dos anos 70 que o ensino de vocabulário passou a considerar a importância de estratégias de inferência lexical (28). Como consequência, até o início dos anos 80 eram escassos os estudos que tratavam dessa questão, principalmente em LE. Sem dúvida alguma, as técnicas introspectivas parecem ter contribuído de forma efetiva para esses estudos, uma vez que permitiram uma maior aproximação ao processo de leitura.

São muitas as taxonomias de estratégias de inferência lexical propostas na literatura. De uma maneira geral, elas se utilizam dos vários níveis de contexto propostos na Seção 2.3. Uma das estratégias que mais se destaca é o reconhecimento de cognatos, embora, como salienta Holmes, (1986) não se tenha notícias de muitos estudos que apresentem uma justificativa teórica para o uso dessa noção, enquanto um elemento facilitador. Cognatos são itens lexicais ou palavras isoladas que têm uma semelhança, em forma e em sentido, com as palavras na LM do leitor. Falsos-cognatos são semelhantes na forma mas diferentes no significado. A estratégia consiste na procura dos cognatos do texto como uma atividade facilitadora a ser conduzida anteriormente à leitura. As justificativas para o uso dessa estratégia estão relacionadas a um aumento de motivação e confiança dos alunos no início dos cursos de leitura.

A inferência lexical é ainda considerada sob o ponto de vista de produto ou de processo, ou ambos. Nas investigações em que é vista enquanto produto, é analisado seu grau de sucesso/insucesso, por vezes relacionado à proficiência lingüística ou, mais especificamente, lexical. Os efeitos desse conhecimento para a inferência parecem variar. Nos trabalhos em que é considerada enquanto processo geralmente salientam-se os tipos de conhecimento que o leitor utiliza para inferir palavras, ou as estratégias usadas. Algumas implicações para o ensino de vocabulário ou uma relação entre inferência em contexto e retenção de significados de palavras são apresentadas em algumas dessas investigações (vide, por exemplo, Van Parreren & Schouten-van Parreren 1981). Relações entre inferência e compreensão não são em geral estabelecidas nesses estudos. Muitos desses resultados são contraditórios, uma vez que, em alguns deles, a inferência é possível sem um limiar de conhecimento de vocabulário e, em outros, ela depende desse limiar, embora muitas vezes não

explicitado.

Apesar dessas evidências, grande parte da literatura sobre ensino de vocabulário e ensino de leitura parece se concentrar na inferência através do contexto, vista como solução para os problemas de vocabulário insuficiente apresentados pelo falante não nativo, independentemente de um nível limiar de conhecimento lingüístico, ou, em outras palavras, como se esse conhecimento pudesse ser compensado, através da transferência automática de estratégias de inferência da leitura em LM. Reitero, mais uma vez, que esses trabalhos, em geral baseados em uma analogia que pode ser questionável entre leitura em L2/LE e modelos descendentes do leitor fluente em LM, parecem ignorar as dificuldades envolvidas nesse processo e os resultados inconsistentes que caracterizam os estudos sobre inferência, que mostro mais adiante.

Nation & Coady (1988) expressam muito bem a situação colocada pela pesquisa sobre vocabulário e leitura em LE:

Tipicamente, aprendizes de inglês como L2 são maus decodificadores, uma vez que seu conhecimento de vocabulário é fraco, enquanto, ao mesmo tempo, eles são leitores competentes em sua LM e já familiarizados com um processamento descendente. Portanto, torna-se importante considerar se nosso ensino deveria enfatizar um processamento descendente ou ascendente, assim como a ênfase apropriada no uso de contexto (51c).

Mais recentemente, entretanto, tem havido uma série de artigos sobre ensino de leitura (vide Eskey 1988; Eskey & Grabe 1988) que colocam a inferência em uma perspectiva mais adequada. Esses artigos se ancoram em resultados de pesquisa que discutem a questão de uma maneira mais sistemática (vide Haynes 1983, dentre outros), apontando para os inúmeros problemas que o leitor de L2 ou LE têm que enfrentar durante a leitura, ou durante a aquisição do vocabulário através do contexto. Muitas dessas experiências têm sido ineficientes e frustrantes para o aluno, que não consegue construir um contexto para se apoiar.

A revisão de alguns estudos que apresento a seguir tem, por objetivo, caracterizar em linhas gerais, o estágio atual da pesquisa na área que considera a interface vocabulário/leitura (Van Parreren & Schouten-van Parreren op. cit.; Haarstrup 1989a; Haynes op. cit.; Parry 1988). Incluo, nessa revisão, alguns trabalhos em contexto de LE no Brasil (Scott 1984 e 1990; Kleiman 1985).

Scott (1984), em um estudo com 25 alunos universitários lendo em inglês com propósitos específicos, discute diversos aspectos da inferência lexical, nesse caso, restrita a inferências de substantivos. Dos pontos discutidos, focalizo a nível limiar e as correlações. Apesar de o nível de conhecimento de vocabulário verificado antes da leitura do texto estar fortemente relacionado à compreensão, não parecem haver evidências, nesse caso, para um nível limiar. As correlações, entretanto, entre inferência e conhecimento de vocabulário, e entre inferência e habilidade de

sumarização, são fracas. Para o autor,

Esse resultado parece indicar que mesmo alunos fracos são capazes de inferir bem, o que os ajudou a produzir resumos ligeiramente melhores, embora não tão melhores quanto se já soubessem as palavras antes da leitura (52c).

O estudo de Scott (1990) sobre inferência é bastante abrangente e focaliza questões que incluem sucesso/insucesso, atitudes com relação à estratégia de inferência e outros tipos de estratégias para lidar com palavras desconhecidas, tais como o uso do dicionário. Embora o autor focalize a inferência e o conhecimento de vocabulário, não os relaciona com a compreensão. Este estudo, assim como o anterior, é particularmente relevante para esta investigação, uma vez que seus leitores são alunos brasileiros em uma situação de leitura em inglês como LE. Como o próprio autor salienta, "... a primeira língua dos sujeitos tem um papel extremamente importante na sua aprendizagem de LE e especificamente no processo de inferência lexical" (53c). Um outro aspecto saliente do estudo de Scott (op. cit.) é a consideração da inferência também sob o ponto de vista de processo.

Os resultados de Scott (1990) são bastante otimistas, mostrando que a inferência não só é possível com leitores de nível de proficiência mais avançados, ou seja, dependentes de um nível limiar de conhecimento de vocabulário, mas também quando usada por leitores menos proficientes. O autor, contudo, não encontra evidências para um nível limiar de conhecimento de vocabulário que impeça a inferência. As evidências apresentadas, portanto, sugerem que mesmo sujeitos que apresentam vocabulário bastante limitado em L2 são razoavelmente bem sucedidos na inferência de palavras desconhecidas. Isso implica na necessidade, para a maioria das inferências, de muito menos contexto do que se poderia imaginar ou seja, um contexto restrito a algumas palavras à esquerda e à direita do item alvo. Inferências implausíveis por parte do sujeito/leitor são resultado do reconhecimento incorreto das formas das palavras, ou do fato de se ignorar quantidades mais extensas de contexto. O sucesso na inferência também parece relacionado à familiaridade do conceito, uma vez que conceitos desconhecidos diminuem grandemente a chance de a inferência ser bem sucedida.

Com relação ao segundo aspecto pesquisado, ou seja, tipos de estratégias usadas, a conclusão desse estudo é que

(...) não parece haver um padrão de uso firmemente associado com as maiores taxas de sucesso (...) Pelo contrário, os sujeitos fizeram grande uso do conhecimento prévio e estratégias baseadas no léxico e pouco uso de características gramaticais ou discursivas (Scott op. cit.:224) (54c).

Infelizmente o autor não relaciona a habilidade de inferência à compreensão, o que ele mesmo reconhece como uma limitação do trabalho.

Um outro estudo já mencionado na Seção 2.3 é conduzido por Haarstrup (1989a). A autora contrasta dois grupos de leitores dinamarqueses de níveis alto e baixo de proficiência lendo em inglês como LE e produz uma taxonomia de fontes de conhecimento. Seus resultados mostram que os leitores mais proficientes usam, em média, duas fontes de conhecimento em um processamento paralelo, que integra processamento descendente através do uso de pistas ascendentes, enquanto o leitor menos proficiente usa apenas uma, contradizendo os resultados de Van Parreren e Shouten-van Parreren (op. cit.), abaixo apresentados, que mostram um processamento serial.

As fontes mais usadas, em ordem decrescente de utilidade, são:

**Quadro 2.4 – Fontes de conhecimento mais usadas
(Haarstrup 1989a)**

-
1. Co-texto
 2. A forma da palavra
 3. Conhecimento de mundo
 4. Conhecimento da LM
-

Enquanto o grupo de proficiência alta usa melhor as fontes de conhecimento, o grupo mais fraco distingue, em geral, itens com fonte em LM. A correlação entre a proficiência em LE e sucesso de inferência foi bem baixa (0,24 para o grupo forte e 0,19 para o outro grupo), confirmando os resultados de Scott (1990). Entretanto, considerando-se os dois grupos, pode-se observar uma tendência mais forte de associação entre proficiência em LE e sucesso de inferência. Para a autora, esses resultados poderiam ser considerados evidências da existência de um nível limiar de proficiência em L2 para a inferência lexical.

O estudo de Kleiman (1985), de natureza exploratória, investiga as estratégias de inferência lexical utilizadas por alunos brasileiros principiantes na leitura em inglês. Os alunos demonstram utilizar uma estratégia de uso do contexto global ou de identificação temática (tema primário ou tópico) e duas estratégias de utilização do contexto imediato. A autora conclui que

Não há evidência generalizada, como no caso das estratégias de identificação temática e identificação de contexto próximo, de que o aluno utilize contextos lineares de maior abrangência, mais distantes, que envolvam o reconhecimento e análise de paralelismos estruturais (Kleiman op. cit.:75).

Nesse estudo, embora o aluno consiga, apesar de problemas de processamento, chegar a

um significado aproximado para certos trechos do texto que contém palavras não inferidas ou incorretamente inferidas, verifica-se que muitos deles são distorcidos ou apagados nos resumos.

Os estudos de Van Parreren & Schouten-van Parreren (op. cit.) focalizam o processo da inferência, analisando, como o de Haarstrup (op. cit.), as fontes de conhecimento usadas nesse processo, entretanto, de maneira indireta, ou seja, através da análise dos erros ou inferências mal-sucedidas dos sujeitos. Estes foram classificados em quatro níveis: sintáticos, semânticos, lexicais e estilísticos. Com base nessa classificação, os autores afirmam que a inferência lexical deve ser vista como uma hierarquia de quatro sub-habilidades. Um sujeito pode agir corretamente em um certo nível apenas se não houver problema nos outross. Um leitor competente é capaz de estimar quais são as dificuldades envolvidas no processo, e entrar no nível mais apropriado (Van Parreren & Shouten-van Parreren (op. cit.), sem passar por todos os níveis da hierarquia.

Scott (1990:81), entretanto, não parece concordar com o modelo proposto, uma vez que, na sua opinião, não parece haver "... evidência clara de que a hierarquia opera como se fosse rua de mão única ... O fato de se poder classificar os erros de acordo com seu esquema não prova que o leitor, que se engana na classe de um item, por exemplo, não possa fazer uma estimativa plausível de seu significado" (55c).

Embora o estudo de Van Parreren & Schouten-van Parreren (op. cit.) não mencione níveis limiares de conhecimento, pode-se notar que a inferência exige vários pré-requisitos. Esse estudo não só considera a importância da estratégia de inferência para o acesso aos significados das palavras mas também para o ensino de vocabulário (vide observação de Haarstrup mais adiante).

O que é importante salientar no trabalho de Van Parreren & Shouten-van Parreren é que os contextos que levam à inferência são contextos especiais, denominados *grávidos* (*pregnant*), que oferecem muitas pistas para a inferência. Também importantes são fatores das próprias palavras a serem inferidas, tais como classe gramatical, grau de abstração, transparência de estrutura, dentre outros, e, obviamente, fatores do leitor, tais como conhecimento e habilidade que permitam perceber esses fatores enquanto pistas, ou usar o contexto semântico, sintático, morfológico, discursivo, além de conhecimento de mundo. Essa visão *rica* de contexto, portanto, poderia ser associada a uma visão rica de competência lexical, que abordo na próxima seção, embora essa relação não seja normalmente explicitada na literatura. Um contexto *rico*, mais bem definido como *potencialmente rico*, como já salientei, seria aquele que traz uma variedade de pistas, oferecendo ao leitor, ampla oportunidade de acesso aos vários níveis desse contexto. Entretanto, apenas através de uma competência lexical *rica*, que não se restringe apenas aos significados das palavras é que o leitor poderia explorar esses recursos.

Embora o uso de contextos *ricos* possa ser importante sob o ponto de vista da motivação e exemplificação da tarefa de inferência, assim como de um tratamento *rico* ao vocabulário, oferecendo ampla oportunidade de acesso ao leitor, concordo com Haynes (1983) que a seleção e

uso constante de contextos ricos em sala de aula pode dar, ao leitor, uma falsa idéia da facilidade da inferência, uma vez que os contextos em geral, encontrados em textos lidos fora da sala de aula, não são, na maioria das vezes, desse tipo. Nesse trabalho, a autora também mostra que a tarefa de inferência de palavras para leitores menos proficientes é mais complexa do que tem sido afirmado, principalmente, quando há necessidade de compreensão de longos trechos de contexto, que dependeriam, além do contexto local (conhecimento de co-texto) e global (conhecimento prévio, crenças, valores e restrições do contexto de leitura) também do conhecimento acumulado. O contexto imediato (contexto local) não parece ser problemático, mesmo para leitores menos proficientes. A autora reconhece, nesse processo, entretanto, a necessidade de conhecimento lingüístico.

Nesse sentido concordo com Haarstrup (1989b) quando salienta que

(...) enquanto o processamento (intake) para a compreensão é sempre facilitado por um contexto rico, parece provável que o processamento para a aprendizagem é, por sua vez, enfraquecido (56c).

Essa posição também é defendida por Mondria & Wit-de Boer (1991), contrariando os resultados de Van Parreren & Schouten-van Parreren (op. cit.) e outros que reconhecem a importância do contexto na aprendizagem das palavras.

Os estudos que acabei de examinar têm a inferência como seu aspecto central. Os resultados a seguir mostram ênfase em outros aspectos. Alguns desses estudos, entretanto, já foram mencionados na seção que discutiu o nível limiar.

Bensoussan et alli (1984) estudaram o efeito do uso de dicionários durante testes de leitura. Os autores não encontraram diferença significativa entre aqueles que usaram o dicionário e aqueles que não o usaram, o que os levou a concluir que o acesso aos significados das palavras não, necessariamente, leva à compreensão.

Alderson & Richards (1977b), em um estudo das dificuldades encontradas por estudantes mexicanos lendo em inglês investigaram fatores como conhecimento de vocabulário, organização retórica, complexidade sentencial e referência contextual em sua relação com a compreensão. Sua conclusão foi que embora a correlação mais forte tenha sido apresentada por vocabulário, ela ainda é fraca (0,47).

Singer e Crouse (1981), embora trabalhando com leitores em LM, propõe um limiar para a compreensão, argumentando que o efeito do vocabulário é quase que totalmente independente do uso do contexto.

Laufer (1987), em um estudo com 100 alunos (falantes de árabe e hebraico) lendo em inglês com objetivos acadêmicos conclui que as chances da leitura serem bem sucedidas são muito maiores, se o leitor souber 95% ou mais do vocabulário do texto, embora também contribuam para isso, as pistas gramaticais, a organização textual e a familiaridade com o assunto. A autora se refere

a um nível limiar de conhecimento de vocabulário que é relativo, ou seja, considera a proporção de palavras desconhecidas/conhecidas no texto e não a extensão total do vocabulário do leitor.

Liu & Nation (1985), em um estudo com aprendizes avançados de L2 mostram que os aprendizes proficientes conseguem inferir entre 85 e 100 por cento das palavras desconhecidas. Essa porcentagem diminui bastante no caso de aprendizes com proficiência mais baixa.

Clarke & Nation (1980), em um estudo, cujo objetivo era propor estratégias para a inferência de itens lexicais em contexto, sugerem que um vocabulário de 3000 palavras permite ao leitor inferir, em média, de 60 a 70% das palavras desconhecidas de um texto, podendo, até mesmo, chegar a 80%. Entretanto, não salienta nenhum tipo de explicação para alunos com um vocabulário inferior a 3000. Este seria, embora não explicitado pelo autor, um nível limiar que leva em conta o vocabulário total do leitor.

Para Johns (1978), esse limiar consistiria em aproximadamente 50 palavras desconhecidas em 1000, resultando em um bloqueio na percepção das estruturas do texto e na impossibilidade de inferência através do contexto. Enquanto o limiar de Clarke & Nation (op. cit.) refere-se à capacidade de inferência, o proposto por Johns refere-se à compreensão.

Alguns desses estudos, como o de Van Parreren & Schouten-van Parreren (op. cit.), vão além da investigação da relação entre compreensão e vocabulário, uma vez que também incluem a relação entre o ensino de significado de palavras e a compreensão. Alguns deles mostram, por exemplo, que textos com muitas palavras desconhecidas não são bem compreendidos mas que a subsequente aprendizagem dessas palavras nem sempre levam a uma compreensão melhor (Parry 1988; Tuinman & Brady 1974; Pany & Jenkins 1978; Jenkins, Pany & Schreck 1978).

O estudo de Parry (op. cit.) foi conduzido com apenas um informante, uma falante de japonês de nível intermediário lendo inglês em uma situação de L2. O texto usado é bastante longo (48.000 palavras). Seus resultados mostram inferências feitas a partir de grandes porções do texto. A extensão é, dessa forma, vista como um elemento facilitador, possibilitando ao leitor criar redundâncias. Esta investigação tem implicações diretas para o ensino/aprendizagem que contrariam os resultados de Van Parreren & Schouten-van Parreren: apesar do grande número de inferências bem sucedidas, não houve aprendizagem dessas palavras.

Johnson (1982) procura avaliar o efeito do conhecimento prévio (visto enquanto familiaridade cultural), assim como do ensino de vocabulário, na compreensão de um texto em inglês como LE. A passagem sobre Halloween usada na coleta de dados continha uma parte considerada pelo pesquisador como familiar para os sujeitos e uma outra não-familiar. Seus resultados mostram que o impacto da familiaridade cultural, na compreensão, foi maior do que o ensino do vocabulário.

A pesquisa de Hudson (1988) apresenta resultados semelhantes aos de Johnson (op. cit.). Ele mostra que as pistas oferecidas aos leitores através de figuras são mais úteis, sob o ponto de vista da compreensão, do que o ensino de vocabulário, sendo particularmente útil aos alunos considerados

fracos.

Apesar de todos esses resultados conflitantes, Nation & Coady (1988:108) salientam que

Em geral, a pesquisa nos deixa pouca dúvida sobre a importância do conhecimento de vocabulário para a leitura e sobre o valor da leitura como um meio de melhorar o vocabulário. A natureza exata dessas relações, e como podemos fazer uso delas em nosso ensino, ainda são áreas que merecem ser investigadas (57c).

Entretanto, o que é mais sério não são resultados conflitantes – que podem, a meu ver, ser explicados, uma vez que parecem partir de arcabouços teóricos diferentes ou refletir posicionamentos de leitura/ensino diversos – mas a dificuldade em compará-los pela falta de explicitação, de uma maneira geral, da base teórica que os fundamenta. Um conceito central em estudos que investigam a relação entre o vocabulário e a leitura é o conceito de conhecimento/competência lexical, ou, em outras palavras, o que significa *conhecer uma palavra*. A definição desse conceito, a priori, é necessária para a avaliação desse conhecimento/competência, da mesma forma que um conceito ou visão de leitura é essencial quando se vai avaliar a leitura, como mostrei no início deste capítulo. Sua explicitação, entretanto, não parece preocupação de alguns dos trabalhos acima salientados. Alguns conceitos ou modelos de competência/conhecimento lexical serão abordados, portanto, na próxima seção.

2.4.3.3 – A competência/conhecimento lexical e sua avaliação

Nesta seção, focalizo alguns modelos ou conceitos de competência/conhecimento lexical existentes na literatura, e abordo questões relativas à sua avaliação. As duas questões são fundamentais para justificar o modelo de competência lexical deste estudo, completando desta forma, a base teórica que o fundamenta. Ao invés de iniciar esta discussão, apresentando os modelos, entretanto, tomo como ponto de partida sua avaliação, uma vez que os modelos estão a ela subjacentes.

Os trabalhos na área de avaliação de vocabulário são escassos. Essa situação, sem dúvida alguma, pode ser explicada pela negligência com relação ao vocabulário que salientei no início deste capítulo. Os avanços observados durante os anos 80, referentes ao ensino e aquisição/aprendizagem do vocabulário, ainda não parecem ter sido incorporados à avaliação, apesar da importância desta área para o ensino e principalmente para a pesquisa. A falta de instrumentos válidos e confiáveis para a avaliação de vocabulário, juntamente com a não explicitação do modelo de competência/conhecimento da palavra que os tem fundamentado, como mostrei acima, têm

dificultado a interpretação de seus resultados e impedido comparações. Isso significa que embora a avaliação possa ser culpada por alguns desses problemas, há outros, que são anteriores a ela. Além disso, embora os argumentos mais freqüentemente pareçam favorecer, principalmente na avaliação de rendimento e proficiência, testes integradores, em que o vocabulário é avaliado juntamente com outras competências e habilidades, não se pode deixar de reconhecer a importância de testes de apenas de vocabulário, principalmente para fins de diagnóstico e de pesquisa.

A primeira distinção feita na avaliação do vocabulário, visto enquanto um componente da leitura, é com relação ao objetivo, isto é, de reconhecimento ou receptivo, ou de produção ou produtivo. A dicotomia produtivo/receptivo tem sido usada, portanto, na área de ensino/aprendizagem de vocabulário, geralmente para representar tipos diferentes de conhecimento: conhecer uma palavra para uso produtivo pressupõe conhecê-la melhor do que apenas para uso receptivo (Nation 1984), e, nesse sentido, pode ser vista como um tipo de sequência de desenvolvimento (vide Melka Teichroew 1982; Crow 1986 e Corson 1983, para uma discussão mais detalhada dessa questão). Para Crow (op. cit.:242)

(...) o controle receptivo sobre o vocabulário em uma L2 não envolve a aprendizagem de novos conceitos ou a aprendizagem de uso de palavras. Ele envolve, entretanto, a aprendizagem de novos rótulos para áreas semânticas e o desenvolvimento da habilidade de reconhecimento desses rótulos em contexto, como um processamento ascendente inicial (58c).

Para Corson (op. cit.:6),

Um vocabulário ativo é talvez mais bem descrito como um vocabulário motivado. Ele consiste em todas aquelas palavras necessárias para um uso sem restrições, para nos comunicar com outras pessoas no dia a dia (...) Um vocabulário passivo pode ser considerado como incluindo o vocabulário ativo do indivíduo como um importante sub-conjunto. Ele difere do vocabulário ativo, uma vez que também incorpora o vocabulário não motivado do indivíduo (59c).

A justificativa freqüentemente citada para o uso da dicotomia é que ela parece mostrar, de maneira mais clara, uma perspectiva de desenvolvimento, isto é, de que a palavra é aprendida receptivamente antes de passar para o vocabulário produtivo do aprendiz, sendo, por esse motivo, melhor do que a dicotomia ativo/passivo empregada anteriormente com o mesmo sentido. Apesar de concordar com o fato de que, em geral, entende-se um número muito maior de palavras do que se é capaz de produzir, e que em situações de LE, muitas vezes, realmente, não há necessidade de se aprender conceitos, uma vez que eles já são conhecidos em LM, não me parece suficientemente

clara a extensão e o significado da lacuna entre um vocabulário receptivo e produtivo. A observação de Melka Teichroew (op. cit.) é pertinente ao mostrar "... como é insatisfatório tentar definir a distinção em termos operacionais e tentar especificar o que a distinção significa no nível conceitual ..." (60c). Ela argumenta que qualquer tentativa de classificar palavras como sendo receptivamente ou produtivamente conhecidas, ou tentar identificar um ponto específico ou um grau de familiaridade que faz a transição entre vocabulário receptivo e produtivo resulta em uma divisão arbitrária que, na opinião de Read (1987) também não parece muito útil. Uma melhor alternativa seria conceber o conhecimento de vocabulário como um contínuo de níveis ou graus de familiaridade com palavras, em que palavras mais familiares seriam mais conhecidas.

Além da distinção produtivo/receptivo não ser satisfatória em termos teóricos, também não o é em termos práticos. Crow (op. cit.:245), apesar de enfatizar a necessidade da dicotomia e discuti-la teoricamente, cita problemas práticos na tentativa de classificação de testes ou exercícios em receptivos ou produtivos:

Os professores freqüentemente assumem que qualquer exercício de vocabulário ou teste pressupõe uma escolha entre conhecimento ativo ou passivo das palavras. Isso simplesmente não acontece: julgar os pre-requisitos de uma atividade de vocabulário significa julgar sua posição em um contínuo de dois polos (...) abordar o ensino de vocabulário através dessa perspectiva pode impedir professores de fazer julgamentos adequados sobre os requisitos das várias formas do vocabulário" (61c).

Várias maneiras têm sido utilizadas na literatura para a avaliação do vocabulário. A prática padrão, entretanto, tem sido combinar uma palavra com um sinônimo ou com uma definição analítica do tipo encontrada em dicionários, usando métodos variados tais como múltipla escolha, preenchimento de lacunas ou ainda tradução (vide Nation 1983). A implicação desse tipo de medida é que cada palavra pode ter apenas um significado e que conhecimento de uma palavra é funcionalmente equivalente a ser capaz de reconhecer ou lembrar uma definição (Read op. cit.).

O uso de escalas de verificação (*checklists*) também têm sido freqüentes. Esse instrumento pode ser considerado não apenas uma das medidas de vocabulário mais antigas, mas uma das mais controvertidas (usadas com falantes nativos desde 1890), uma vez que, para muitos, elas apenas constituem testes de familiaridade ou auto-avaliações de familiaridade e não de conhecimento, sendo, portanto, de valor limitado.

Várias listas de verificação do tipo *sim/não* (*yes/no method*), com dois níveis de familiaridade, têm sido propostas na literatura para a avaliação da extensão do conhecimento de vocabulário (vide Champion & Elley 1971; Anderson & Freebody 1983 e mais recentemente Mcara & Jones 1988). Para um de seus proponentes (Mcara no prelo), uma de suas grandes vantagens está

no fato de se constituir uma medida aproximada de proficiência global em LE que é simples de ministrar, principalmente se usada como teste de filtragem ou de seleção (*screening test*), quando a pesquisa envolve um grande número de sujeitos. Uma versão computadorizada tem sido usada pelos Eurocentres como teste classificatório. Sua base teórica está em trabalhos conduzidos em inglês como LM (Zimmerman et alii 1977; Anderson & Freebody op. cit.; Meara & Jones op. cit.; Meara & Jones 1990).

Esse procedimento consiste na apresentação de uma lista de palavras isoladas (reais e inventadas, na proporção de duas reais para uma inventada) uma de cada vez, que o avaliado deverá indicar se conhece ou não. As palavras reais fazem parte de uma lista de palavras, elaborada por Nation (1986), enquanto as inventadas são nomes próprios reais, selecionados da lista telefônica do Reino Unido, ou então construídas com base em morfemas latinos e gregos. A inclusão de palavras inventadas justifica-se como uma forma de minimizar as chances de adivinhação, uma vez que, ao identificar como conhecida uma palavra inventada, o avaliado é punido. Cálculos matemáticos complexos são usados para projetar a extensão total do vocabulário do avaliado.

O teste é realizado através do computador, não exigindo a intervenção do professor nem mesmo para a correção, pois a nota é dada automaticamente. O resultado do teste é apenas um escore, que estima a extensão total do vocabulário do avaliado. Segundo Meara (op. cit.), esse escore apresenta correlações bastante altas com outras habilidades linguísticas, principalmente compreensão em leitura, compreensão oral e precisão gramatical.

Devo salientar que esse tipo de instrumento tem subjacente um modelo de conhecimento de vocabulário que se baseia em um contínuo de níveis ou estágios ordenados de desenvolvimento – que poderia incluir o contínuo produtivo/receptivo – ilustrando uma das formas de se caracterizar a composição do conhecimento/competência de vocabulário identificadas na literatura (vide Dale 1965; Russell & Saadeh 1962 e Melka Teichroew op. cit.). Uma outra forma, entretanto, seria como um conjunto de componentes ou aspectos (vide também Cronbach 1942; Richards 1976; Nation 1984).

Antes de apresentar algumas características desses modelos, entretanto, fazem-se necessárias duas observações: a primeira, diz respeito à definição do termo *palavra* e, a segunda, a uma distinção entre *palavra* e *conceito*.

No primeiro caso, uma maneira um tanto simplificada mas conveniente, emprestada de Scott (1990), e adotada neste estudo, consiste em denominar de *palavra* qualquer constituinte que seja separado, no texto escrito, dos dois lados, por um conjunto de separadores, ou seja, ou espaço em branco. Verbos frasais seriam itens lexicais compostos de duas ou mais palavras (ver Halliday e Hasan 1989; Sinclair & Renouf 1988; Nagy & Anderson 1984 para uma discussão mais detalhada). Halliday & Hasan (op. cit.) ainda distinguem entre *itens lexicais* e *gramaticais*: itens lexicais são palavras de conteúdo, enquanto itens gramaticais são palavras de função. São lexicais

porque funcionam em conjuntos abertos, enquanto as palavras de função funcionam em conjuntos gramaticais, ou fechados. A linguagem escrita caracteriza-se por apresentar uma proporção muito maior de itens lexicais; ela é, por isso, mais densa do que a linguagem oral, que por sua vez é mais difusa.

Ainda em referência à definição de palavra, observa-se que, para Leech (1981:9 e 13), por exemplo, há dois tipos de significado: conceitual e denotativo, que é contrastivo e tem estrutura, referindo-se ao objeto; e conotativo, que é periférico, indeterminado e aberto. Lyons (1968:196) distingue três usos do termo *palavra* em termos de realização. *Sang*, por exemplo, é a realização ortográfica da forma mais gramatical e abstrata do passado da palavra *sing*, que por sua vez, é menos abstrato do que a noção de lexema ou lema *sing*. O conceito *sing* representa o significado, ou *produzir sons musicais com a boca*.

Embora *conceito* e *palavra* sejam teoricamente separados – é possível saber uma palavra sem conhecer o conceito, da mesma forma que é possível conhecer um conceito sem saber a palavra, uma vez que essas situações ocorrem em LE– há, sem dúvida alguma, uma relação estreita entre conceitos e palavras, embora não se saiba muito bem como essas relações funcionam.

Apresento a seguir alguns modelos de conhecimento/competência lexical existentes na literatura.

Dale (1965), em uma pesquisa sobre medidas de vocabulário para falantes nativos, identifica quatro estágios no conhecimento ou familiaridade com uma palavra:

Quadro 2.5 – Conhecimento lexical (Dale 1965)

-
1. Distinguir entre palavras e "não palavras" ou reconhecer que é uma palavra, sem, no entanto, saber seu significado
 2. Colocar a palavra vagamente em contexto
 3. Conhecer o significado da palavra
 4. Ser capaz de distinguir palavras relacionadas em significado e/ou forma
-

Para esse autor, "... o conhecimento de uma palavra pode ser colocado em um contínuo, que começa com 'eu nunca vi essa palavra antes' (op. cit.:898) (62c). Esse estágio distingue entre palavras existentes e não existentes na língua. O segundo estágio refere-se à colocação da palavra em um contexto, que não precisa ser específico, fazendo referência, portanto, a regras de uso. Nesse estágio as palavras estariam em uma "zona crepuscular" entre o reconhecimento e conhecimento. Com relação ao terceiro aspecto, Dale (op. cit.:898) afirma "... e finalmente nós atingimos o estágio em que identificamos a palavra. Nós sabemos seu significado. Nós a reconheceríamos novamente se a vissemos, e provavelmente lembraríamos seu significado" (63c). Não fica claro se Dale refere-se aqui a *ser capaz de definir a palavra* ou apenas *identificar a palavra*. O quarto e último estágio vai além do conhecimento de um significado da palavra:

(...) pode-se elaborar testes ainda mais precisos exigindo discriminações mais finas, como entre controlar (arrest) e curar (cure) uma doença; ironia (irony), sarcasmo (sarcasm) e sátira (satire); compaixão (sympathy) and tolerância (empathy) (...) (Dale op. cit.:898) (64c).

Embora uma série de níveis ou um contínuo de desenvolvimento seja uma maneira interessante de se conceber a competência lexical, não ficam claras, na proposta acima, as diferenças entre esses níveis. Além da distinção entre reconhecimento e conhecimento, que, em princípio, poderia implicar na dicotomia receptivo/produtivo, o autor não faz qualquer referência explícita a ela.

Um outro modelo, que também se baseia em uma série ordenada de estágios, é o modelo de Haarstrup (1989a) Para a autora, o conhecimento do aprendiz pode ser representado em uma escala de quatro níveis:

Quadro 2.6 – Modelo de conhecimento da palavra (Haarstrup 1989a)

1. O significado central (*core*) da palavra
2. O significado central mais partes do significado periférico da palavra
3. O significado central mais potencialidades associativas remotas da palavra
4. O significado total da palavra

O modelo de Scott (1990) difere do de Haarstrup (op. cit.), uma vez que é visto como um conjunto de componentes. Para Scott (op. cit.), conhecer uma palavra pressupõe pelo menos três condições necessárias: saber o conceito (denotação), saber o que a palavra não é (também denotação), e saber quais as outras palavras ou conceitos com os quais a palavra pode se relacionar (paradigmaticamente ou sintagmaticamente), ambos envolvendo denotações e conotações. Isso significa estar consciente das relações entre a palavra (em um dos três níveis de realização acima mencionados) e o conceito ao qual está relacionada. Conhecer a palavra, nesse contexto, está, portanto relacionado com seu significado. Esta é uma visão centrada no significado. Para o autor,

Conhecer o conceito não implica em saber muito sobre o conceito, e não implica também ter experimentado o referente; isso implica em saber apenas em que 'gancho' (peg) superordenado se vai 'dependurar' o conceito (65c).

De acordo com essa visão, conhecer uma palavra pressupõe um conhecimento que é parcial e difuso (Aitchinson 1987). Isso significa dizer que as pessoas não poderiam nunca saber tudo sobre uma palavra denotativamente, uma vez que os conjuntos de relações conceituais têm fronteiras e formas conotativas imprevisíveis. Assim, esse *saber* não seria quantificável.

O modelo de Hastrup (op. cit.), de acordo com Scott (op. cit.), também pressupõe uma visão de conhecimento lexical que é parcial. Entretanto, nota-se que, apesar de reconhecer estágios intermediários em uma escala de conhecimento, e justificá-los, quando o conhecimento ainda é parcial, o modelo também permite um conhecimento total, demonstrado pelo nível 4. A maior semelhança do modelo de Hastrup com o de Scott está no fato de ambos estarem centrados no significado, enquanto outros modelos que apresento a seguir, com exceção de Cronbach (1942), são mais abrangentes, incluindo outros aspectos. É interessante observar que, embora o modelo de Hastrup e de Scott sejam centrados no significado, os autores reconhecem (vide Seção 2.3 deste capítulo), em sua caracterização de contexto, vários níveis de conhecimento associados à palavra, níveis esses que extrapolam o significado, como se as duas coisas não estivessem relacionadas.

Um outro modelo de conhecimento lexical composto de um conjunto de componentes é o de Cronbach (op. cit.), distinguindo cinco tipos de comportamento envolvidos na compreensão de uma palavra. São eles:

Quadro 2.7 – Conhecimento lexical (Cronbach 1942)

-
1. Generalização: ser capaz de definir a palavra
 2. Aplicação: selecionar um uso apropriado para a palavra
 3. Ampliação de sentidos: lembrar os diferentes significados da palavra
 4. Precisão de significado: aplicar a palavra corretamente a todas as situações possíveis
 5. Disponibilidade: ser capaz de usar a palavra produtivamente
-

Embora o autor se refira à capacidade de definir a palavra, não especifica, contudo, como ela deve ser definida. A relação entre os diferentes aspectos não é suficientemente clara, principalmente entre a precisão de significado ou quarto aspecto e a disponibilidade ou quinto, ou ainda entre a aplicação ou segundo e a ampliação de sentidos ou terceiro. Embora mencione *uso produtivo* no quinto aspecto, não explicita qual é o seu conceito de uso produtivo e receptivo.

Enquanto medidas como as escalas de avaliação que apresentei acima possam ser consideradas apropriadas para alguns propósitos, como por exemplo, estimativas de extensão de vocabulário, não o são para outros, tais como quando se deseja entender a natureza do conhecimento de vocabulário do aprendiz ou do leitor e também sua habilidade de uso, principalmente em LE. Há, nessas situações, a necessidade de medidas mais qualitativas e funcionais.

Alguns trabalhos mais recentes têm reconhecido essas lacunas e tentado ampliar as abordagens usadas na área de avaliação de vocabulário, levando-se em conta a base teórica existente na área. O trabalho de Read (1987) é uma tentativa nesse sentido. Para isso, desenvolveu um procedimento de entrevista para ser usado com alunos aprendendo inglês com propósitos acadêmicos na Nova Zelândia. Trinta substantivos e trinta verbos de baixa frequência são selecionadas de uma lista de palavras não-técnicas para uso acadêmico (*University Word List* ou Xue & Nation 1984) e

apresentadas aos avaliados (nesse caso, professores de inglês não-nativos) durante duas entrevistas de meia hora cada uma, conduzidas por diferentes avaliadores. As palavras são apresentadas por escrito, e isoladamente, ou seja, sem contexto. A tarefa do avaliado é pronunciar a palavra e demonstrar sua familiaridade com ela em uma escala de três níveis: *consigo entender a palavra, não tenho certeza e desconheço a palavra*. O próximo passo é responder a uma série de questões focalizando os seguintes aspectos:

Quadro 2.8 – Conhecimento lexical (Read 1987)

1. Pronúncia da palavra
2. Explicação do significado
3. Identificação da área de estudo
4. Colocações e associações
5. Conhecimento de outras formas da palavra

As respostas foram classificadas em três categorias: analíticas, exemplos de uso da palavra ou colocações. Para palavras mais técnicas notou-se uma tendência a associá-las aos seus campos acadêmicos. Muitas confusões foram observadas entre palavras que têm sons e formas semelhantes, ou o que Laufer (1990) denomina *sinformes* (*synforms*).

Respostas na LM do avaliado, através de sinônimos, definições ou exemplos de uso é uma outra possibilidade, que poderia ter sido usada se a situação fosse de LE, uma vez que esse tipo de resposta exige do avaliador conhecimento da LM do avaliado. Imagino não ter sido usada no estudo de Read por se tratar de uma situação de L2. Entretanto, essa seria a opção que mais se aproxima da tarefa exigida pela leitura ou compreensão oral.

Apesar de alguns problemas observados, relacionados principalmente aos aspectos psicológicos da entrevista (o que responder ao avaliado, quando a resposta for errada, por exemplo) o procedimento mostrou-se válido, apesar de exigir algumas alterações. Uma delas consistiu na ampliação da entrevista, que passa a incluir, num segundo momento, também uma fase em que as palavras são definidas em contexto. Um aspecto, não explorado por Read (op. cit.), seria verificar até que ponto as respostas dadas para as palavras da lista coincidem com as respostas em que os avaliados têm realmente de fornecer suas definições, isto é, em que medida a auto-avaliação coincide com a avaliação do entrevistador. Essa informação seria importante para uma validação das escalas de verificação.

Embora o procedimento desenvolvido por Read (op. cit.) possa consumir muito tempo daqueles nele envolvidos, restringindo, conseqüentemente, sua aplicação, ele é positivo por introduzir uma dimensão qualitativa na avaliação do conhecimento lexical, dimensão essa que considera também sua profundidade ou qualidade (*width*), ampliando, portanto, as medidas existentes, que apenas consideravam sua extensão (*size* ou *breadth*), baseadas em um conceito de

conhecimento de palavra definido de forma estreita ou como formas equivalentes em duas línguas (*paired-associate*).

O conceito ou modelo que fundamenta o trabalho de Read parece ter sido proposto inicialmente por Richards (1976) (vide também Nation 1984 e Wallace 1982, para aplicações desse conceito no ensino), sendo identificado, na literatura, como um conceito *rico* de vocabulário. Embora composto de um conjunto de componentes ou aspectos, esse modelo difere do de Scott (op. cit.) e de Cronbach (op. cit.) por ser mais abrangente, apresentando um leque de componentes que envolve aspectos lingüísticos, psicolingüísticos e sociolingüísticos.

Esse conceito constitui uma base teórica promissora para uma ampliação do escopo do vocabulário para o ensino/avaliação e também pesquisa em LE. Digo promissora porque seus resultados ainda não se fizeram sentir. Investigações mais sistemáticas do conhecimento de vocabulário apresentado por leitores em LE, que exploram a natureza desse conhecimento através de dados sobre sua profundidade, não foram até o momento apresentadas. Os relatos que se tem sobre esse conhecimento, são, em geral, parciais e não sistemáticos, baseados na extensão do vocabulário de falantes nativos. Com relação a uma situação de LE, mais especificamente em inglês no Brasil, os dados são ainda mais escassos. Uma das poucas tentativas é apresentada por Scott (op. cit.). Buscando caracterizar a extensão do conhecimento de vocabulário de seus alunos universitários da área de odontologia, que identifica como falsos-principiantes, usando uma lista de palavras baseada em Carroll et alii 1971, o autor afirma que muitos estacionaram no nível de 100 palavras, enquanto a maioria parou no nível de 500 palavras.

O componentes de um conceito *rico* são definidos a partir de uma série de pressuposições, que apresento no Quadro 2.9 abaixo, sobre a competência lexical de um falante nativo definido em termos ideais, e que, segundo Richards (op. cit.) poderiam servir de parâmetro para se estimar que tipo de conhecimento e produto final de aprendizagem podem ser esperados de um aprendiz em LE. Entretanto, como Richards (op. cit.) mesmo afirma, sua expectativa não é de que toda essa informação seja traduzida diretamente em procedimentos de sala de aula, uma vez que é muito mais complexa do que se pode supor:

Uma consideração de parte do conhecimento que se pressupõe como parte de uma competência lexical é oferecido como uma referência para a determinação de objetivos para o ensino de vocabulário e para a avaliação de técnicas de ensino projetadas para a realização desses objetivos (Richards op. cit.:78) (66c).

Não fica claro porque Richards aborda o primeiro aspecto no quadro acima. Levando em conta outras considerações que faz no texto, depreende-se que o autor focaliza, nesse caso, a importância de um vocabulário amplo e de uma expansão contínua, e, como tal, refere-se à extensão

Quadro 2.9 – Pressuposições sobre o conhecimento lexical do falante nativo (Richards 1976)

1. O falante nativo de uma língua continua expandindo seu vocabulário na idade adulta, embora haja, comparativamente, pouco desenvolvimento da sintaxe
2. Conhecer uma palavra significa saber o grau de probabilidade de encontrar essa palavra na fala ou na escrita. Para muitas palavras nós também "sabemos" os tipos de palavras mais prováveis de se associarem a ela
3. Conhecer uma palavra implica em conhecer as limitações impostas ao seu uso de acordo com as variações da função e de situação
4. Conhecer uma palavra significa conhecer o comportamento sintático associado àquela palavra.
5. Conhecer uma palavra pressupõe conhecimento de sua forma subjacente e das derivações que podem ser feitas a partir dela
6. Conhecer uma palavra pressupõe conhecimento de uma rede de associações entre essa palavra e outras na língua
7. Conhecer uma palavra significa conhecer o seu valor semântico
8. Conhecer uma palavra significa conhecer muitos dos diferentes significados associados a ela

do vocabulário, enquanto as demais pressuposições/aspetos dizem respeito à profundidade desse vocabulário. O segundo aspecto está relacionado à frequência e à colocabilidade das palavras. Segundo o autor,

(...) dada uma lista de palavras, com exceção de substantivos concretos, um falante nativo é capaz de classificá-las em 'frequente', 'moderadamente frequente', 'não frequente' com uma grau de precisão semelhante ao de suas frequências reais (Richards op. cit.:79) (67c).

Além disso, esse falante nativo também é capaz reconhecer a probabilidade de ocorrência de certas palavras em combinação com outras palavras, ou seja, sua colocabilidade ou relações sintagmáticas. Usando um exemplo de Richards, posso dizer que a palavra *fruta* coloca-se com *madura, verde, gostosa* e assim por diante. O terceiro aspecto está relacionado às limitações impostas ao uso das palavras de acordo com as variações de função e situação, incluindo variação temporal, social, geográfica, da área ou modo do discurso. Em outras palavras, refere-se às características de registro. O quarto componente, por sua vez, significa que, quando se aprende uma palavra, não se aprende apenas seu conceito, mas também suas propriedades gramaticais e estruturais específicas. De todos os aspectos, esse é o que mais contraria a divisão tradicional entre vocabulário e gramática, ao qual retornarei mais adiante. O quinto aspecto não diz respeito apenas às derivações mas também às flexões, salientando *"...regras que permitem construir diferentes formas de uma palavra ou mesmo diferentes palavras"* (Richards op. cit.:81) (68c). O sexto componente relaciona-se às redes de associações ou relações paradigmáticas entre as palavras de uma língua. Por exemplo, *acidente* pode se associar com *carro, ambulância, hospital* e tantas outras palavras. Os sétimos e oitavos componentes fazem referência ao significado. Enquanto o sétimo considera o denotativo, o oitavo considera outros aspectos que a palavra assume em contexto, ou conotativos.

Tendo explicitado minha interpretação dos componentes de Richards, reapresento-os, de modo sintetizado, no Quadro 2.10 a seguir, com o primeiro aspecto entre parênteses:

Quadro 2.10 – Conceito *rico* de vocabulário
(Richards 1976)

-
1. (Extensão)
 2. Frequência/colocações
 3. Registro
 4. Comportamento sintático
 5. Derivações e flexões
 6. Associações
 7. Significado denotativo
 8. Significado conotativo
-

O modelo de Nation (1984) assemelha-se ao de Richards, podendo ser considerado uma ampliação. Seu conjunto de aspectos é organizado no Quadro 2.11 abaixo, através de questões-chaves. Conhecer uma palavra significa ser capaz de responder a cada uma das questões. Estas subdividem-se de acordo com quatro critérios: forma, posição, função e significado. O quadro incorpora a distinção receptivo/produtivo, identificada pela abreviação R/P, não mencionada por Richards (op. cit.) e levemente tocada por Cronbach, em cada uma das quatro categorias de questões. Dessa forma, esse modelo pode ser considerado um conjunto de componentes com um contínuo receptivo/produtivo em cada um deles.

Como nos modelos anteriores, alguns aspectos não parecem ficar muito claros também no modelo de Nation (op. cit.). Por exemplo, no item *significado para uso receptivo*, não está explícito se o autor se refere a um significado ou aos vários diferentes significados de uma mesma palavra.

Embora um conceito *rico* de vocabulário tenha sido usado, na literatura, muitas vezes para caracterizar a competência lexical do falante nativo, observa-se que ele ainda é um conceito "estático", na medida em que não faz referência à uma dimensão ou habilidade/capacidade de uso. Embora incorpore, como parte do conhecimento lexical, vários outros níveis que na lingüística tradicional, nos outros modelos acima examinados e também na LA, mais especificamente no ensino/avaliação de LE, são considerados separadamente, ou como parte da gramática (flexões/derivações, relações gramaticais), embora faça menção à regras de uso (variações de função e situação, incluindo variação temporal, social, geográfica, da área ou modo do discurso ou aspecto 3, ainda está restrito ao aspecto semântico (denotativo/conotativo), não mencionando um nível pragmático-discursivo, ou como o conhecimento nos níveis mencionados é efetivamente atualizado na construção da compreensão ou negociação do significado, seguindo a visão apresentada em um modelo interativo, conforme proposto por Cavalcanti (1983, 1989) e Moita Lopes (1986).

Uma consideração mais detalhada do conceito de competência lexical, em oposição a um conceito de conhecimento, pode ser elucidativa neste ponto da discussão.

Quadro 2.11 – Conhecimento da palavra (Nation 1984)

FORMA	1. Falada	R: Qual é o som da palavra? P: Como a palavra é pronunciada?
	2. Escrita	R: Qual é a forma da palavra? P: Como a palavra é escrita e soletrada?
POSIÇÃO	1. Contextos gramaticais	R: Em que contextos a palavra ocorre? P: Em que contextos devemos usar a palavra?
	2. Colocações	R: Que outras palavras podem ser usadas antes ou depois da palavra? P: Que outras palavras podem ser usadas com essa palavra?
FUNÇÃO	1. Frequência	R: Quão frequente é a palavra? P: Quão é a frequência de uso da palavra?
	2. Adequação	R: Onde se esperaria encontrar a palavra? P: Onde a palavra pode ser usada?
SIGNIFICADO	1. Conceito	R: O que a palavra significa? P: Que palavra deve ser usada para expressar esse significado?
	2. Associação	R: Que outras palavras estão associadas a essa palavra? P: Que outras palavras poderiam ser usadas ao invés desta palavra?

Uma consideração mais detalhada do conceito de competência lexical, em oposição a um conceito de conhecimento, pode ser elucidativa neste ponto da discussão.

Competência tem sido um conceito controvertido, interpretado de formas diferentes na Lingüística e principalmente na LA. Originalmente proposto por Chomsky (1957, 1965), refere-se ao "... conhecimento de um sistema de regras, de parâmetros ou princípios, de configurações na mente para o que a linguagem apenas serve de evidência ..." (Widdowson 1989:129) (69c), algo, portanto, muito mais abstrato do que a linguagem. Para Hymes (1972), cujas preocupações estão voltadas para a linguagem propriamente dita, competência é a capacidade de usar a língua, introduzindo, portanto, o termo *competência comunicativa*, que marca uma dissidência a Chomsky ampliando seu conceito em dois aspectos: além de incluir conhecimento de outros aspectos da linguagem, que extrapolam a gramática ou sistema de regras da língua (conhecimento do que é possível, viável) também inclui a capacidade de uso, ou uma dimensão de desempenho, ou de comunicação real (vide artigos de *Applied Linguistics*, volume 10/2, 1989) para uma discussão mais aprofundada desse conceito; vide também Patrocínio (1993) para uma crítica ao conceito na perspectiva do ensino de português-LE).

Entretanto, é necessário salientar que ao propor esse conceito, Hymes não estava preocupado com o ensino, e muito menos, com L2 ou LE. A adaptação desse conceito para um contexto de ensino de L2, ou sua operacionalização e utilização enquanto arcabouço teórico para o ensino/avaliação de L2 foi feita primeiramente por Canale & Swain (1980), retomada a seguir por Canale (1983). Para esse autor, esse quadro teórico tem quatro componentes, cada um com duas dimensões (conhecimento e habilidade de uso): competência gramatical (regras de linguagem, vocabulário, formação de palavras, dentre outros, ou seja, o conhecimento e habilidades necessários para o entendimento e expressão apropriados do sentido literal das frases); competência sociolingüística (regras de uso sócio-culturais que levam em conta fatores tais como status dos participantes, propósitos e normas da interação); competência discursiva, envolvendo o domínio de formas gramaticais e significados para realizar um texto oral ou escrito em seus diferentes gêneros (coesão e coerência); e, finalmente, competência estratégica (estratégias verbais e não verbais usadas para compensar possíveis falhas na comunicação ou realizá-la de forma mais efetiva).

Tanto para Hymes (op. cit.) como para Canale (op. cit.) o conceito inclui duas dimensões, uma de conhecimento e outra de uso. Entretanto, enquanto para o primeiro a dimensão de uso é vista como capacidade (*ability*), para o segundo é vista enquanto habilidade (*skill*), o que pode ser considerado um estreitamento do conceito, pois, enquanto habilidade parece pressupor algo que se desenvolve através de um treinamento, comportando-se de maneira semelhante ou estável nas várias situações, ou o uso de formas rotinizadas, e que poderia ser associada a uma visão behaviorista (vide nota 1 deste capítulo), capacidade parece incluir também a adaptação a novas situações e poderia estar associada a uma visão mais cognitivista. Nesse sentido, considero mais dinâmica a conceitualização de Hymes.

Uma outra diferença, entre os conceitos propostos pelos dois autores, diz respeito à distinção competência vs. desempenho. O conceito de Hymes parece aproximar esses dois polos da dicotomia, uma vez que menciona, além das regras de gramática, as regras de uso (ou um desempenho virtual ou competência), assim como capacidade de uso (desempenho real). Em outras palavras, isso significa incorporar ao conceito de competência também um conceito de desempenho (real). Entretanto, Canale parece retomar a dicotomia, dado que descreve a competência comunicativa (equivalente a um desempenho virtual) como parte de uma competência real (desempenho real). Ao invés de incorporar o conceito de desempenho real, ele o coloca em oposição ao conceito de competência comunicativa. Mais uma vez minha opção é pela conceitualização de Hymes.

Uma competência lexical definida nesses termos, portanto, significa incorporar, em um mesmo conceito, os vários aspectos especificados em um conceito *rico*, ou seja, os níveis lexical, sintático, morfológico, semântico, ao qual adiciono um nível discursivo/pragmático, que inclui o conceito de capacidade de uso de Moita Lopes (1986), que permitiria a realização ou atualização dos

outros níveis na negociação dos sentidos do texto. Entretanto, também significa incorporar uma dimensão de uso, incluindo, em um mesmo conceito, também uma dimensão cognitiva, ao invés de considerá-la um componente separado, levando em conta a interação entre esses dois componentes. Essa dimensão de uso, entretanto, também envolve questões de processamento. Nesse sentido, devo salientar que são várias as teorias que enfatizam o processamento de vocabulário (vide a *Hipótese do Gargalo* ou *Bottleneck Hypothesis* em Perfetti 1986a, 1986b) ou uma competência lingüística que vai além do conhecimento lingüístico. A compreensão, nesse caso, não é afetada apenas por um conhecimento insuficiente de vocabulário mas também por uma dificuldade em processamento, que exige, portanto, decodificação ou reconhecimento de palavras, que nesse caso, deverá ser rápido, automático e preciso. Nesse sentido, Beck et alii (1982:148) salientam que

(...) para que a aprendizagem de vocabulário tenha um efeito positivo na compreensão, esse aprendiz não requer apenas precisão de significado, mas também fluência de acesso a significados na memória (...) ele precisa ir além de simplesmente respostas precisas a palavras. Ele deve desenvolver fluência de acesso (fluency of access) ao significado de palavras, integrando as palavras do aprendiz às redes semânticas existentes (70c).

Além desse processamento *automático*, que parece caracterizar uma competência lexical bem desenvolvida, Shiffrin & Schneider (1977), embora se referindo ao processamento cognitivo ou da informação em geral, identificam outra modalidade, que denominam de *controlado*. Essas duas modalidades de processamento baseiam-se em vários experimentos cujos resultados levaram os autores a elaborar uma *Teoria de processamento de informação*. A percepção automática da informação, para esses autores, parece refletir um processo de longo-termo, algo bem estabelecido na memória, virtualmente não afetado pelo volume de informação. Por outro lado, a percepção controlada mostra uma dependência à capacidade de atenção e à quantidade de informação.

Um conceito de competência lexical, visto dessa maneira, portanto, é um conceito mais abrangente, que incorpora o componente lingüístico e cognitivo, a gramática (sintaxe, morfologia) e o léxico, mais compatível com a visão de contexto e de o conceito de interação em leitura conforme o quadro teórico proposto por Cavalcanti e Moita Lopes. Essa visão consistiria, portanto, no uso de outros níveis a serviço do léxico, que apenas seria o ponto de partida para resolver problemas que podem não se restringir ao nível lexical. Essa competência lexical mais desenvolvida tornaria, portanto, possível a exploração das várias pistas que, por exemplo, contextos *potencialmente ricos*, oferecem, na efetiva construção de contextos ricos.

Uma competência lexical envolvendo duas dimensões, portanto, de conhecimento e uso será, orientadora desta investigação, fundamentando os testes de vocabulário do estudo de produto

e a análise do estudo de processo. Enquanto o foco do primeiro estudo ou de produto será o conhecimento, o do segundo ou de processo será o conhecimento/uso. Especificamente para o segundo estudo, proponho uma categorização do que denomino *recursos bilaterais do leitor*, definidos enquanto a inter-relação entre as fontes de conhecimento ou níveis propostos nesse conceito de competência e as restrições do texto e do contexto de leitura, que permitiram construir o que é denominado contexto do texto e contexto geral, para reconhecimento de palavras e construção de um sentido para o texto. A denominação *recurso* é, a meu ver, mais consistente com uma visão de contexto na perspectiva de um modelo interativo, uma vez que não enfatiza nem o leitor nem o texto, mas sua inter-relação. Como já mostrei anteriormente neste capítulo, as categorizações de contexto propostas na literatura variam, não apenas com relação às categorias apresentadas mas principalmente ao seu foco. Poder-se-ia dizer que a escolha das categorias é determinada pelo foco. A de Haarstrup (1987, 1989), apresentando *fontes de conhecimento*, parece enfatizar o contexto sob o ponto de vista do leitor; a de Scott (1992) por sua vez, focalizando os *vários níveis de co-texto* (palavras imediatamente vizinhas à palavra a ser inferida, a oração, o parágrafo, dentre outros), embora também se refira ao contexto geral que inclui conhecimento prévio (*background*), parece se concentrar no contexto sob a perspectiva do texto, ou ainda, das restrições ou pistas. Van Parreren & Shouten-van Parreren (1981) referem-se aos níveis (estilístico, lexical, semântico e sintático), enfatizando, mais uma vez, o conhecimento do leitor. É também "dinâmico", mais bem definido como um processo, constantemente reformulado, não enfatiza apenas o conhecimento mas também os procedimentos de uso, sendo portanto, mais coerente com um conceito de competência e não apenas de conhecimento.

Embora a base para essa categorização seja a competência definida em termos de um conceito *rico*, ela é ampliada para incluir também o conhecimento prévio ou de mundo (conhecimento conceitual ou do assunto, e ainda o conhecimento sócio-histórico, cultural, crenças, valores ou conhecimento esquemático). Ao invés de considerá-lo como um nível separado, como em outras categorizações e taxonomias, ele é incorporado aos vários níveis. Embora muitas vezes não explicitado, ele aparece subjacente a todos eles. Por sua vez, essa competência lexical é vista como parte de uma competência mais ampla, que também envolve conhecimentos textuais e superestruturais.

Os recursos que apresento a seguir podem ser intralingüísticos, quando as relações estabelecidas envolvem conhecimentos apenas na LE, e interlingüísticos quando envolvem a LM e outras línguas:

- Lexicais, relacionados à forma (ou ortografia) e ao som (ou aspectos fonéticos) da palavra; nesta categoria ainda podem ser considerados aspectos morfológicos (especialmente o prefixo e radical) uma vez que esses elementos, usados para acessar o conhecimento semântico da palavra, envolvendo seu conceito e vários significados, levam diretamente ao significado.

- Sintáticos, resultado da interação das restrições do contexto imediato da palavra ou dentro da oração, ou ainda entre orações ou porções maiores do texto, levando em conta a posição da palavra e sua função, os seus padrões de comportamento sintático na oração, classes gramaticais e morfologia, principalmente os sufixos. Enquanto os prefixos e radicais possibilitam um acesso direto ao significado, sufixos, através da indicação da classe gramatical das palavras, permitem o acesso ao significado de maneira indireta, ou seja, através de relações gramaticais ou sintáticas;
- Semânticos, relacionados ao significado, não apenas envolvendo a palavra desconhecida, mas outras palavras dentro da oração, relações intersentenciais ou ainda porções maiores de discurso; inclui o significado da palavra, seu conceito, assim como seus vários significados, que formam redes de associação; associações sintagmáticas ou de colocações, e sintagmáticas (de contraste ou antonímia, similaridade ou sinonímia, classificação subordinativa, classificação coordenativa, classificação superordenada). Nesta categoria, ainda podem ser considerados aspectos como a adequação/frequência das palavras em relação à outras para uma previsão de seu significado. Os recursos de adequação referem-se às variações de uso da palavra em registros diferentes, contextos sociais e temporais, área do discurso (discurso científico vs. geral, por exemplo), dentre outros, delimitados pelas restrições do texto;
- Conhecimento acumulado em vários níveis, ou seja, de oração, sentença, parágrafo, do texto como um todo. Esse conhecimento pode envolver todos os outros recursos acima;
- Textuais, que se utilizam da organização ou macro-estrutura do texto ou de sua estrutura retórica, subtítulo, título;
- Superestruturais ou referentes à estrutura canônica de cada tipo de discurso;
- Conhecimento prévio do assunto, de mundo, do leitor, do tipo de revista ou livro onde o texto está inserido, embora ele possa ser visto como fazendo parte de todos os níveis;
- Gráficos e tipográficos tais como letras maiúsculas, grifos, aspas.

Foi o objetivo deste capítulo construir uma base teórica para esta investigação, através de uma revisão crítica da literatura, da explicitação dos modelos de leitura/competência lexical que fundamentam este estudo, assim como de seus conceitos chaves e limitações teóricas e metodológicas. Essas limitações são o ponto de partida para o próximo capítulo, que discute a metodologia e o desenho de pesquisa desta investigação.

Notas

(1) A divisão dos objetivos de cursos de línguas em quatro habilidades (*skills*) tem sido criticada por Widdowson (1991:96), dentre outros, por pressupor uma visão segmentada e behaviorista de aprendizagem de línguas, referindo-se "... à maneira pela qual o sistema da língua se manifesta, ou é percebido se

manifestar, como forma". Apesar dessa crítica, são numerosos os cursos de língua incluindo o contexto desta pesquisa) que focalizam apenas a leitura. Uma justificativa para tais cursos estaria na necessidade dos aprendizes nessas situações.

(2) A distinção entre L2 e LE, muitas vezes ignorada na literatura, é feita neste trabalho, por reconhecer as diferentes condições de aquisição de leitura que caracterizam esses dois contextos.

(3) Essa posição parece coincidir com a de Moita Lopes (1986:118) quando afirma que "*... os objetivos da leitura em L1 e LE ... devem ser os mesmos. Entretanto, com respeito aos processos reais de aquisição desses objetivos, a leitura em L1 e L2 pode diferir*" (71c).

(4) As citações que aparecem no corpo do texto foram traduzidas. As originais se encontram depois das notas, no final de cada capítulo, identificadas por números seguidos da letra *c* (1c, 2c, e assim por diante). O mesmo procedimento é usado em todos os capítulos.

(5) O termo *interpretação pragmática* neste contexto seria o extremo da visão de pragmática de Cavalcanti (op. cit.) ou de Moita Lopes (1989), definida a partir da negociação do sentido no texto, e relacionada, portanto, com uma interpretação semântica ou apoiada em elementos de decodificação (vide nota 15). Para a autora, a direcionalidade de âmbito, definida como a propriedade dos itens-chaves de ativar estruturas de conhecimento ou sistemas de valores do leitor, pode ser semântica, pragmática ou ainda semântico/pragmática, e, portanto, definida em termos de um contínuo. A denominação *interpretação pragmática* para se referir às visões descendentes poderia ser explicada através da predominância, nesse caso, de uma direcionalidade pragmática (Cavalcanti, 1983, 1989).

(6) Note-se a recorrência do termo *mensagem* proveniente da Teoria da Comunicação.

(7) Essa visão também tem sido interpretada como *construção do significado*.

(8) Embora situações onde a leitura é ensinada independentemente das outras "habilidades" sejam bastante comuns, e possam ser justificadas principalmente em contextos de LE (vide nota 1) observa-se hoje uma tendência de ensino mais integradora, que tenta buscar um equilíbrio entre o desenvolvimento da expressão oral e da leitura.

(9) A não explicitação da importância do componente linguístico no ensino comunicativo (de certa forma sujeito às mesmas condições ou ao mesmo momento histórico) pode ser responsável por muitas das interpretações radicais que tendem a eliminá-lo totalmente da sala de aula.

(10) O levantamento de hipóteses ou de possíveis explicações é consistente com a natureza exploratória e descritiva deste trabalho, constituindo uma de suas maiores contribuições. Generalizações, contudo, não devem ocorrer sem a prévia confirmação dessas hipóteses.

(11) Um terceiro uso para o termo (ainda exploratório) está sendo desenvolvido em estudos de Biber (1984) e Grabe (1988) e diz respeito à *interação textual* entre as modalidades escrita e oral.

(12) Uma outra proposta para a questão da adequação de leituras ou interpretações e que também se baseia na interação é a do *contrato*, de Tierney & Lazansky (1980). Elaborada a partir de teorias de atos de fala ou do conceito de intencionalidade e do Princípio Cooperativo de Grice (1975) essa teoria é uma tentativa de conciliar os dois polos de tensão no texto, ou seja, o determinismo da convenção e o indeterminismo da intenção, pressupondo um acordo contratual entre autor e leitor, que lhes dá direitos e deveres. Isso implica, para o leitor, em limites em sua liberdade interpretativa e para o autor, o dever de respeitar o leitor, sendo informativo, sincero, relevante e claro. Apesar de suas contribuições, principalmente no que diz respeito às intenções comunicativas e tipologia textual, a proposta não considera o papel da gramática e a estrutura textual (vide Beaugrande 1978 para considerações sobre tipologia textual como um fator determinante da compreensão na leitura).

(13) Essa posição poderia ser coincidente com o que Dascal (1981), sob o ponto de vista da linguística pragmática, denomina *contextualismo moderado*. Para esse autor, essa visão, que consiste em modificar a noção de significado literal, sem excluí-la, poderia ser uma alternativa entre dois extremos, o literalismo radical e o contextualismo. "*A força ilocucionária de um ato de discurso está relacionada com a frase*

enunciada só provisoriamente. Algumas frases são muito explícitas em suas especificação dos atos de discurso que servem para desempenhar, enquanto outras oferecem umas sugestões muito gerais com relação aos possíveis tipos de atos de discurso desempenhados por seu uso. Tais sugestões têm que ser completadas por informação contextual com o fim de produzir uma determinação da força ilocucionária" (72c).

(14) Parece haver um ponto de convergência entre posições, apesar de Coracini et alii (1988) considerar a questão sob o ponto de vista da linha francesa da Análise do Discurso.

(15) A tendência dessa literatura é dar centralidade ao *Princípio de Cooperação de Grice* (op.cit.). O *Princípio de Relevância* proposto origina-se da máxima de relação transformada em princípio (Sperber & Wilson op. cit.). A noção de relação parece depender da relevância e perceber uma relação depende da relevância percebida (Sperber & Wilson op. cit.; Cavalcanti op. cit.).

(16) A delimitação da semântica e da pragmática tem gerado controvérsias. Adoto, nesse sentido, a visão de Leech (1980) recuperada em Cavalcanti (1983, 1989) que considera a semântica e a pragmática como níveis diferentes mas complementares dentro da lingüística. Para Leech (op. cit.:2), que se define como representante da posição *funcionalismo formalista*, a pragmática é "...o estudo da relação entre o significado ou o sentido da expressão lingüística [ou seu sentido em abstração ou semântico] e a força comunicativa que ela tem para falantes e ouvintes em dadas situações de enunciado" (73c). Enquanto a tarefa da semântica é preocupar-se com a compreensão, a da pragmática é a interpretação do sentido, ou a explicação da relação entre "... dois tipos de significado: o sentido (que freqüentemente é descrito como 'literal') e a força ilocucionária" (Leech 1983:35) (74c). Entretanto, outras tendências podem ser identificadas: a pragmática englobando a semântica (Searle 1969); a semântica englobando a pragmática (Lyons 1977, 1981); ou ainda as duas como áreas distintas (Grice 1968). Para Mey (1985:17) "...a lingüística tem que ser pragmática; os usuários da língua e não os lingüistas, são as pessoas mais importantes na lingüística" (75c). Entretanto, a pragmática de Mey parece se distinguir da pragmática proposta por esses outros autores, estando mais próxima da análise crítica do discurso (vide nota 19). O autor critica a falta de dimensão humana na lingüística, que, na sua pretensão de se tornar uma ciência, objetiva e portanto "neutra", acaba por se distanciar da sociedade e do indivíduo. Entretanto, essa neutralidade da lingüística é apenas aparente e ilusória, na medida em que opressões e manipulações são conseguidas a partir não apenas de seu conteúdo mas também a partir de sua forma (correção). O veículo, no caso do conteúdo, é o léxico, usado, embora de forma indireta, para estabelecer a falsa idéia de "conceitos universalmente aceitáveis (as entradas de dicionário) ... os conceitos, como todas as coisas criadas pelo homem, são entidades históricas e, como tais, intimamente associadas à organização de uma sociedade em uma dada época" (Mey, op. cit.:36) (76c). Entretanto, não trazem consigo "instruções de uso" e, portanto assumem significados diferentes daqueles originais através dos contextos de uso (relexicalização). Daí a impossibilidade de se pensar no conhecimento em geral e também lingüístico, como algo que possa ser distanciado de seu contexto de uso, que possa ser considerado em abstração, ou algo como a semântica. Para esse autor, portanto, a lingüística (ou a pragmática) é dinâmica, e deve estudar a língua juntamente com as transformações da sociedade ou com suas categorias políticas, econômicas e sociais. A posição de Mey poderia se aproximar do que Leech denomina de *funcionalismo extremo*.

(17) A concordância intersubjetiva na leitura, para Koch e Travaglia (op.cit.), é explicada, como na teoria do contrato (vide nota 10), pela *intencionalidade* e *aceitabilidade*. Koch (1989:80), referindo-se a Koch (1984) salienta que "... quem produz um texto tem determinadas intenções, consistindo a intelectualidade na captação dessas intenções, o que leva a prever, por conseguinte, uma pluralidade de interpretações ... a aceitabilidade diz respeito à atitude dos receptores de aceitarem a manifestação lingüística como um texto coesivo e coerente, que tenha para eles alguma aceitabilidade". Para Cavalcanti, por outro lado, a concordância intersubjetiva seria explicada através do estabelecimento de um elo entre a relevância-autor, sinalizada por itens lexicais chaves e a relevância-leitor (sinalizada por itens contextualmente relevantes).

(18) Na leitura, o componente lingüístico é conhecido como *primeiro componente*, enquanto a proficiência ou fluência (*reading abilities*) constitui o *segundo componente*.

(19) A análise de discurso crítica (vide Fairclough 1985) distingue-se da análise de discurso descritiva na medida em que esta tem objetivos que não são exploratórios, ou que o são apenas dentro da situação imediata e não nos níveis acima da instituição social ou formação social, ou seja, como o nome diz, descreve sem explicar.

(20) Parte da afirmação de Van Dijk e Kintsch (1983:23-24) reproduzida a seguir (vide página 42 para citação completa) sugere essa visão de contexto. "... o que realmente está em questão são a velocidade e a precisão de operações de reconhecimento de palavras independentemente de contexto". O comentário de Eskey & Grabe (1988:224) nesse sentido é oportuno: "[Falar em 'reconhecimento de palavras sem necessidade de contexto'] é simplificar demais. A noção de 'contexto' é um assunto complexo. O reconhecimento automático ... na realidade indicam algum tipo de processamento contextual. Não é, no entanto, o mesmo tipo de contexto que a maioria dos pesquisadores se referem quando falam da importância do contexto" (77c).

(21) Parece haver uma incoerência quando se compara a caracterização de contexto de Haastруп (1989a) com seu modelo de conhecimento lexical (vide Seção 2.3 deste capítulo).

(22) Leffa (1984) através de instrumentos metacognitivos, concluiu que para estudantes universitários brasileiros, alunos de Inglês Instrumental da Universidade Federal do Rio Grande do Sul, a competência sintática é mais significativa para a compreensão de textos em inglês do que uma habilidade de leitura em LM. Em Maciel (1989), a questão da ênfase na proficiência lingüística vs. proficiência em leitura é retomada. Nesse artigo a autora especula sobre a necessidade de se considerar os dois componentes, sem, entretanto, apresentar dados. Em um artigo intitulado "Gramatiquice ou estrategice: é esta a questão?" Barbosa (1990) recoloca a questão. A autora salienta a tensão existente entre os dois componentes, mostrando a inexistência de estudos sobre sua distribuição ou dosagem nos cursos de Inglês Instrumental no Brasil. Através de um estudo piloto, propõe-se a investigar, dentre os itens privilegiados pelo ensino instrumental, a existência de uma tendência a se enfatizar o componente lingüístico ou estratégias de leitura. Entretanto, dado o escopo limitado do estudo, os resultados foram inconclusivos.

(23) A tradução de *reading abilities* por *habilidades de leitura* parece-me problemática, podendo ser confundida com *reading skills*. Tenho me referido, portanto, ao componente cognitivo da leitura como *proficiência ou fluência*, embora esses termos também sejam, confusos, pois estão relacionados não somente à competência cognitiva mas à competência em geral. *Capacidade* seria ainda outra possibilidade.

(24) A partir daí, vários níveis limiares ou levantamentos de necessidades comunicativas e vocabulário básicos têm sido propostos para várias línguas na Europa (vide Malaca Casteleiro et alii (1988) para o português).

(25) A popularidade das estratégias de leitura também se deve à tendência de se enfatizar o processo da leitura ou de construção do sentido e da compreensão. Como salienta Carrell (1989:121), essa pesquisa tem-se concentrado no uso de estratégias, porque elas "... são de interesse pelo que revelam sobre a maneira pela qual os leitores interagem com o texto escrito e como essas estratégias estão relacionadas à compreensão do texto" (78c).

(26) Neste caso, o critério para seleção de palavras é a frequência, diferindo, portanto, daquele usado nos níveis limiares, isto é, a análise das necessidades comunicativas do aprendiz.

(27) Uma associação entre estratégias de inferência/conhecimento *procedimental* (*procedural*) pode ser feita. A dicotomia conhecimento *procedimental* vs. *declarativo* parece ter-se originado na teoria de aquisição de habilidades proposta por Anderson (1982). A partir daí, tem sido usada por inúmeros autores (vide Faerch & Kasper 1985; Haastруп 1987; Robinson 1989). Para seu proponente, a aquisição de uma habilidade passa por dois estágios: *declarativo*, em que os fatos são interpretados, e *procedimental*, em que o conhecimento é diretamente incorporado em procedimentos para o desempenho dessa habilidade.

(28) Embora os termos *adivinhar* (*guess*) e *inferir* (*infer*) refiram-se ao mesmo processo, alguns autores (vide Scott 1990) preferem distingui-los, usando o primeiro para *adivinhação pura* (*wild guessing*) e o último para o uso quando há alguma hipótese ou evidência, seja ela bem sucedida ou não. Neste trabalho, sigo a sugestão de Scott (op. cit.). Para não confundí-la com o outro tipo de inferência feita na leitura uso o termo *inferência lexical*.

Citações originais

(1c) "... such models are, in general, models of the 'ideal', completely fluent reader with completely developed knowledge systems and skills; whereas the second language reader is, almost by definition, a developing reader with gaps and limitations in both these categories".

(2c) "Podría decirse que una grande parte de la historia de la filosofía del lenguaje, en este siglo, ha consistido en el descubrimiento progresivo de las dificultades inherentes a dicha concepción. Entre ellas, las limitaciones del principio de la concepción de significados, la importancia del contexto en la interpretación de la frase, el dilema de trazar con precisión la línea divisoria entre diccionario y enciclopedia, o entre lo analítico y lo sintético, y la insuficiencia crónica del 'significado literal' como medio de determinar las condiciones de verdad y las propiedades lógicas de las frases".

(3c) "... the reader-based model has almost totally disregarded the printed text ... the reader interprets the text only pragmatically and ignores its semantic interpretation".

(4c) "... the reader reconstructs ... the message that was encoded by the writer through a graphic display".

(5c) "This view of reading should not be considered as an alternate version of 'reading as an interactive process'. The issue is not the relation of the reader to the text but the processing relations among various component skills in reading".

(6c) "... interactive refers to the interaction between information obtained by means of bottom-up decoding and information provided by means of top-down analysis, both of which depend on certain kinds of prior knowledge and certain kinds of information-processing skills".

(7c) "... if there is a deficiency at an early print-analysis stage, higher-order knowledge structures will attempt to compensate. For the poor reader, who may be inaccurate and slow at word recognition, but who has knowledge of the topic, top-down processing may allow for this compensation ... On the other hand, if the reader is skilled at word recognition but does not know much about the text topic, it may be easier to simply recognize the words on the page and rely on bottom-up processes".

(8c) "... places discourse analysis in the position of ('uncritically') reproducing certain ideological effects".

(9c) "Knowledge' implies facts to be known, facts coded in propositions which are straightforwardly and transparently related to them. But 'ideology'... involves the representation of 'the world' from the perspective of a particular interest, so that the relationship between proposition and fact is not transparent, but mediated by representational activity. So ideology cannot be reduced to 'knowledge' without distortion".

(10c) "The 'competent' subject of cognitive conceptions of interaction is the autonomous subject of ideology".

(11c) "Finding a shortcut method to identify prior knowledge clearly remains a problem. In this study the contradictions between topic familiarity ratings and actual recall performance raise the question of whether a subject's simple self-reporting that topic A is more familiar to him than topic B was indeed accurate. One reason for the lack of significance for the familiarity variable, therefore, may be inaccuracies in such reports".

(12c) "The theme of prior knowledge and its role in comprehension is repeated in many discussions of language proficiency. Most current theories of foreign language proficiency incorporate prior knowledge within the realm of 'context' as either influencing proficiency or as a part of the very construct of proficiency ... Current language theories imply that more proficient language users rely less on prior knowledge of topic and rely more on knowledge of elements of language. Proficiency is defined, in part, as a measure of context dependency in measuring comprehension".

(13c) "The important corollary is that if the learners cited were able to guess a majority of the words, then

the necessary clues are there for other, perhaps less gifted, learners to use. It is not unrealistic goal to expect learners to guess between 60 per cent and 80 per cent of the unknown words in a text if the density of the unknown words is not too high".

(14c) "... although comprehension of a text can be predicted to a significant extent by the amount of exposure one has of the topic domain of the text and the grasp of the language of the text, it is really the latter that determines how much the foreign language reader can extract and interpret the meaning of the text".

(15c) "... it is their linguistic competence of the target language that predicts their comprehension success rather than prior knowledge of the subject matter."

(16c) "... pupils who have attained the threshold may perform very differently on cognitive and academic tasks from pupils who failed to attain the threshold".

(17c) "... there seems to be a 'threshold effect' by which, when more than approximately 50 per 1,000 words are unknown, perception of overall structures may effectively be blocked, which in turn means that there is not enough in the way of context to allow successful guessing ...".

(18c) "Our experience shows that learners with a vocabulary of around 3,000 words are capable of guessing, on average, 60-70% of the unknown words in a reading passage. Some learners regularly achieve around 80%".

(19c) "There must be a 'bump' in the line, just as there must be a doorway and some difference between 'outside' and 'inside' for one to be at a threshold ...there will not only be a clear difference between subjects on either side of the point labelled 'threshold', but also a noticeable alteration in the line of the graph. The position of the bump itself will indicate where qualitative change occurs".

(20c) "Such a usage merely imply that subjects at the left part of the x axis (horizontal) perform less well than subjects further right, and that the difference is worth labelling. This is all that is claimed, for example, by Haarstrup".

(21c) "... as part of a larger interacting network of information needed by the L2 reader. It is necessary but not, by itself, sufficient for successful L2 reading".

(22c) "It is necessary to look at what second component factors are involved in the nontransference of good reading behaviors and how these interact with different levels of a first component linguistic ceiling to form, in effect, a combined linguistic and psycholinguistic ceiling".

(23c) "... while the results from Cziko and Clark would indicate that basic language competence is involved in reading comprehension, the fact that good L1 readers maintained an advantage over poor L1 readers at their proficiency level when both were reading L2 indicates that some nonlanguage skill operates. That is, it seems likely that reading skills per se, second component factors, affect the degree to which the proficiency ceiling restricts comprehension, just as the proficiency ceiling may restrict the degree to which good reading skills are applied. In this symbiosis, either component can affect the degree to which the other affects comprehension".

(24c) "Under certain conditions poor readers exhibit greater sensitivity to contextual constraints than do good readers. They do so in those circumstances where featural, orthographic and/or lexical knowledge ... sources are weak in comparison to syntactic and semantic knowledge. The reason good readers are sometimes less sensitive to contextual effects is that their knowledge sources are seldom weak ...".

(25c) "The greatest facilitation of word recognition by meaningful context is observed with poor readers, not with good readers. Furthermore, it is simply not true that good readers take decoding lightly; they fixate almost every content word It has been found over and over again that the best discriminator between good and poor readers in performance are simple letter and word identification tasks. What is really wrong with poor readers is that they recognize isolated words inaccurately and too slowly and compensate for lack in decoding skills with context, dependent guessing or hypothesis testing ... Good readers with their superior decoding skills can decode letters and words rapidly in a bottom-up fashion,

and therefore do not need to resort to guessing strategies - what is really at issue are the speed and accuracy of context free word recognition operations".

(26c) "... are mental operations involved when readers approach a text effectively and make sense of what they read".

(27c) "... the selection of a given strategy under a certain set of circumstances may create reading failure even though other strategies promoting comprehension are also in use".

(28c) "... if a reader is aware of what is needed to perform effectively, then it is possible to take steps to meet the demands of a reading situation more effectively. If, however, the reader is not aware of his or her own limitations as a reader or of the complexity of the task at hand, then the reader can hardly be expected to take preventive actions to anticipate or recover from problems".

(29c) "... system of assumptions through which experiences are organized ... schemata which govern behavior".

(30c) "... first, that a reader's theoretical orientations towards reading may determine, to some extent, the degree to which low proficiency in the language restricts second language reading ability and, second, that the models that readers hold may be of critical importance in allowing them to strike a successful balance between bottom-up (or data driven) and top-down (or conceptually driven) processing necessary for the interpretation of a text".

(31c) "It is often difficult for readers (and language users more generally) to articulate models of the mental processes; these mental operations may simply not be available for introspection. For L2 readers, such as those in this study, the difficulty may be compounded by lack of fluency in the second language. In addition, as is often noted, complex cognitive processes very probably involve a number of interacting strategies. Readers, for example, may employ one strategy for reading a novel, another for reading scientific or technical material, and a combination of strategies for reading a newspaper. To posit a single model, however intricate, may be to overlook this complexity".

(32c) "... behaviors used for centering, arranging, planning, and evaluating one's learning. These 'beyond-the cognitive strategies' are used to provide 'executive control' over the learning process".

(33c) "... refer to those mental processes that readers consciously choose to use in accomplishing readings tasks. Such strategies may contribute to successful comprehension or detract from it. In principle, what distinguishes strategies from other processes is the element of choice involved in their selection. In reality, the amount of attention that readers pay to their choice of strategies falls on a continuum from total attention to total lack of attention. If readers are requested to indicate the strategies that they use, it is likely that they would be able to describe even the ones that they are attending to the least, since these are, by definition, within the realm of conscious awareness. They would not, however, be able to describe certain unconscious reading processes".

(34c) "... have preferred to study grammar and phonology since these are close systems and therefore lend themselves to much more abstraction and generalization than vocabulary, which is not a closed system but an open set ... and since a good theory is the one that accounts for the largest number of events as simply as possible, this means that the theory of grammar is more powerful than the theory of lexis".

(35c) "... have been reacting against the earlier tendency of learning-theory-oriented linguists to concentrate on vocabulary learning, explainable as associative learning, rather than grammar acquisition which needs rules".

(36c) "As for the interaction between methodologists and language acquisition researchers, until recently the main concern was concerned to be in the beginning stages of language learning. It was assumed that at those stages it was more important to concentrate on grammar".

(37c) "In learning a second language, you will find that vocabulary is comparatively easy, in spite of the fact that it is vocabulary that students fear most. The harder part is ... structure ... the context and expression".

(38c) "... vocabulary was somewhat peripheral, an irritating irregularity in an otherwise ordered grammar".

(39c) "... without grammar very little can be conveyed, without vocabulary nothing can be conveyed".

(40c) "... the relationship between thought and words is a dynamic one: it is a process, not a thing".

(41c) "... most of the impetus to future research on vocabulary acquisition is likely to stem from some principles of communicative language teaching. There can be no fluency without a solid vocabulary base; no comprehensible, interesting and relevant input with poor lexis ... Learner-oriented approach to teaching and work on motivation gave rise to the analysis of learner's needs by both the teacher and the learner. Such an analysis is bound to show the need for better vocabulary learning ...".

(42c) "... to know a word so that it may be recognized, recalled at will, related to an object or a concept, correctly used, pronounced and spelt, appropriately collocated, used at the right level of formality and with awareness of its connotations and associations."

(43c) "... not a corpus of technical language but the ability to mobilize the resources of general English in the solving of technical problems. The 'procedural' or more general vocabulary helps the learner to get at the specific, technical vocabulary".

(44c) "...[the] research ... has been largely atheoretical and unsystematic. There are no clear theories of vocabulary acquisition ... and on the whole ... has avoided the serious theoretical questions that arise once one moves away from this very basic level. All these pieces are small-scale, however, and none of them adds up to a coherent and impressive body of knowledge at this stage."

(45c) "The effective teaching of vocabulary, especially at the earlier stages, is, to a considerable degree, dependent on satisfactory answers to this question ... how can the second language speaker mental lexicon be characterized and how it differs from the lexicon of a monolingual native speaker?".

(46c) "... attempts to justify the selection of vocabulary items for inclusion in courses and examinations on the basis of frequency counts and similar objective measures".

(47c) "Learning vocabulary is not just a matter of acquiring translation equivalents ... and completely ignore the complex patterns of meaning relationships that characterise a proper, fully formed lexicon, as opposed to a mere list ... some lexical structuring must go on when even the shortest word list is learned".

(48c) "... learning from context is still a default explanation; evidence that individuals actually learn word meanings from contextual experiences is notably lacking".

(49c) "... even the most ruthlessly systematic direct vocabulary instruction could neither account for a significant proportion of all the words children actually learn, nor cover more than a modest proportion of the words they will encounter in school reading materials".

(50c) "... to the working out of entailments, presuppositions and implicatures, using default suppositions and the like, procedures that are needed for all the 'reading between the lines' we do all the time".

(51c) "Typically, ESL learners are poor decoders since their vocabulary knowledge is weak while, at the same time, they are already literate in their mother tongue, and are familiar with top-down processing. Therefore, it becomes important to consider whether our instruction should emphasize top-down or bottom-up processing, as well as an appropriate emphasis on the use of context".

(52c) "This seems to indicate that even weak students were able to infer quite well, and that this helped them to produce slightly better summaries, though not as much as knowing the words in advance would have".

(53c) "... the subjects first language plays an extremely important role in his or her language learning and specifically in the lexical inference process".

(54c) *"No one pattern of strategy use was firmly associated with greater lexical inference success rates. Instead, subjects made great use of background knowledge and lexis-based strategies and little use of discursal or grammatical features".*

(55c) *"... no clear evidence that the hierarchy operates like a one-way street ... The fact that errors can be classified according to their scheme does not prove that a reader who mistakes the word class of an item, for example, cannot make a plausible estimate of its meaning".*

(56c) *"... whereas intake for comprehension is always facilitated by rich context, it seems likely that intake for learning is correspondingly weakened".*

(57c) *"In general, the research leaves us little doubt about the importance of vocabulary knowledge for reading and the value of reading as a means of increasing vocabulary. The precise nature of these relationships, and how we can make use of them in our teaching, are still fruitful areas of investigation".*

(58c) *"... the receptive control over vocabulary in a second language normally does not involve learning new concepts or learning how to use words. Rather, it involves learning new labels for semantic areas and developing the ability to recognize these labels in context as an initial bottom-up processing".*

(59c) *"An active vocabulary is perhaps better described as a motivated vocabulary. It consists in all those words we need to use and have no reservations about using to communicate with one another on an everyday basis ... A passive vocabulary could be said to include the individual's active vocabulary as an important subset of itself. It differs from active vocabulary, though, in also including the individual's unmotivated vocabulary".*

(60c) *"... how unsatisfactory it is trying to define the distinction in operational terms and to specify what the distinction is at the conceptual level ...".*

(61c) *"Teachers quite often assume that any vocabulary exercise or test assume a choice of either active or passive knowledge of the words. This is simply not the case: judging the requirements of a vocabulary activity means judging its position on a bi-polar continuum ... Approaching vocabulary instruction from an either-or perspective can prevent teachers from making proper judgements about the requirements of various vocabulary forms".*

(62c) *"... the word knowledge can be placed on a continuum, starting with 'I never saw this word before'".*

(63c) *"... and finally we reach the stage where we have pinned the word down. We know it. We would recognize it again and again if we saw it, and we are likely to remember it".*

(64c) *"... we can make tests still more precise by requiring finer discriminations as between the arresting and curing of a disease; irony, sarcasm and satire; sympathy and empathy ...".*

(65c) *"... knowing the concept does not imply knowing much about the concept, just as it does not imply having experienced the referent; it may imply knowing only which superordinate peg to hang the concept onto".*

(66c) *"A consideration of some of the knowledge that is assumed by lexical competence is offered as a frame of reference for the determination of objectives for vocabulary teaching and for the assessment of teaching techniques designed to realize these objectives".*

(67c) *"Given a list of words, with the exception of concrete nouns a native speaker can classify them into 'frequent', 'moderately frequent', 'not frequent', to a degree of accuracy reasonably close to their actual frequencies".*

(68c) *"... the rules that enable us to build up different forms of the word or even different words, from that word".*

(69c) *"... knowledge of something much more abstract than language: it is a knowledge of systems of rules, of parameters or principles, configurations in the mind for which language simply serves as evidence".*

(70c) "... it must go beyond establishing accurate responses to words. It must develop fluency of access to word meaning and must integrate the learner words into existing semantic networks"

(71c) "The objectives of L1 e L2 reading ... must be the same. However, in respect to the actual processes of acquiring these objectives, L1 e L2 reading may well differ".

(72c) "... que consiste en modificar aligerándola la noción de significado literal, sin excluirla ... la fuerza ilocucionaria de un acto del discurso está relacionada con la frase enunciada sólo provisionalmente. Algunas frases son mui explícitas en su especificación de los actos del discurso que sirven para desempeñar, mientras otras ofrecen unas sugerencias muy generales en cuanto a los posibles tipos de actos del discurso desempeñados por su uso. Tales sugerencias han de ser completadas por información contextual con el fin de producir una asignación de la fuerza ilocucionaria".

(73c) "... the study of the relation between the abstract meaning or sense of linguistic expressions, and the communicative force which they have for speakers and hearers in given utterance situations".

(74c) "... two types of meaning: the meaning (which is frequently described as 'literal') and the illocutionary force".

(75c) "... linguistics has to be pragmatic: the users of language, not the linguists, are the most important persons in linguistics".

(76c) "... concepts, like all things created by humans are historical entities, as as such, intimately associated with the organizational of a particular society at a given time".

(77c) "This is somewhat oversimplified. The notion of 'context' itself is a complex issue. Automatic recognition and spreading activation do, in fact, involve subconscious textual processing of some sort. This is not, however, the sort of context that most reading researchers discuss when they refer to the importance of context".

(78c) "... are of interest for what they reveal about the way readers manage their interaction with with written text an how these strategies are related to text compreehnsion".

CAPÍTULO III

O DESENHO E A METODOLOGIA DE PESQUISA

No capítulo anterior busquei construir, através de uma revisão crítica da literatura e da explicitação de modelos/visões de leitura e de competência lexical, uma base teórica para esta investigação. Este capítulo, organizado em três seções, tem por objetivo apresentar sua base metodológica. Na Seção 3.1 é abordado o desenho da pesquisa e sua justificativa. Esse desenho pressupõe dois estudos, um de ênfase quantitativa e foco no produto da leitura em LE, e outro de ênfase qualitativa, com foco no processo. As Subseções 3.1.1 e 3.1.2 explicitam, respectivamente, o sentido das dicotomias produto/processo e qualitativa/quantitativa na literatura em geral e particularmente neste estudo. A pesquisa de ênfase quantitativa, a descrição dos sujeitos participantes, assim como dos instrumentos e procedimentos para a coleta dos dados são apresentados na Seção 3.2. A Subseção 3.2.3 apresenta e justifica as técnicas estatísticas usadas na análise desses dados. A terceira e última seção (3.3) aborda a pesquisa de ênfase qualitativa, incluindo a seleção e descrição dos sujeitos e a técnica usada para elicitação dos dados.

3.1 – O desenho de pesquisa: a escolha da metodologia e sua justificativa

O desenho desta pesquisa procura levar em conta as limitações teóricas e metodológicas dos estudos que investigam a relação entre competência lingüística/lexical e leitura apresentadas no Capítulo II. Elas se referem ao desenho de pesquisa e também à metodologia de coleta dos dados. Esta geralmente é orientada para o produto da compreensão, baseando-se, de forma geral, em resultados de testes e os estudos portanto, de ênfase quantitativa. Os poucos dados de processo são obtidos através da análise de lapsos orais (*miscue analysis*), um instrumento controvertido pelo fato de se basear na premissa, não comprovada, de que a leitura oral é equivalente à leitura silenciosa. Os desenhos de pesquisa, por sua vez, não parecem levar conta a proficiência em leitura em LM dos sujeitos, considerando apenas a proficiência de leitura em LE. Uma outra limitação diz respeito ao fato de não compararem leitores de níveis de competência lingüística diferentes com sua proficiência em leitura nas duas línguas.

Com relação às questões teóricas, é necessário salientar que a pesquisa que aborda a contribuição do vocabulário para a leitura considera esse componente isoladamente dos outros componentes, não levando em conta sua inter-relação. Além disso, parece basear-se em modelos

ideais de leitura em LM que nem sempre conseguem explicar a leitura em LE. Uma outra limitação seria não explicitar, na maioria das vezes, o conceito de competência lexical que fundamenta as medidas usadas, como já salientei no capítulo anterior. O que se pode inferir, entretanto, é que esse conceito enfatiza apenas o conhecimento enquanto um sistema, não levando em conta uma outra dimensão igualmente importante, isto é, a de uso. Além disso, esse conceito de vocabulário tem-se limitado apenas ao nível semântico ou de significado de palavras, focalizando a extensão do vocabulário ou o número de palavras desconhecidas/conhecidas, avaliadas geralmente através da combinação de uma palavra com um sinônimo (mais especificamente uma tradução, no caso de LE), ou através de uma definição analítica (como as encontradas em dicionários) ou ainda através de uma lista de verificação (*checklist*) (vide Capítulo II, Seção 2.4.3.3). Considerações mais qualitativas ou de profundidade desse conhecimento, que levam em conta não apenas o significado mas outros aspectos têm sido ignoradas.

Procurando preencher as lacunas salientadas, o desenho desta pesquisa considera:

- A leitura na perspectiva de um modelo interativo, e, portanto, resultado de uma inter-relação de fatores, que incluem a competência lexical, a proficiência em leitura em LM, o conhecimento conceitual ou prévio e diferentes concepções de leitura. Enquanto a competência lexical e a compreensão constituem as variáveis independente e dependente deste estudo, respectivamente, as outras podem ser vistas como moderadoras;
- Diferentes níveis de proficiência linguística em LE;
- O conhecimento lexical enquanto parte de uma competência lexical mais ampla, abrangendo, além do conhecimento, também seu uso. A dimensão de uso focaliza as habilidades necessárias ao acesso a significados de palavras através do reconhecimento automático/não automático (ou habilidades de nível mais baixo ou de decodificação), assim como os procedimentos de nível mais alto ou aqueles usados na construção do sentido do texto. Além disso, considera um conceito *rico* de vocabulário, mais consistente com uma visão de leitura interativa, que leva em conta a profundidade desse conhecimento e não apenas sua extensão, envolvendo os níveis lexical, semântico, morfológico, sintático e pragmático-discursivo; e, finalmente,
- Uma triangulação de informações como um recurso de validação - mesmos dados coletados através de instrumentos diferentes.

Levando em conta esses aspectos, esta pesquisa, portanto, apresenta dois focos de análise ou uma divisão em duas fases:

- Fase I, de ênfase quantitativa, dirigida para o produto de testes de leitura em LE e vocabulário, num primeiro plano e testes de proficiência linguística, conhecimento conceitual, leitura em LM e concepções de leitura num segundo; e
- Fase II, de ênfase qualitativa, dirigida para o processo da leitura.

Enquanto a primeira fase mostra uma dimensão quantitativa, horizontal, focalizada no

universo de 49 sujeitos/leitores, a segunda introduz uma dimensão qualitativa, vertical, focalizada no universo de três leitores, seleccionados a partir da primeira fase.

Antes de focalizar cada uma das fases, explico o que as dicotomias qualitativo/quantitativo e processo/produto significam no contexto deste estudo, uma vez que seus sentidos parecem variar na literatura. Também justifico o uso, em uma mesma pesquisa, de métodos aparentemente vistos como conflitantes e irreconciliáveis. Uma breve consideração sobre paradigmas de pesquisa subjacentes a esses métodos, apresentada na seção seguinte, é meu ponto de partida.

3.1.1 – A distinção quantitativo/qualitativo como paradigmas de pesquisa

A distinção entre métodos quantitativos e qualitativos na avaliação de pesquisas tem gerado um grande debate. Por um lado, embora o protótipo de método quantitativo seja um experimento verdadeiro (*true experiment*), elaborado para testar hipóteses através do uso de instrumentos objetivos e análises estatísticas (Larsen-Freeman & Long 1991), observa-se que também tem incluído os chamados métodos objetivos, usados não apenas em experimentos, mas também em quase-experimentos, testes de lápis e papel e análises estatísticas em geral (Reichardt & Cook 1979). Por outro lado, embora o método qualitativo seja melhor exemplificado por um estudo etnográfico, em que os pesquisadores não se propõem a testar hipóteses, mas apenas a observar o que está presente através de seu foco (Larsen-Freeman & Long op.cit.), nota-se que são considerados nessa categoria estudos de caso, entrevistas verticais, observação participativa e mais recentemente também a introspecção. Como pode-se notar, os métodos quantitativos/qualitativos de pesquisa abrangem desde métodos para elicitação de dados até formas de análise desses dados.

Para Reichardt & Cook (op. cit.) esse debate, embora na superfície centrado na distinção entre os métodos é, na realidade, um conflito entre os paradigmas metodológicos aos quais esses métodos estão associados. Em outras palavras, o método é de tal maneira ligado a um paradigma que ao se escolher o paradigma se está inevitavelmente definindo o método. Na prática, identifica-se uma associação entre método quantitativo e paradigma positivista (ou ainda nomológico-analítico para Grotjahn 1987), por um lado, e método qualitativo e paradigma interpretativista (ou ainda exploratório-interpretativista, para Grotjahn op. cit.) de outro (vide também Baltra Montaner (1982) para uma discussão mais detalhada dessa questão). Como os dois paradigmas são considerados mutuamente exclusivos, mostrando visões de mundo consideradas antagonistas, tem-se a impressão de que uma reconciliação entre dois métodos de paradigmas opostos em uma mesma pesquisa se torna impossível. Para Reichardt & Cook op. cit.:9):

Basicamente, a questão não diz respeito às estratégias de pesquisa

propriamente ditas. A adesão a um paradigma como oposto a um outro nos predispõe a ver o mundo e seus eventos de maneiras muito diferentes (1c).

A associação paradigma/método é tão forte que a dicotomia quantitativo/qualitativo passa a caracterizar paradigmas, ou seja, paradigma qualitativo e paradigma quantitativo (1).

Um outro aspecto controvertido desse debate refere-se à definição de um paradigma enquanto um conjunto fixo de atributos. Dessa forma, o paradigma positivista ou quantitativo apresentaria uma visão do mundo das ciências naturais, que é hipotético-dedutiva, particularista, objetiva, orientada para o produto, assumindo uma realidade estável, generalizável, confirmatória, expansionista e inferencial. O paradigma interpretativista ou qualitativo, por sua vez, mostraria uma visão do mundo antropológico, que é fenomenológica, indutiva, holística, subjetiva, orientada para o processo, assumindo uma realidade dinâmica, não generalizável, exploratória e descritiva.

A visão de paradigma como um conjunto rígido de atributos tem sido criticada por esses mesmos autores, dentre outros:

(...) a escolha de métodos não deveria ser determinada pela fidelidade a um paradigma arbitrário. Isso porque um paradigma não está inerentemente ligado a um conjunto de métodos e porque as características de um contexto de pesquisa específico são, da mesma forma que os atributos de um paradigma, importantes na escolha do método (Reichardt & Cook op.cit.:19) (2c).

Para esses autores, portanto, todos os atributos acima podem ser aplicados aos dois tipos de métodos, uma vez que não parece haver nenhuma ligação inerente ou lógica entre atributos e métodos. A única associação seria aquela determinada pela prática, o que poderia ser considerado uma simplificação, pois a prática já é geralmente informada por uma visão ou teoria.

Uma posição bastante semelhante é defendida por Larsen-Freeman & Long. Também discordando de uma divisão rígida entre paradigmas, os autores apresentam os métodos em um contínuo, que têm nos seus polos os paradigmas quantitativo e qualitativo.

Jacob (1989), ao invés de paradigmas, sugere o termo tradição, que a autora define através de dois critérios: um conjunto de pressuposições filosóficas sobre o universo assim como um conjunto de métodos apropriados. Contudo, propõe esse termo por discordar da classificação em qualitativas de todas as alternativas de pesquisa que não seguem a linha tradicional de pesquisa positivista. Para ela, há várias linhas ou tradições que distinguem essa pesquisa e qualitativo é um termo muito abrangente e genérico que não capta as especificidades de cada uma delas. Entretanto, devo salientar que, embora usando um termo diferente, tradição, a autora não deixa de reconhecer a importância de se adotar um paradigma.

A rigidez e inadequação dos dois paradigmas também é criticada por Grotjahn (op. cit.). Sua proposta é uma ampliação dos dois paradigmas em suas formas puras para seis formas mistas, identificadas através de três critérios: o método de coleta ou forma dos dados (quantitativo/qualitativo), o método de análise (interpretativo/estatístico) e ainda a maneira da coleta ou o desenho da pesquisa (desenho não experimental vs. (quase-) experimental).

Além desses critérios, esse mesmo autor, considerado um representante da pesquisa de base introspectiva, também reconhece a importância dos objetivos do processo de coleta dos dados, isto é, a exploração da realidade para a formação/testagem de hipóteses ou para a aplicação da teoria tendo como metas a introdução de mudanças em certos fenômenos.

O uso dos termos quantitativo/qualitativo para definir dois paradigmas implica, também para Erickson (1991), do ponto de vista da etnografia educacional, numa dicotomia enganosa. O autor propõe o uso do termo *interpretativista* (*interpretive*) para caracterizar uma família de abordagens que inclui etnografia, estudos de caso e pesquisa qualitativa, e os termos levantamentos (2) e experimentos, para se referir à pesquisa normalmente caracterizada como quantitativa. A razão dessa mudança é que grande parte dessa pesquisa, dita qualitativa, ou que tem um objetivo geral qualitativo, envolve técnicas quantitativas e até mesmo análises estatísticas. Para Erickson (op. cit.) *contar ou não contar coisas* não seria um critério relevante para contrastar levantamentos e experimentos, de um lado, e pesquisa interpretativista, de outro. Mais importante nessa distinção seria a textura dos dados e o papel da análise causal. Com relação ao primeiro aspecto, ele afirma:

Em levantamentos e experimentos de grande escala, coleta-se relativamente pouca informação sobre muitos sujeitos, enquanto na pesquisa interpretativista coleta-se muita informação sobre relativamente poucos sujeitos Erickson (op. cit.:339) (3c).

Considerando também o papel da análise causal, o autor salienta:

Em levantamentos e pesquisa experimental, a 'causa' é entendida como nas ciências naturais, ou seja, uma noção física de causa. Na pesquisa interpretativista, não se está muito interessado na causa propriamente dita, mas na explicação em termos de significado e entendimento daquilo que está sendo estudado (4c).

Embora reconhecendo que dificilmente essas controvérsias deixarão de existir, considero necessário salientar que as críticas colocadas têm sido importantes no sentido de redefinir o debate para aspectos mais produtivos, mostrando possibilidades de combinação de métodos e usos mais criativos, que vão além da rigidez de dois paradigmas. Como salientam Reichardt & Cook (op. cit.:27),

(...) chegou a hora de impedir que o pêndulo mova de um extremo para o outro. Chegou a hora de parar de construir muros entre os métodos e começar a construir pontes. Talvez tenha até mesmo chegado a hora de ir além da linguagem dialética de métodos qualitativos e quantitativos. O desafio real é combinar os métodos de pesquisa ao problema de avaliação sem 'paroquialismo' (3) (5c).

O fato de as críticas serem colocadas por autores de tradições de pesquisa completamente opostas (Erickson e Jacob, por exemplo, são representantes de uma tradição de pesquisa qualitativa, enquanto Grotjahn e Reichardt & Cook vêm de uma tradição de pesquisa quantitativa) faz com que elas extrapolem o viés da visão de seu proponente e tenham um peso ainda maior. Elas mostram um reconhecimento de que todos os métodos são falhos, mas que esse fato não seria uma razão para descartá-los, mas sim, encontrar maneiras de minimizar os aspectos negativos e tentar melhorá-los.

Antes de caracterizar melhor meu posicionamento assim como justificar minha proposta para este trabalho, considero necessário discutir mais detalhadamente também a dicotomia produto/processo.

3.1.2 - A distinção produto/processo na pesquisa em LA

A costumeira associação entre produto e paradigma positivista ou quantitativo e processo e paradigma interpretativista ou qualitativo parece encobrir diferentes sentidos de uso da dicotomia não apenas na pesquisa em geral, mas também na pesquisa no ensino/aprendizagem de línguas assim como em outras sub-áreas da LA.

No contexto de avaliação da pesquisa em geral, a análise de processo, na opinião de Reichardt & Cook (op. cit.:21) tem dois propósitos distintos, identificados por rótulos diferentes.

O primeiro significado de processo é monitoração: descrever o contexto e a população do estudo, descobrir de que maneira o tratamento ou o programa foi implementado, fornecendo realimentação formativa imediata, e coisas do tipo. O segundo sentido de processo é uma explicação causal: descobrir ou confirmar o processo pelo qual o tratamento teve o efeito que teve. Certamente, medir o efeito do programa é o resultado somativo ou avaliação de impacto. Para uma compreensão ou entendimento completo do programa, uma avaliação teria que incluir pelo menos estas três tarefas: monitoração, avaliação de impacto e explicação causal (6c).

No contexto da pesquisa de ensino/aprendizagem de línguas, observa-se a tendência de se associar um atributo do paradigma qualitativo ou orientado para o processo com uma abordagem

longitudinal ou estudo de caso, e um atributo do paradigma qualitativo, ou orientado para o produto com uma abordagem transversal (*cross-sectional*). A descrição de cada uma das abordagens oferece subsídios para que se possa entender melhor a dicotomia nesse contexto. Segundo Larsen-Freeman & Long (op. cit.:11),

Uma abordagem longitudinal (...) tipicamente consiste em observar o desenvolvimento do desempenho lingüístico, geralmente a fala espontânea de um sujeito, quando os dados da fala são coletados em intervalos periódicos durante um certo período de tempo. Em uma abordagem transversal (cross-sectional), o desempenho lingüístico de um número maior de sujeitos é estudado, e os dados de desempenho são geralmente coletados em apenas uma sessão (7c).

Dessa maneira, orientado para o processo significa que o desempenho lingüístico do sujeito é avaliado *durante um certo período de tempo (it takes place over time)*, enquanto, orientado para o produto significa que esse desempenho é avaliado *em apenas um ponto no tempo*. Entretanto, como salientam Faerch & Kasper (1987), os estudos longitudinais são, na realidade, uma maneira indireta de se avaliar o processo, constituindo uma tentativa de reconstrução do desenvolvimento do aprendiz através de produtos sucessivos. Vários métodos de análise têm sido utilizados com essa função, incluindo a análise contrastiva e a análise de erros. Mais recentemente, entretanto, tem-se buscado um acesso direto a esse processo, não mais através de fontes externas do aprendiz, mas através de depoimentos dos próprios informantes sobre as maneiras como organizam e processam informações, ou através de métodos introspectivos.

A dicotomia processo/produto tem ainda outros sentidos nas várias sub-áreas da LA. O ensino da produção escrita, a pesquisa em sala de aula, a pesquisa em leitura, a tradução e a avaliação são algumas das áreas que têm feito referência e se utilizado dessa distinção. Uma breve consideração de alguns desses usos pode ser elucidativa para este trabalho.

Originalmente, a dicotomia aparece associada à distinção forma/função da língua. Uma ênfase no produto pode ser observada na medida em que a forma da língua gozava de prestígio e era enfatizada. A partir do momento em que a função passa a ser prestigiada, observa-se uma ênfase maior no processo. O ensino voltado para a produção de textos escritos parece ilustrar bastante bem essa alternância de valores e de foco.

Há aproximadamente dez anos, o ensino e a pesquisa sobre produção escrita, tradicionalmente centrada no produto, ou na composição final, passa também a considerar o processo através do qual as pessoas constroem seu texto, deslocando, dessa forma, o ensino para os processos mentais subjacentes ao texto. Um ensino centrado no produto consistia apenas na correção de erros sintáticos, morfológicos, lexicais e semânticos; qualquer revisão editorial ou que envolvesse reformulações relativas à estruturação do texto era vista como desnecessária. Conseqüentemente, as

avaliações eram baseadas na precisão técnica e na avaliação da coesão e coerência como um produto final. Kaplan (1988) salienta que a ênfase no produto, nesse contexto, foi culpada por todas as falhas no ensino da escrita.

Um ensino centrado no processo, por sua vez, muda o foco para uma abordagem em que os conceitos de coesão/coerência são discutidos através de exemplos extraídos do texto do aluno, explorando, de forma explícita, as características sintáticas que levam à coesão e coerência, ou as estratégias retóricas subjacentes a elas (Kaplan op. cit.).

Um outro exemplo de uso da dicotomia produto/processo pode ser observado na pesquisa em sala de aula no ensino/aprendizagem de L2. Nesse caso, a pesquisa orientada para o produto focaliza os resultados do programa, ou uma testagem de hipóteses relativas à sua eficiência quando comparado com outro, ou ainda aquilo que ele produz em termos da aprendizagem do aluno, sendo, portanto, uma pesquisa essencialmente prescritiva. A pesquisa orientada para o processo, por outro lado, é de natureza descritiva, tendo por objetivo gerar hipóteses relativas às relações entre práticas instrucionais específicas e resultados de aprendizagem (Spada 1987).

A dicotomia ainda é observada num contexto de ensino/aprendizagem de LE, vista agora na perspectiva da pedagogia de sala de aula. Tarvin & Al-Arishi (1991), observando que as atividades desenvolvidas no ensino comunicativo têm privilegiado ou promovido a sensação ou ações visíveis (respostas espontâneas) em detrimento da reflexão, propõem três tipos de atividades que, em sua opinião, preencheriam essa lacuna, ou encorajariam a reflexão (4). Duas dessas atividades parecem se fundamentar na dicotomia produto/processo: atividades orientadas para a tarefa (*task-oriented activities*) e orientadas para o processo (*process-oriented activities*). Uma atividade orientada para a tarefa focaliza o produto, uma vez que sua ênfase é no resultado, no ponto de descoberta, geralmente na solução de um problema, que é "fechado" ou único. Uma ênfase no produto, entretanto, não significa que o processo seja totalmente ignorado. Este também é considerado, uma vez que os alunos são incentivados a explicitar as bases das hipóteses que os conduziram aos resultados. Uma atividade orientada para o processo, por outro lado, como o nome diz, enfatiza o valor dos procedimentos, não mais seu resultado, e, nesse sentido, considera as várias possibilidades de solução.

O terceiro tipo de atividades mencionadas pelos autores são atividades orientadas para a síntese (*synthesis-oriented activity*), aparentemente não relacionadas com a dicotomia produto-processo. Entretanto, a meu ver, esse tipo de atividade poderia ser mais eficiente se também explorasse a dicotomia estabelecendo a relação entre atividades orientadas para o produto e para o processo.

Em avaliação, a distinção também é usada em duas situações distintas. A avaliação somativa ou de produto focaliza, como no caso da pesquisa em sala de aula, o resultado do ensino/aprendizagem enquanto a avaliação formativa ou de processo focaliza o processo da

aprendizagem. Ainda no contexto da avaliação, agora mais especificamente da leitura, um foco no processo significa atenção aos processos mentais de construção do sentido ou compreensão, enquanto um foco no produto pressupõe atenção ao resultado da leitura, ou à compreensão.

A ênfase no produto da leitura é criticada por Yamamoto (1991:9) sob o ponto de vista do ensino da leitura:

(...) a escola focaliza apenas o produto da leitura, isto é, o aspecto mensurável da compreensão do texto didático, geralmente através de perguntas e respostas, ou de testes de múltipla escolha, em que o aluno ao ler utiliza-se de um processo automático na busca do reconhecimento da resposta correta extraída do texto, numa atitude mecânica.

Entretanto, sob a ótica da avaliação da leitura, mais especificamente da proficiência, o foco no processo também pode ser criticado, uma vez que não é relevante procurar saber que rotas ou caminhos o leitor usou para chegar ao produto, o que realmente interessa é o produto ou a compreensão.

Na área de tradução, o uso da dicotomia assemelha-se ao da área de leitura. Enquanto o produto refere-se ao texto traduzido, terminado, o processo refere-se aos processos mentais em que as pessoas se engajam ao traduzir, ou o que fazem, passo a passo, enquanto traduzem. Entretanto, grande parte da vasta literatura em tradução tem recebido várias críticas por focalizar o produto (Gerloff 1987; Krings 1987) ou por se basear em medidas externas de um fenômeno que é interno, medidas essas geralmente de características lingüísticas, fazendo com que muito pouco se saiba sobre o processo e suas características psicolingüísticas. Os poucos dados sobre o processo são inferidos a partir do produto.

Deixando de lado as especificidades de cada situação, convém ressaltar que em todos esses exemplos, as críticas mais fortes à dicotomia têm-se referido à tendência de se enfatizar um aspecto em detrimento ao outro. No caso da pesquisa em sala de aula, salienta Spada (op. cit.:138), apesar de cada abordagem de pesquisa apresentar problemas distintos, pode-se observar que elas são relacionadas.

(...) o problema mais sério é que estudos orientados para o produto e para o processo são vistos muito freqüentemente como competidores ao invés de abordagens complementares (...) Como resultado, temos estudos orientados para o processo sem um componente de produto. Claramente, há a necessidade de ambas as investigações no ensino e aprendizagem em contextos de sala de aula (8c).

Crítica semelhante é feita por Kaplan (1988:35) para a escrita. Concordo com o autor

quando salienta que a distinção entre processo e produto, nesse caso, estabelece uma falsa dicotomia, que obscurece fatos e mascara variáveis importantes, uma vez que todo processo conduz a um produto, embora provisório, e cada produto é o resultado de um processo.

Tendo caracterizado as dicotomias qualitativo/quantitativo e produto/processo na literatura, justifico, a seguir, o seu uso neste trabalho.

3.1.3 – As dicotomias qualitativo/quantitativo e processo/produto neste trabalho

É importante salientar que a visão acima apresentada, em que as alternativas metodológicas de pesquisa são classificadas de acordo com apenas dois paradigmas, ou que a escolha do paradigma logicamente determina a escolha do método, ou, ainda, em que o conjunto de atributos é inerentemente ligado a apenas um dos paradigmas, não podendo se aplicar aos dois, não é a que fundamenta esta pesquisa. Neste caso, ela se distingue especialmente pela possibilidade de separação entre método e paradigma e pela livre escolha entre os vários atributos dos dois tipos de paradigmas; a decisão vai ser, nesta situação, definida pelo problema de pesquisa, que pode permitir a combinação de vários métodos e atributos se for o caso, ou se essa solução for aquela que melhor ofereça condições de solucionar o problema. Essa posição parece coincidir com a de Larsen-Freeman & Long (op. cit.:14) quando afirmam que

(...) o que é importante para os pesquisadores não é a escolha de paradigmas a priori ou até mesmo de metodologias, mas sim a clareza do propósito do estudo e escolher os atributos que sejam melhores para alcançar tal propósito. Em outras palavras, o desenho metodológico deve ser determinado pela questão de pesquisa (9c).

No caso específico desta pesquisa, o tópico a ser investigado, ou seja, a relação entre leitura e vocabulário, tem duas dimensões: a primeira, que focaliza o conhecimento de vocabulário enquanto um sistema, com contribuições para o resultado dessa leitura e daí, o foco no produto; e uma segunda, que focaliza o uso desse conhecimento e conseqüentemente, o foco no processo. O conhecimento, nesse caso, é visto como um recurso e não como uma configuração mental que existe por si própria como uma estrutura (Widdowson 1989) e, portanto, precisa de uma dimensão que explore como essas abstrações mentais são usadas. As duas dimensões, conhecimento lexical (ou representação mental de informação sistemática e organizada do léxico da língua) e habilidade ou capacidade de uso (procedimentos para uma recuperação efetiva e eficiente desse conhecimento) compõem o que chamaria de competência lexical (5). Essa competência lexical é parte de uma competência ainda mais ampla, que incorpora o conhecimento prévio ou esquemático assim como

seu uso (*content schemata*) e ainda o conhecimento/habilidade de como o texto funciona (*formal schemata*). A leitura, portanto, envolve a interação da competência lexical com outras fontes de conhecimento ou competências.

A tendência geral observada na literatura é que essas duas dimensões não só são independentes mas competidoras, ora se enfatizando a língua, ora as habilidades ou estratégias. Além disso, alguns modelos, como por exemplo o modelo descendente e também o interativo-compensatório, concebem a dimensão de habilidades como compensatória, o que poderia significar o reconhecimento de um papel menos relevante para a primeira dimensão. Na minha visão, as duas dimensões são importantes e se complementam. A variação de desempenho na leitura depende desses dois fatores. O foco em um fator em detrimento do outro tem implicações diferentes para o ensino e conseqüentemente, para a avaliação.

A fase I deste estudo, portanto, focaliza, num primeiro plano o produto de testes de vocabulário e de leitura, e, num segundo plano, o produto de testes de proficiência linguística, de conhecimento conceitual, de leitura em LM e de uma medida de concepções de leitura. O objetivo, neste caso, é uma medida do conhecimento lexical (e dos outros tipos de conhecimento) de vários sujeitos para se estabelecer relações com seu desempenho na leitura em LE. Mais especificamente, o propósito é identificar um estágio ou um ponto fixo no processo de aquisição do léxico (e outros aspectos) e da leitura que possam ser comparados, ou, em outras palavras, investigar quanto do sistema da língua é conhecido/desconhecido e quanto da leitura pode ser compreendido. Para isso, este estudo se utiliza de um método de coleta de dados que é quantitativo, com foco no produto. Apesar do foco no produto, entretanto, podem-se identificar algumas tentativas de se inferir também o processo, embora indiretamente, ou seja, através de uma análise dos erros dos sujeitos nos testes (6).

Neste trabalho, também faço uma distinção entre métodos de coleta e análise de dados. Essa distinção é mais uma vez eclipsada pela definição de método como um conjunto fixo de atributos ou ligações inerentes entre métodos e paradigmas que salientei acima. A seleção de um método de coleta não determina necessariamente a escolha do procedimento analítico e vice-versa, uma vez que tanto dados de desempenho conversacional como relatos introspectivos podem ser submetidos a tratamentos qualitativo-interpretativos ou quantitativo-estatísticos (Grotjahn 1987).

Na fase I deste estudo é utilizada, para análise dos dados, também um método predominantemente quantitativo, que usa técnicas de análise estatísticas para se chegar a uma quantificação que valorizam o indivíduo pela sua contribuição à tendência estatística geral dos dados. Nesse sentido, entretanto, três observações precisam ser feitas. Em primeiro lugar, a análise estatística não elimina a necessidade de um entendimento qualitativo, uma vez que a escolha de um modelo estatístico e a interpretação desses resultados envolvem conhecimento qualitativo. Concordo com Reichardt & Cook (1979) quando afirmam que não se pode beneficiar dos números sem saber o que

eles significam. Além disso, embora haja predominância de uma técnica de análise quantitativa, deve-se ressaltar que paralelamente a essa análise o estudo também salienta aspectos individuais dos sujeitos, isto é, os considera como indivíduos e portanto, de forma também qualitativa.

Em segundo lugar, embora o método de coleta de dados seja quantitativo, é necessário ressaltar que o desenho de pesquisa não é experimental, prototípico do paradigma positivista ou quantitativo, uma vez que não foram observados os dois critérios que definem esse tipo de desenho, isto é, a seleção aleatória dos sujeitos e a existência de um grupo de controle. Como afirma Larsen-Freeman (1991:21),

A idéia básica de um experimento é poderosa. Se um grupo de sujeitos é tratado de uma maneira e um outro grupo de maneira diferente, e não há outros fatores influenciando os dois grupos, pode-se determinar uma relação de causa-efeito entre tratamento e consequência (10c).

Neste estudo, uma metodologia experimental se mostra inapropriada porque as duas condições ou critérios acima não podem ser levados em conta. Além de os sujeitos não terem sido escolhidos aleatoriamente, uma vez que faziam parte de classes pré-existentes, não era pertinente um grupo de controle, uma vez que o desenho não previa um tratamento para esses sujeitos.

Uma terceira observação diz respeito à natureza da análise estatística usada. Nesse sentido, o estudo também não usa técnicas de análise inferenciais ou de confirmação de hipóteses que justifiquem o rótulo quantitativo/positivista. O uso de técnicas estatísticas exploratórias enfatiza seu caráter observacional e de levantamento de hipóteses.

A fase II, por sua vez, focaliza o segundo aspecto do tópico investigado ou a dimensão dinâmica do conhecimento de vocabulário. Essa dimensão envolve a habilidade de uso desse conhecimento na construção do sentido do texto, geralmente mediado pelo acesso automático e não automático a significados, ou pelo reconhecimento de palavras durante a decodificação. Nessa dimensão, portanto, estão envolvidos não apenas as fontes de conhecimento/procedimentos de uso mas também a eficiência e velocidade ou automaticidade com que tal recuperação é colocada em operação pelo sujeito.

A técnica escolhida para captar esse processo é um método introspectivo, ou seja, o protocolo de pausa (vide Seção 3.3.2 para uma descrição mais detalhada desse método), por ser aquele que a meu ver, mais se aproxima do processo de construção da compreensão. Nesse caso, verbalizações concomitantes são feitas pelo sujeito enquanto a informação ainda se encontra na memória de curto prazo. Os dados de processo são complementados por dados de produto, ou de resumos que os sujeitos são solicitados a elaborar no final de cada parágrafo, trazendo maior validade e confiabilidade a esses dados.

O método de análise dos dados na fase II é predominantemente qualitativo na medida em

que valoriza a contribuição de cada indivíduo. Entretanto, esse tipo de ênfase na análise não significa que quantificações tenham sido eliminadas. O conhecimento quantitativo está presente em avaliações qualitativas da mesma forma que o conhecimento qualitativo está presente em avaliações quantitativas. Deixar de mencionar conceitos quantitativos como número de ocorrência de uma estratégia ou uso de uma fonte de conhecimento leva muitas vezes a caracterizações enganosas mesmo em um estudo de caso.

Como pode-se observar, o tópico a ser investigado justifica diferentes métodos, uma vez que o uso de apenas um método não seria suficiente para abranger as especificidades de cada aspecto. Entretanto, embora com objetivos aparentemente diferentes, não posso deixar de salientar que eles estão sendo usados com o mesmo propósito, ou seja, para investigar duas dimensões de um mesmo problema. Nesse sentido, poderia afirmar que os dois métodos diferentes se complementam, oferecendo contribuições que nenhum sozinho seria capaz de oferecer.

Uma outra característica deste estudo é usar métodos diferentes para a coleta de mesmas informações, ou uma triangulação. Como todos os métodos apresentam vieses e limitações que muitas vezes se misturam à realidade observada, poderia dizer que uma combinação de métodos oferece melhores condições de se avaliar um problema. Neste trabalho, por exemplo, a leitura foi avaliada através de métodos diferentes, tais como de perguntas abertas, do tipo cloze e auto-avaliação. O cruzamento de todas essas medidas favorece sua validação.

Considerando-se o número de sujeitos, as variáveis e os métodos tanto de coleta como de análise desta pesquisa, pode-se dizer, portanto, que é de natureza descritiva focalizada, complementada por um estudo introspectivo, fornecendo dados de natureza observacional e exploratória.

3.2 - Fase I: o estudo de produto com ênfase quantitativa

O foco da pesquisa é a relação entre conhecimento de vocabulário, medido por testes de vocabulário e o produto da compreensão em leitura, medido por testes de leitura, num primeiro plano e testes de proficiência lingüística, conhecimento conceitual, leitura em LM e concepções de leitura num segundo plano.

3.2.1 – Seleção e descrição dos sujeitos

Nesta fase, foram sujeitos 49 alunos de graduação, com níveis de proficiência em leitura que variavam de muito bom a muito fraco. Os sujeitos eram, em sua maioria, alunos de Engenharia

Elétrica e Matemática Aplicada em uma universidade estadual paulista cursando Inglês Instrumental I e II no segundo semestre de 1990 (vide Capítulos IV e V para uma complementação dessa descrição).

A seleção dos sujeitos apresenta uma justificativa teórica e prática. Em primeiro lugar, a necessidade de níveis diferentes de proficiência lingüística e de proficiência em leitura tornou necessária a seleção de sujeitos cursando níveis diferentes de inglês instrumental, uma vez que as chances de apresentarem níveis diferentes de proficiência aumentam, nesse caso. Isso não significa, entretanto, que as classes sejam homogêneas, uma vez que o critério para seleção de alunos para cada classe é puramente administrativo, uma vez que esses alunos não passam por uma pré-seleção ou teste de classificação. Um número mínimo de sujeitos, ou seja, 40, para que generalizações pudessem ser feitas, ou para que se pudesse discernir entre um comportamento típico daquele que é único ao sujeito, tornou necessário a seleção de mais de uma classe, uma vez que esta, em geral, apresenta apenas 20 alunos. Para não introduzir variáveis adicionais, a seleção dos sujeitos procurou levar em conta sua área de estudos. Foram escolhidos alunos, portanto, da área de Ciências Exatas, mais especificamente de Engenharia Elétrica e Matemática Aplicada, em sua maioria, por serem as turmas disponíveis no semestre em que os dados foram coletados.

Devo lembrar que o Inglês Instrumental é obrigatório para esses alunos durante dois semestres, com 60 horas/aula, distribuídas em 4 horas/aula semanais. Essa disciplina, entretanto, é colocada em épocas diferentes no currículo de cada curso. Para Matemática Aplicada, ela se inicia no quarto semestre de curso, enquanto para a Engenharia Elétrica, no primeiro semestre.

Portanto, quando os dados foram obtidos, os alunos de Matemática Aplicada cursavam Inglês Instrumental I, ministrado pela pesquisadora, enquanto os alunos de Engenharia Elétrica cursavam Inglês Instrumental II, subdivididos em três classes, ministradas por professores distintos. Todos os sujeitos, portanto, já haviam sido expostos a uma abordagem de ensino que tinha como objetivos principais torná-los mais conscientes das estratégias usadas na leitura em LM assim como praticá-las na leitura de textos gerais em inglês. Esse trabalho era centrado em estratégias como localização de informação específica, levantamento de hipóteses a partir do título, inferência lexical, e assim por diante, usando-se, em sua maioria, textos de vulgarização científica de temas variados, extraídos de jornais e revistas.

3.2.2 – Procedimentos e instrumentos de pesquisa

Para abranger o complexo de variáveis envolvidas no estudo, foram elaborados questionários e testes para a coleta dos dados. Coletaram-se também notas ou escores em leitura em português como LM e leitura em inglês como LE obtidas por esses alunos no vestibular de ingresso

à universidade como medidas complementares para comparações ou validação de algumas das medidas usadas neste estudo.

A Tabela 3.1, a seguir, apresenta os instrumentos de coleta dos dados e seus objetivos:

Tabela 3.1 – Instrumentos de coleta e seus objetivos

Instrumentos	Objetivos
1. Questionário I	Dados de perfil dos sujeitos
2. Questionário II	Auto-avaliação
3. Tarefa de leitura	Situação de leitura para facilitar o preenchimento do questionário III
4. Questionário III	Fatores metacognitivos/Concepções de leitura
5. Testes de vocabulário I, II, III e IV	Conhecimento de vocabulário
6. Testes de leitura I e II	Proficiência em leitura em LE
7. Teste de conhecimento conceitual	Conhecimento conceitual ou prévio do assunto do texto
8. Cloze I	Proficiência em leitura em LM
9. Cloze II	Proficiência em leitura em LE e proficiência em LE

Além disso, também foram coletados dados usando-se alguns dos testes de vocabulário e leitura, de um grupo de 12 falantes nativos (alunos de pós-graduação e professores na área de Linguística e Linguística Aplicada) com o objetivo de se convencionalizar o critério de correção de alguns dos testes de vocabulário e de leitura aplicados aos 49 sujeitos, e não como parâmetro ou uso da competência ideal do falante nativo para julgamento da competência do leitor em LE. Em alguns dos testes, como por exemplo, o de contextos gramaticais, parecia haver controvérsias até mesmo entre os falantes nativos com relação aos contextos considerados possíveis.

Os instrumentos usados na fase I deste estudo foram desenvolvidos, pré-testados e calibrados durante um período de três meses. A coleta de dados foi conduzida durante um período de cinco meses (7), seguindo o cronograma apresentado na Tabela 3.2 a seguir:

Tabela 3.2 – Cronograma da coleta de dados

Instrumento	Coleta
Questionário I	Semana I
Tarefa de leitura em Inglês, seguida dos questionários II e III	Semana II
Teste de vocabulário I	Semana III
Cloze I	Semana V
Cloze II	Semana VI
Teste de conhecimento conceitual	Semana VII
Teste de vocabulário II e III (seções A, B e C)	Semana VIII
Teste de vocabulário III (seção D)	Semana IX
Teste de vocabulário IV	Semana X
Teste de leitura I e II	Semana XII

As próximas seções descrevem cada um dos instrumentos.

3.2.2.1 – Questionário I

O questionário I (vide Apêndice A, Seção A.1) é constituído de 11 itens combinando perguntas de respostas fechadas e abertas em português. Seu objetivo é coletar dados que permitam uma caracterização ou um perfil mais detalhado dos sujeitos do estudo, complementando a descrição apresentada no início deste capítulo. Os itens de 1 a 3 (nome, RA (8) e idade) identificam o sujeito. Os itens 4, 5 e 6 visam construir o seu perfil com relação à proficiência lingüística em inglês, mais especificamente em termos do tempo e qualidade de exposição à língua. Em geral, observa-se uma variação dessa qualidade, dependendo da escola. Alunos que freqüentaram o colegial em escolas particulares e que tiveram, além disso, aulas em escolas de línguas em geral demonstram um nível mais alto de proficiência lingüística. A inclusão dos itens 7, 8 e 9 se justifica na medida em que permitem obter informações do sujeito a respeito de sua proficiência em leitura em inglês e português, que serão posteriormente relacionadas aos itens 1, 4, 8 e 9 (natureza das leituras em português e inglês) do questionário de auto-avaliação, complementando o perfil dos sujeitos quanto ao conhecimento prévio e proficiência em leitura. Os itens 10 e 11 focalizam a concepção que os sujeitos têm de um "bom leitor" e sobre o seu processo de leitura.

As instruções para o preenchimento do questionário foram dadas oralmente. O limite de tempo permitido foi de 15 minutos, que se mostrou suficiente para os candidatos responderem aos itens sem pressa. Os sujeitos não apresentaram dúvidas explícitas, uma vez que não houve perguntas.

3.2.2.2 – Tarefa de leitura em inglês e questionários II e III

São três os instrumentos usados nesta etapa da coleta. A tarefa de leitura precede o preenchimento dos dois questionários. Ela é realizada através de um texto de 560 palavras em inglês, intitulado *The tuned-out generation*, extraído da revista *Time*, de 9 de julho de 1990 (vide Apêndice A, Seção A.3). Essa tarefa tem por objetivo criar uma oportunidade de leitura para a tarefa de auto-avaliação (questionário II) e para o preenchimento do questionário III.

O questionário II (vide Apêndice A, Seção A.2) é constituído de cinco itens combinando perguntas fechadas e abertas, além de um item de múltipla escolha. Seus objetivos são a auto-avaliação da leitura em LE (item 1), das dificuldades encontradas na leitura (item 2) e do nível de dificuldade do texto (item 3). Uma avaliação da idéia geral do texto (item 4) foi solicitada para uma verificação posterior da auto-avaliação. Os itens 5 e 6 tentam identificar a concepção que o sujeito

tem de sua leitura em LM e LE. Essa informação já havia sido solicitada no questionário I (itens 5 e 6) e foi repetida para uma verificação da consistência das respostas.

O questionário III (vide Apêndice A, Seção A.5 e Tabela 2.1, Capítulo II), foi proposto por Carrell (1989) e traduzido para o português. Ele se subdivide em duas partes: parte A, introdutória, de identificação do sujeito e parte B, de fatores metacognitivos propriamente dita.

A parte A tem como objetivo a identificação do sujeito, ou seja itens 1, 2, 3 e 4 (nome, idade, RA e turma de inglês) e de sua proficiência lingüística e em leitura (itens 5, 6, 7 e 8). Embora muitas dessas informações já tenham sido solicitadas no questionário I, optou-se pela repetição para verificação, mais uma vez, da consistência das respostas dos sujeitos. Além disso, como o questionário tinha sido usado em outro estudo e traduzido para o português, não era desejável se introduzir mudanças e alterar sua estrutura original. Entretanto, em alguns itens, a informação foi solicitada de uma maneira diferente. É o que acontece com o item *leituras em inglês*, que no questionário traduzido é apresentado através de um item de múltipla escolha, para facilitar a correção.

A parte B do questionário consiste em 36 afirmações sobre leitura silenciosa em inglês. As respostas são dadas através de uma escala Likert de 1 a 5, onde 1 indica *acordo total* e 5 *desacordo total* com a afirmação.

O objetivo desse questionário é coletar informações sobre a percepção dos sujeitos com relação ao uso de estratégias de forma que se possa inferir, a partir das respostas nas categorias eficiência e dificuldade, a concepção de leitura dos sujeitos. Essa concepção pode ser *local* (se tiver como aspectos subjacentes som/letra/ palavra significado, sintaxe e detalhes de conteúdo) ou *global* (se tiver como aspectos subjacentes a idéia central, a organização do texto e o conhecimento prévio). O limite de tempo para a realização das três tarefas foi de uma hora.

3.2.2.3 – Testes de vocabulário

Uma bateria de quatro testes (9) de vocabulário foi desenvolvida visando uma medida abrangente (envolvendo extensão e profundidade) principalmente de uma das duas dimensões da competência lexical dos sujeitos, ou seja, de conhecimento, uma vez que nos três primeiros testes (com exceção do subteste de contextos/classes gramaticais) as palavras são apresentadas isoladamente, isto é, sem contexto. Entretanto, devo salientar que esses testes não constituem uma medida absoluta de conhecimento de vocabulário, ou da extensão total ou número de palavras conhecidas do vocabulário dos sujeitos, mas relativa, ou das palavras conhecidas/desconhecidas em um texto. A partir desse conhecimento espera-se fazer inferências ou generalizações com relação ao conhecimento de vocabulário em geral do sujeito.

A Tabela 3.3 abaixo apresenta os testes e seus focos de avaliação:

Tabela 3.3 – Testes de vocabulário e seus focos de avaliação

Testes de vocabulário	Focos de avaliação
Teste I (T1)	Familiaridade
Teste II (T2)	Frequência
Teste III (T3, T4, T5, T6, T7, T8)	Conceito <i>rico</i> de vocabulário
Teste IV (T9)	Inferência em contexto

O teste de vocabulário I (T1) ou de familiaridade (vide Apêndice B, Seção B.5) é uma auto-avaliação da familiaridade de palavras conduzida através de uma escala de verificação (*checklist*), baseada em Dale (1965). Algumas modificações foram introduzidas no que se refere aos níveis de conhecimento. As escalas em geral contém dois níveis de conhecimento: *sim/não* ou *conhece/não conhece a palavra*. Neste estudo, entretanto, incluo uma escala com quatro níveis. (vide Capítulo II para revisão dessa bibliografia).

São 30 as palavras que compõem esse teste — 20 reais (extraídas de um texto) e 10 inventadas — seguidas de uma escala de quatro níveis. A inclusão de palavras inventadas se justifica na medida em que constitui uma medida de consistência e também de controle contra possíveis adivinhações.

As palavras reais (comuns a todos os testes) foram extraídas de um texto de 350 palavras em inglês intitulado *Galactic Birth?*, extraído da revista *Time* de 11 de agosto de 1989. Esse texto também foi usado para o teste de leitura. O cruzamento dos resultados deste teste com os do vocabulário III que descrevo abaixo permite verificar a validade da escala de verificação enquanto medida de conhecimento de vocabulário.

O teste de vocabulário II (T2) ou teste de frequência (vide Apêndice B, Seção B.6) consiste na mesma lista de 30 palavras usada no teste de vocabulário I. Sua finalidade é verificar a percepção dos sujeitos com relação à frequência das palavras da lista. Espera-se, também, verificar a relação entre a frequência e familiaridade, uma vez que palavras mais frequentes tendem a ser mais familiares e vice-versa.

O teste de vocabulário III é composto por seis subtestes com finalidades, formatos e instruções diferentes (vide Apêndice B, Seções B.7, B.8, B.9, B.10 e B.11). O objetivo geral é uma medida abrangente dos vários aspectos que compõem o conhecimento de vocabulário de acordo com um conceito *rico* de conhecimento de vocabulário (Richards 1976). A bateria de seis testes é a operacionalização desse conceito (vide Capítulo II para revisão dessa bibliografia), tendo envolvido algumas alterações.

Todos os subtestes da bateria são elaborados a partir da lista de 20 palavras dos testes I e II, excluindo-se as palavras inventadas. A seleção das palavras para este teste de vocabulário, assim

como para os demais, foi baseada no procedimento reproduzido no Apêndice B, Seção B.2, que consiste na seleção das palavras consideradas desconhecidas por alunos de nível de proficiência semelhante ao dos sujeitos. As inventadas obedeceram dois critérios:

- palavras reais mudando-se uma letra ou duas; e
- combinações diferentes de radical mais afixo (*novel base-and-affix combinations*).

Para a elaboração desse teste, assim como para dos demais, partiu-se de um princípio corrente na área de avaliação, que relaciona extensão de um teste e sua confiabilidade, ou seja, testes mais longos são mais confiáveis. Entretanto, essa extensão não poderia vir a comprometer sua praticidade, introduzindo algum tipo de efeito fadiga. Procurou-se, portanto, um equilíbrio entre extensão, confiabilidade e praticidade. O limite de tempo de 5 minutos mostrou-se suficiente para a realização do teste. A Tabela 3.4, a seguir, apresenta esses subtestes e seus focos de avaliação:

Tabela 3.4. Testes elaborados de acordo com um conceito *rico* de conhecimento de vocabulário

Subteste	Focos de avaliação
III-a (T3)	Derivações (quando a palavra é apresentada em sua forma primitiva) Forma subjacente da palavra (no caso da palavra derivada) Flexões (singular/plural)
III-B (T4)	Significado das palavras Significados adicionais das palavras.
III-C (T5)	Associações
III-D (T6)	Colocações
III-E (T7 e T8)	Contextos gramaticais possíveis para cada palavra, ou seja, as relações gramaticais que cada palavra apresenta, sem uso de metalinguagem Classes gramaticais, usando-se metalinguagem

O subteste III-E, como pode-se notar pelo quadro acima, subdivide-se em duas partes. Na primeira, ou T7, o que realmente se deseja verificar é a função gramatical ou funções gramaticais de cada uma das palavras avaliadas dentro da oração (substantivo, adjetivo por exemplo) sem uso de metalinguagem, que o sujeito pode desconhecer. Para isso apresento uma lista (embora não exaustiva) de contextos possíveis e impossíveis, que foram mantidos a um nível mínimo, de forma a não influenciarem o resultado ou não permitirem inferências a nível de significado. O formato, portanto, difere dos demais testes uma vez que neste caso optou-se por um método de múltipla escolha. Para cada item há um número diferente de opções (*distractors*), que variam de 3 a 10.

Na segunda parte do subteste III-E, ou seja, T8, solicito o mesmo tipo de informação da primeira parte, ou seja, classe ou função gramatical da palavra em contexto. Entretanto, neste caso,

o conhecimento de metalinguagem é necessário, uma vez que o sujeito é solicitado a explicitar a(s) classe(s) gramatical(ais) de cada palavra.

O limite de tempo foi de uma hora para o teste de vocabulário II e subtestes III-A, B, C, D, e 30 minutos para o subteste E.

No teste de vocabulário IV (vide Apêndice B, Seção B.12) ou de inferência em contexto as 20 palavras dos testes anteriores são recolocadas no texto *Galactic Birth?* de onde foram extraídas, para serem inferidas. O tempo limite para a realização desse teste foi de 30 minutos.

3.2.2.4 – Teste de conhecimento conceitual

O teste de conhecimento conceitual (vide Apêndice C, Seção C.1) é composto por seis itens de respostas abertas, a serem respondidos em português, nos quais solicito definições ou explicações de seis conceitos selecionados do texto *Galactic Birth?*. Este teste constitui uma medida do conhecimento prévio dos sujeitos com relação aos conceitos envolvidos no texto e foi desenvolvido para que pudesse ser usado como uma variável moderadora.

Os conceitos são apresentados em português, o que poderia ser considerado uma limitação, vindo a fornecer pistas para a inferência de outras palavras e também para o entendimento do texto. Para se tentar controlar ou minimizar esse efeito houve um intervalo de mais de um mês entre esse teste e o teste de vocabulário IV, ou seja, vocabulário em contexto (vide Tabela 3.3 acima). Optei pela apresentação dos conceitos em português para que não houvesse uma interferência entre conhecimento/significado e conceito. O tempo limite para a realização do teste foi de 15 minutos.

3.2.2.5 – Testes de leitura

Antes de descrever os testes de leitura utilizados neste estudo, considero importante abordar algumas questões relativas à avaliação da leitura. Embora direcionada em grande parte pelos processos de significação e interpretação, ou de concepções de linguagem e de leitura, essa avaliação também é influenciada por fatores relativos ao método usado. Nesse sentido, a literatura em avaliação faz uma distinção entre a característica (*trait*) que se deseja avaliar e o método de avaliação.

Muitos métodos têm sido usados na avaliação da leitura. As perguntas de múltipla escolha, as perguntas de respostas abertas, os testes do tipo cloze e o resumo são alguns exemplos. Em geral esses métodos focalizam a compreensão enquanto produto da leitura, em detrimento do processo ou da busca da compreensão. Isso não quer dizer, entretanto, que também não possam ser direcionados para o processo. Não há dúvida que, nesse caso, são medidas indiretas desse processo, apresentando

limitações. A escolha do método é importante, uma vez que uma escolha errada pode exigir uma sobrecarga naquilo que se deseja avaliar. Perguntas do tipo múltipla-escolha, por exemplo, colocando geralmente uma única possibilidade de resposta, não permitem ao aluno justificar a opção escolhida. A elaboração de um resumo, por sua vez, também envolve habilidades de produção (vide Scaramucci 1990). Como métodos diferentes exigem do aluno processamentos diferentes, uma combinação de métodos seria a forma de se obter uma avaliação menos parcial ou menos enviesada.

Neste estudo são usados três métodos para avaliar a leitura: perguntas de respostas abertas, resumo e procedimento cloze. Dos três métodos, o último parece o mais controvertido, merecendo, portanto, uma consideração mais detalhada.

O procedimento cloze é uma alternativa metodológica em vigência já há alguns anos tanto em LM como em L2 e LE (vide Oller & Perkins 1969; Oller 1971, 1979; Oller & Conrad 1973; Hinofotis 1979; Alderson 1979; Shohamy 1982; Heilenman 1983; Deyes 1984; Fotos 1991; Abraham & Chapelle 1992 dentre outros, para uma revisão dessa literatura). O termo foi originalmente usado por Wilson Taylor (1953) para se referir a um índice de legibilidade de textos em LM: um painel de leitores determinava as proporções de apagamento de um texto que poderiam ser recuperadas. Dessa forma, seu objetivo era fornecer informações sobre o texto e não investigar a habilidade de processamento do leitor (Deyes op. cit.). Entretanto, mais tarde, reconhecido principalmente por Oller (1971) como um teste de habilidades globais ou integrador, passa a ter propósitos mais abrangentes, além da legibilidade, sendo considerado também como teste de leitura e de proficiência geral.

De fácil elaboração e correção, uma vez que consiste apenas na seleção de um texto, onde palavras são apagadas segundo um critério pré-determinado, deixando-se geralmente intacta a primeira oração, o cloze foi por muitos anos tido como um procedimento único e automaticamente válido ou a panacéia para todos os males da avaliação. A base dessa pesquisa eram estudos em que os resultados de testes do tipo cloze correlacionavam fortemente com um número grande de outras medidas de proficiência. Esses resultados também foram a base para se afirmar que o teste media a gramática de expectativa pragmática subjacente ao desempenho lingüístico. Em outras palavras, a habilidade de antecipar o significado das palavras através do contexto para o preenchimento das lacunas é usada para avaliar a proficiência na língua, uma vez que através da redundância do texto o leitor proficiente é capaz de fazer previsões baseadas em informação incompleta. Como os itens a serem recuperados se distribuem pelo texto, os testes do tipo cloze são ainda considerados testes de processamento ao nível de discurso. Como salienta Oller (1979:346),

Aparentemente itens do tipo cloze refletem uma compreensão global do texto. Não é todo item que é sensível a restrições de longo-alcance, mas um número suficiente de itens são sensíveis a tais restrições de maneira a afetar o desempenho global (...) é difícil imaginar alguém

completando as lacunas corretamente em um teste cloze sem entender os significados do texto no sentido de mapeá-lo ao contexto extralingüístico (...) (11c).

Entretanto, essa situação privilegiada do procedimento cloze foi logo desafiada, e pode-se dizer que continua sendo até hoje. Uma revisão dessa literatura parece indicar uma situação bastante controvertida.

Alderson (op. cit.), em trabalho de pesquisa que representa uma contribuição diferenciada à discussão sobre o teste cloze, mostrou que o procedimento, pelo menos em situações de LE, não produzia testes automaticamente válidos e confiáveis, uma vez que, tanto os procedimentos de apagamento quanto os de correção, assim como o texto escolhido, ou seja, variáveis metodológicas, resultavam em diferentes testes, medindo inclusive habilidades distintas. Para que o teste fosse válido, ele teria que ser validado como qualquer outro teste.

A propósito, são dois os procedimentos de apagamento geralmente usados na literatura: o apagamento fixo (*fixed-ratio method*), também denominado aleatório ou pseudo-aleatório (*pseudo-random deletion*) e o apagamento variável (*variable-ratio method*) também conhecido como apagamento racional (*rational deletion*). No primeiro caso, se apaga cada sexta palavra, ou cada oitava palavra, ou ainda cada segunda preposição ou sexta preposição, por exemplo. No segundo caso, as palavras são selecionadas de acordo com outros critérios, tais como apenas palavras que carregam significado, como substantivo, adjetivos, deixando de lado, por exemplo palavras de função.

São dois os métodos de correção mais comuns: o método da palavra exata (*exact-word method*), que aceita como correta apenas a palavra que foi apagada, deixando de lado erros menores de ortografia; e o método da palavra aceitável (*acceptable-word method*), mais comum em situações de LE, onde, como o nome indica, considera corretas outras palavras que podem ser sinônimos, palavras semanticamente aceitáveis ou contextualmente apropriadas. Outros procedimentos, tais como o escore da classe de palavras (*form class score*), em que são aceitas palavras que pertencem à mesma classe da palavra apagada; e o escore da clozentropia (*clozentropy score*, também conhecido como *communality of response score*), que dá crédito à respostas dadas por um grupo considerado um grupo de critério, geralmente falantes nativos da língua.

Os resultados de Alderson (op. cit.) indicam ser o texto uma variável importante, embora, devo salientar, o texto é sempre importante na avaliação, independentemente do método usado. Através de correlações de testes do tipo cloze com várias medidas externas de proficiência, o pesquisador mostrou que textos mais difíceis parecem fornecer uma melhor medida de proficiência global, enquanto textos mais fáceis têm como resultado correlações mais altas com o ditado.

Uma mudança no procedimento de correção também parece afetar os resultados. Nesse

caso, as correlações mais altas foram obtidas através do método da palavra aceitável. Comparando apagamentos pseudo-aleatórios de cada sexta, oitava, décima e décima segunda palavra, o estudo conclui que resultados diferentes são obtidos dependendo da taxa de apagamento escolhida. Esses resultados lhe permitiram concluir que, se o cloze é sensível ao apagamento de palavras individuais, pode não ser capaz de medir habilidades de nível mais alto, podendo ser considerado apenas uma medida de habilidades de nível mais baixo, e, portanto, mais uma medida de proficiência global do que de proficiência em leitura. Essa conclusão parece ser confirmada pelas correlações mais altas com testes de gramática e vocabulário do que com testes de compreensão em leitura.

Com base nesses resultados, a sugestão apresentada pelo autor é que

(...) o princípio da aleatoriedade deveria ser abandonado em favor de uma seleção racional de apagamentos, baseada em uma teoria sobre a natureza e processamento da linguagem. A noção de aleatoriedade pode ter sido justificada quando o objetivo do procedimento era caracterizar a dificuldade do texto, quando a seleção das dificuldades poderia ter dado a falsa impressão da natureza do texto. Mas agora que o foco é naquele que processa a linguagem ao invés da linguagem processada, agora que o procedimento cloze está sendo usado para produzir testes de compreensão escrita e de proficiência lingüística, o princípio não mais pareceria importante (12c).

A sugestão de Alderson, ou seja, substituir o procedimento pseudo-aleatório de apagamento de palavras por um outro que tem como critério o valor comunicativo dessas unidades ou sua recuperabilidade e relevância, permitiria, portanto, maximizar uma medida de sensibilidade a restrições de longo-alcance. Alguns trabalhos foram propostos seguindo essa sugestão. Na opinião de Deyes (op.cit.), por exemplo, um de seus proponentes, esses seriam verdadeiros clozes discursivos (*discourse clozes*).

3.2.2.5.1 – Cloze I

O teste denominado Cloze I (vide Apêndices D, Seções D.1 e D.2) foi elaborado a partir de um excerto de um texto de 540 palavras em português, extraído da revista *Isto É/Senhor* de 7 de julho de 1990, com 73 lacunas ou apagamentos, usando-se um procedimento de apagamento não-aleatório.

Os critérios que nortearam a seleção do texto foram:

- Tópico familiar e de interesse potencial do leitor.
- Texto bem organizado, apresentando uma extensão e grau de dificuldade que são compatíveis com a dificuldade do texto usado para o teste Cloze II, ou seja, a medida de proficiência em LE

e proficiência em leitura em LE. Não foi usado nenhum índice formal de dificuldade.

A finalidade do Cloze I é medir a proficiência de leitura dos sujeitos em sua LM, português. O teste foi pré-testado por três leitores de níveis diferentes de proficiência em leitura e algumas modificações foram realizadas em caráter de calibração. A primeira versão consistia em 64 itens, portanto, 9 itens a menos do que no teste usado. Alguns itens se mostraram muito fáceis, não discriminando leitores de níveis diferentes de proficiência, e outros muito difíceis, impossíveis de serem recuperados através do contexto, tendo sido substituídos. O critério para o apagamento das palavras obedeceu a uma seleção racional de apagamentos, seguindo a sugestão de Alderson (op. cit.) acima apresentada.

O tempo para a realização do teste foi de 35 minutos, que se mostrou suficiente. As instruções foram dadas oralmente e por escrito no quadro-negro.

3.2.2.5.2 – Cloze II

O teste denominado Cloze II (vide Apêndices D, Seções D.3 e D.4) foi elaborado a partir de um texto de 580 palavras, extraído da revista *Time*, consistindo de 36 itens-lacunas correspondentes a palavras, apagadas de acordo com um procedimento não aleatório, como no Cloze I.

Os critérios que nortearam a seleção do texto foram semelhantes aos do teste anterior:

- Tópico familiar e de interesse potencial do leitor.
- Texto bem organizado, apresentando um grau de dificuldade, que, na intuição do pesquisador, era compatível com o nível de dificuldade dos textos lidos durante as aulas ou típicos das aulas de Inglês Instrumental. Nesse sentido, é necessário lembrar que o pesquisador também é o professor de uma das turmas, de forma que a adequação do texto foi julgada em relação a uma aula regular.
- A extensão e nível de dificuldade o aproxima do texto usado para o teste Cloze I, ou seja, a medida de proficiência de leitura em português. Como no teste anterior, não foi usado nenhum índice formal de dificuldade.

O objetivo deste teste é uma medida aproximada de proficiência global em inglês e também uma medida de proficiência em leitura. Usando-se um procedimento de correção que considera corretas não apenas as palavras originalmente apagadas mas qualquer palavra que preserve o sentido (em princípio até mesmo palavras em português), tem-se também uma medida de proficiência em leitura (vide Capítulo IV para procedimentos de correção). O tempo permitido para a realização do teste foi de 45 minutos. Os demais procedimentos são iguais aos do Cloze I.

3.2.2.5.3 – Testes de leitura (resumo/perguntas de respostas abertas)

Foram elaborados dois testes de leitura: leitura I e leitura II. Para o teste de leitura I (vide Apêndice D, Seção D.5), o texto é o mesmo do teste de vocabulário IV. As vinte palavras selecionadas para os testes de vocabulário foram substituídas por palavras inventadas, usando-se o mesmo procedimento usado anteriormente para os testes de vocabulário. Dez dessas palavras foram as mesmas usadas nos testes de vocabulário I e II.

O teste é constituído de três partes:

- Parte I: Grifar palavras
- Parte II: Resumir
- Parte II: Responder a cinco perguntas de respostas abertas

Os objetivos desse teste são:

- Através da parte I, uma medida de todas as palavras desconhecidas do texto, além das já previstas pelos pré-testes e escolhidas para os testes de vocabulário.
- Através das partes II e III, medidas de compreensão de níveis diferentes, assumindo que pelo menos 20 palavras no texto são desconhecidas (palavras que foram substituídas por palavras inventadas, mais as outras grifadas na parte I deste teste). Convém lembrar que as palavras inventadas trazem consigo informações com relação à classe de palavras e flexões, que poderiam fornecer pistas ao sujeito com relação à função da palavra na sentença e também ao seu significado. O objetivo da tarefa de resumo é uma medida de compreensão da idéia geral do texto. Ele é elaborado e entregue ao pesquisador antes de o sujeito tomar conhecimento da tarefa da parte III, numa tentativa de se evitar que as perguntas dessa última tarefa pudessem direcionar a compreensão da idéia geral. Procurou-se estabelecer um propósito para a tarefa de resumo, de forma que o sujeito/leitor pudesse adequá-lo em termos de detalhes e extensão.

Níveis diferentes de compreensão são exigidos para as diferentes perguntas ou itens. O objetivo desse procedimento é verificar possíveis variações no desempenho em cada questão e tentar relacionar essas medidas à medidas de conhecimento de vocabulário. Uma competência lexical menos desenvolvida pode ser suficiente quando o objetivo é a compreensão da idéia geral, mas pode não ser suficiente quando as perguntas focalizam os detalhes do texto. O tempo para a realização deste teste foi de 45 minutos.

A Tabela 3.5 abaixo apresenta cada uma das questões ou itens da parte III, acompanhados de seus objetivos. Os itens foram elaborados de forma que pudessem avaliar problemas de uso da língua e não se limitassem a problemas puramente lingüísticos. Nesse sentido, não se restringiram apenas ao nível conceitual ou proposicional, ou na localização de informações explícitas e significados em orações isoladas, mas voltaram-se principalmente para o nível contextual e pragmático. As instruções com relação ao objetivo esperado para cada item foram claras para não

introduzirem nenhuma variável. Além disso, foram elaborados de maneira a possibilitar não apenas as respostas mas também suas justificativas.

Tabela 3.5 – Itens (questões) dos testes de leitura I e II e seus objetivos

Item	Objetivos
1	Tópico do texto, idéia principal. Relação título/texto.
2	Detalhes de apoio para o desenvolvimento da idéia principal.
3	Implicações da idéia principal. Complementação da idéia principal.
4	Detalhes de apoio. Aparentemente irrelevante à idéia principal. Entretanto, oferece a oportunidade de se observar a relação entre compreensão e conhecimento do conceito (Pedra de Rosetta), medido pelo teste de conhecimento conceitual.
5	Detalhes de apoio, exigindo nível de compreensão bem detalhado.

O teste de leitura II (vide Apêndice D, Seção D.6) consiste no mesmo teste de leitura I no que se refere às tarefas, diferindo apenas no texto, que, neste caso, é o texto original. Seu objetivo é uma medida da compreensão de leitura baseada na leitura do texto real. Essa medida é comparada aos resultados do teste de leitura I e teste de vocabulário I para verificar se a compreensão foi melhor no texto que traz as palavras reais, ou seja, o texto original.

Para que uma comparação de resultados fosse viável, houve a necessidade de se usar as mesmas questões do teste I. Esse procedimento apresenta duas desvantagens, ou seja, um acréscimo de familiaridade com o texto na segunda leitura, e conseqüentemente, desmotivação dos sujeitos.

O tempo limite para a realização do teste de leitura II foi exatamente igual ao do teste de leitura I.

3.2.2.6 – Escores do Vestibular

Como medidas adicionais e mais grosseiras, foram obtidos alguns escores que os sujeitos obtiveram no vestibular de entrada na universidade. Esses resultados têm que ser vistos com uma certa reserva, uma vez que não se tem informações dos critérios de correção adotados na época. Além disso, entre o vestibular e a coleta dos dados deste estudo houve um período de tempo (que varia entre os sujeitos, podendo ser de um a quatro semestres), que poderia ser responsável por alterações em sua proficiência, resultado da exposição à língua inglesa na universidade ou em escolas de línguas.

Foram obtidos dois tipos de escores: escores de primeira fase e de segunda fase. O escores da primeira fase apresentam uma medida do desempenho do candidato em produção escrita em LM. Embora essa produção seja responsável por 70% desse escore, ele também envolve questões de conhecimento geral.

Os escores da segunda fase referem-se ao desempenho do candidato em leitura/produção escrita em LM e em leitura em inglês como LE.

3.2.3 – Técnicas estatísticas

Três técnicas estatísticas descritivas foram usadas na análise dos dados. A análise de correspondência (AC doravante) foi utilizada principalmente nas variáveis de conhecimento de vocabulário, conhecimento conceitual e concepção de leitura. Depois dessa análise, que considerou cada variável separadamente, usou-se a análise de correlação para relacionar todas as variáveis do estudo. Na matriz de correlação resultante foi aplicada a análise fatorial.

Breve caracterizações de cada uma dessas técnicas são apresentadas nas seções a seguir.

3.2.3.1 – Análise de correspondência

A AC foi escolhida por ser o método mais apropriado para uma realidade multivariada resultante do desenho desta pesquisa, isto é, um número grande de variáveis categóricas com possíveis relações/associações.

Variáveis categóricas são aquelas em que o atributo está ou não presente. A preocupação nesse caso não é com *quanto*, mas sim com *quantos* ou *quão freqüente*. Daí o resultado apresentado em termos de freqüência, proporções, percentagens, e não em termos de escores. Para isso usa-se tabelas de freqüência, também denominadas tabelas de contingência ou ainda de freqüências cruzadas.

A AC é uma técnica de análise exploratória ainda pouco conhecida até mesmo dentro da estatística. Na França, essa técnica já foi utilizada na década de 60 por Benzècri para analisar uma tabela de consoantes e vogais finais da língua chinesa. Na área de LA não se tem notícia de nenhum estudo que se utilize dessa técnica.

O termo *correspondência*, nesse caso refere-se a um sistema de associações entre os elementos de dois conjuntos constituídos pelas categorias da tabela (linhas e colunas). Através dessa análise pode-se gerar hipóteses, uma vez que ela fornece uma idéia da associação entre as variáveis envolvidas, o que é especialmente útil quando as tabelas são grandes. Com experimentos confirmatórios posteriores, pode-se testar as hipóteses, empregando as metodologias usuais de estatística inferencial.

Um conjunto de variáveis categóricas em que muitas categorias estão envolvidas geralmente contém várias inter-relações, impossíveis de se interpretar à primeira vista. A abordagem empregada

na AC se baseia em representações geométricas, revelando a estrutura dos dados de forma ótima, sem a necessidade de se assumir modelos nem distribuições fundamentais. Assim, a AC é considerada um algoritmo de redução de dados que fornece imagens simplificadas da realidade multidimensional. Ela busca a melhor representação simultânea indicativa da dependência de dois conjuntos de categorias, compreendidos pelas linhas e colunas da tabela de contingência. Ela pode ser comparada a um aparelho radiográfico, que fornece imagens a partir de uma realidade não observável. Para obter-se um diagnóstico preciso, há necessidade não só de um preparo prévio do paciente (ingerindo medicamentos de contraste, por exemplo) como também do posicionamento do aparelho no paciente da forma mais conveniente possível, ou seja, que melhor focalize os órgãos ou ossos de interesse. A interpretação dos resultados está ligada ao conhecimento dos princípios de funcionamento do aparelho e à opacidade dos raios X, dependente, por sua vez, da densidade, do volume e da composição química dos órgãos, alterada a partir da absorção do medicamento (Villalobos-Aguayo 1993).

Podem-se distinguir vários tipos de AC: AC simples, AC múltipla (ACM) ou de matrizes indicadoras e/ou matrizes de Burt, e AC generalizada (ACG). Atualmente, a AC simples é conhecida por diferentes nomes, devido aos contextos em que foram criadas: ponderação recíproca, *dual* ou *optimal scaling*, análise de correlação canônica e regressão linear simultânea. Neste trabalho foram usadas AC simples e múltiplas.

Aparentemente, a curiosa história dessa técnica começa em 1935, quando Hirschfeld (1935) publicou um artigo intitulado *Uma conexão entre correlação e contingência* (*A connection between correlation and contingency*). Nesse artigo, ele propõe uma solução algébrica ou em termos de uma equação para a correlação entre linhas e colunas de uma tabela de contingência. Na mesma época, idéias similares usando abordagens diferentes foram sugeridas por Richardson & Kuder (1933) e depois por Horst (1935), que acabou criando o Método da média recíproca (*Method of reciprocal averaging*) para aplicação na psicometria. Posteriormente, Fisher (1940) desenvolveu a teoria fundamental e propôs uma análise discriminatória para tabelas de contingência. Por outro lado, Guttman (1941) desenvolveu um método para construir escalas para dados categóricos e que, partindo de um contexto diferente, resultou num método equivalente ao da AC. Dessa maneira, pode-se dizer que a AC foi redescoberta em diferentes circunstâncias mas foi logo esquecida. Sua popularização somente teve lugar em meados da década de 1960, graças a um grupo de franceses liderados por Jean-Paul Benzècri. O grupo desenvolveu o método como uma técnica gráfica multidimensional, cujas implicações geométricas facilitariam a interpretação. Deu ao método o nome de *Análise fatorial de correspondências* (*Analyse factorielle des correspondances*). Pouco depois, o nome foi simplificado para *Análise de correspondências* (*Analyse des correspondances*).

Como sugere Greenacre (1984),

(...) uma descrição gráfica é mais facilmente assimilada e interpretada do que uma numérica e pode ter as três funções (...) resumir uma grande massa de dados numéricos, simplificar o aspecto dos dados apelando para nossa habilidade natural de absorver imagens visuais, e, possivelmente fornecer uma visão geral da informação, e, por conseguinte, estimular explicações possíveis (...) aumentando a importância da análise de dados exploratória (...) (13c).

A partir daí, a literatura no assunto passa a ser abundante (vide Lebart & Fenelon 1973; Benzècri 1973; Hill 1974; Nishisato 1980; Verdinelli 1980; Gifi 1981; Greenacre 1984; Souza 1982, 1990; Maquera Sosa 1989; Villalobos-Aguayo 1993). Desde 1986 existe um periódico intitulado *Cadernos de Análise de Dados (Les Cahiers de l'Analyse des Données)* dedicado quase que exclusivamente às aplicações da AC.

3.2.3.2 – Análise de correlação

A análise de correlação é um método bastante comum, usado na maior parte das pesquisas em Ciências Sociais. Nesses estudos, o interesse é determinar-se o grau de relação entre pares de duas ou mais variáveis. A estatística usada nesse caso é chamada coeficiente de correlação. Se escores altos de uma variável estão associados a escores altos de outra variável, tem-se uma correlação positiva entre as variáveis. Se escores altos de uma variável estão associados a escores baixos de outra variável, tem-se uma correlação negativa. Se não há um padrão sistemático entre escores altos e baixos, não se tem relação entre dois conjuntos de escores.

3.2.3.3 – Análise fatorial

A análise fatorial é uma técnica para análise de escores em termos de fatores subjacentes. Basicamente, a técnica se baseia na pressuposição de que, em qualquer teste ou escore, há provavelmente um ou mais traços subjacentes que estão sendo avaliados. O objetivo é construir esses fatores subjacentes e decompor as variâncias do escore em termos da correlação dos fatores e dos escores observados. Em outras palavras, a análise fatorial fornece informação sobre os fatores subjacentes de variáveis através de um exame da variância comum entre essas variáveis. Após obter-se os agrupamentos, é necessário encontrar um rótulo para os fatores gerados pela análise.

3.3 – A fase II: o estudo de processo com ênfase qualitativa

O foco da pesquisa nesta fase é o processo de compreensão de leitura de sujeitos/leitores de níveis diferentes de proficiência em leitura e vocabulário (caracterizados a partir de escores de leitura, de vocabulário, questionário metacognitivo e outros questionários, conduzidos na fase I.

3.3.1 – Seleção e descrição dos sujeitos

Participaram desta fase três sujeitos, apresentando níveis de leitura caracterizados como intermediário avançado, intermediário e iniciante, selecionados a partir de seu desempenho no teste de leitura cloze em LE e escores do vestibular usados na fase quantitativa acima mencionada. Para essa seleção, além do critério teórico, ou seja, necessidade de sujeitos de níveis diferentes de proficiência em leitura, adotou-se também um critério prático ou a disponibilidade dos sujeitos para a coleta dos dados, uma vez que ela não foi conduzida durante as aulas, mas em sessões especiais. Um perfil detalhado de cada um dos sujeitos é apresentado no início do Capítulo V, mais especificamente na Seção 5.1.

3.3.2 – Descrição dos procedimentos e instrumentos de pesquisa

Uma técnica mentalista, mais especificamente o protocolo de pausa (Cavalcanti 1983, 1989), foi escolhido para a elicitación dos dados. Essa técnica consiste em solicitar aos sujeitos que leiam silenciosamente (monitorando seu processo de leitura), com o objetivo de fazer um resumo oral, e pensem alto toda vez que observarem pausas no seu processo. O protocolo de pausa é uma adaptação do protocolo verbal, usado por Olshavsky (1976-7) e Kavale & Schreiner (1979) em LM e Hosenfeld (1984) em LE. O protocolo verbal, por sua vez, é uma técnica de auto-revelação.

De acordo com Cohen (1987) as técnicas mentalistas podem ser classificadas em três categorias: auto-relato, auto-observação e auto-revelação.

O auto-relato refere-se às afirmações do sujeito sobre seu próprio comportamento enquanto lê ou aprende, e, como tais, refletem os conceitos que tem de seu comportamento enquanto leitor e aprendiz. Tais afirmações são generalizadas, não se baseando em nenhum acontecimento lingüístico específico.

A auto-observação, por outro lado, diz respeito às afirmações que o sujeito faz de seu comportamento a partir de algum acontecimento lingüístico específico e que pode se dar de duas maneiras: introspectivamente ou retrospectivamente. A auto-observação introspectiva ocorre

enquanto a informação ainda se encontra na memória de curto prazo, enquanto a auto-observação retrospectiva ocorre geralmente depois de vinte minutos. A retrospectiva pode ainda ser imediata (por exemplo, uma hora depois do acontecimento) ou retardada (algumas horas ou mesmo semanas depois).

A auto-revelação, por sua vez, refere-se às afirmações que o sujeito faz enquanto processa suas informações, consistindo numa espécie de *pensar alto* sem a preocupação de editar ou analisar os dados.

Apesar de algumas limitações, normalmente associadas à familiaridade com a tarefa, vontade de cooperar e motivação dos sujeitos, e confiabilidade dos dados (vide Cavalcanti 1983, 1989) para uma análise mais detalhada dessas técnicas), o protocolo de pausa foi escolhido por possibilitar uma maior aproximação ao processo de leitura dos sujeitos. Na realidade, ele permite aos sujeitos relatarem seu comportamento, e não propriamente o seu processo. Este é inferido e analisado posteriormente pelo pesquisador, evitando-se assim dados dependentes da descrição e análise verbal do sujeito, que não seriam necessariamente dados de seu processo mental.

3.3.3 – Coleta de dados

A coleta de dados foi precedida de uma fase de aquecimento (Cavalcanti 1989), cujo objetivo era familiarizar os sujeitos com a técnica e orientá-los a focalizar em pausas (vide Apêndice E, Seção E.1, para o texto usado).

Antes de proceder à primeira tarefa (Tarefa I), os leitores foram solicitados a ler e analisar o título do texto, o que possibilitaria um controle sobre o conhecimento ou familiaridade com o tópico assim como uma antecipação do seu conteúdo.

Como Tarefa I, os leitores foram solicitados a ler o texto-base silenciosamente, com o objetivo de fazerem um resumo oral, e pensar alto toda vez que observassem uma pausa ou problema em seu processo de leitura. Uma vez localizado o problema, a próxima etapa seria tentar descrevê-lo e solucioná-lo, sempre em voz alta. Ao final de cada parágrafo (sinalizados por setas), o leitor deveria elaborar um resumo oral.

O objetivo da Tarefa II era confirmar as palavras não inferidas durante a primeira tarefa. Foi solicitado aos leitores que lessem mais uma vez o texto e grifassem as palavras.

Os dados foram coletados durante uma sessão de aproximadamente uma hora e meia, incluindo a fase de aquecimento. Para a elicitação dos dados, foi usado um mini gravador *cassette* e as instruções foram dadas oralmente. A pesquisadora decidiu não permanecer no local durante toda a coleta, por perceber que, durante a fase de aquecimento, sua presença parecia inibir os sujeitos, além de encorajar uma interação leitor-texto-pesquisadora. Entretanto, é necessário observar que

esteve presente em alguns momentos dessa coleta, para se certificar de que a tarefa estava sendo elaborada de acordo.

O texto-base (vide Apêndice E, Seção E.2) consistiu em um artigo extraído da revista *New Scientist* de 1 de setembro de 1990 cujo tema é Ecologia na União Soviética, com aproximadamente 700 palavras, apresentando uma estrutura problema-solução (Hoey 1979). Os critérios que nortearam sua seleção foram os mesmos usados para a seleção de textos na fase I, ou seja:

- Tópico familiar e de interesse do leitor. O tema já havia sido discutido durante as aulas e despertado o interesse dos alunos;
- Texto bem organizado, apresentando um grau de dificuldade que possivelmente apresentaria alguns desafios para os sujeitos, principalmente para os de nível intermediário e iniciante.

O objetivo deste capítulo foi apresentar o desenho e a metodologia de pesquisa usados neste estudo, assim como sua justificativa. Para isso, abordou questões relativas às dicotomias produto/processo e qualitativa/quantitativa, uma vez que o desenho pressupõe dois estudos, um de ênfase quantitativa e foco no produto e outro de ênfase qualitativa, com foco no processo. Num segundo momento, focalizou cada um dos estudos individualmente, apresentando descrições dos sujeitos assim como dos instrumentos e procedimentos para a coleta dos dados. O próximo capítulo apresenta a análise e discussão dos dados de ênfase quantitativa, ou com foco na compreensão enquanto produto da leitura.

Notas

(1) Na avaliação de pesquisas pode-se observar uma tradição bastante grande do paradigma quantitativo. Essa tradição, entretanto, também mostra suas falhas e viéses, levando à insatisfação com esse método e a conseqüente e mais recente popularidade dos métodos qualitativos.

(2) O termo em inglês é *survey*.

(3) O termo *paroquialismo*, tradução do inglês *parochialism*, embora possa ser considerado um anglicismo, está sendo usado aqui na falta de um termo melhor em português.

(4) *Tarefa* é um conceito corrente em LA, também usado em avaliação. Esses autores não são seus proponentes, mas apenas se utilizam dele para o desenvolvimento de suas atividades.

(5) Esse mesmo tipo de conceitualização também poderia ser usado para os outros tipos de conhecimentos envolvidos na leitura, como por exemplo, conhecimento prévio.

(6) Análise de erros, neste contexto, diz respeito à exploração de alguns dos erros dos sujeitos como um recurso para tentar recuperar elementos de seu processo, não se referindo à análise de erros usada no período estruturalista.

(7) Embora a presença dos alunos em aula fosse alta, alguns ocasionalmente deixaram de comparecer nos dias em que os dados foram coletados. Assim, alguns dos testes tiveram que ser ministrados em sessões individuais.

(8) RA (registro acadêmico) é um código de seis algarismos usado na universidade para identificação dos alunos. Enquanto os quatro últimos algarismos identificam o aluno, os dois primeiros indicam seu ano de ingresso.

(9) Estou usando a denominação *testes* para me referir aos instrumentos de avaliação usados na coleta de dados, instrumentos esses que apresentam os mais variados métodos e não são apenas do tipo múltipla escolha.

Citações originais

(1c) *"Ultimately, the issue is not research strategies, per se. Rather, the adherence to one paradigm as opposed to another predisposes one to view the world and the events within it in profoundly different ways".*

(2c) *"... the choice of methods should not be determined by allegiance to an arbitrary paradigm. This is both because a paradigm is not inherently linked to a set of methods and because the characteristics of the specific research setting are equally important as the attributes of a paradigm in choosing a method".*

(3c) *"In survey research and in large-scale experiments, one collects relatively few bits of information about many research subjects, whereas in interpretive research, one collects many bits of information about relatively few subjects".*

(4c) *"In survey and experimental research, one usually conceives of cause in ways akin to those held in the natural sciences, that is, a physical notion of cause. In interpretive research one is not so much interested in cause per se as one is interested in explanation in terms of meanings and understandings held by those who one has studied".*

(5c) *"... it is now time to stop the pendulum from swinging from one extreme to the other. It is time to stop building walls between the methods and start building bridges. Perhaps it is even time to go beyond the dialectic language of qualitative and quantitative methods. The real challenge is to fit the research methods to the evaluation problem without parochialism".*

(6c) *"The first meaning of process is monitoring: describing the context and population of the study, discovering the extent to which the treatment or program has been implemented, providing immediate feedback of a formative type, and the like. The second meaning of process is causal explanation: discovering or confirming the process by which the treatment had the effect that it did. Of course, measuring the effect of the program is the summative outcome or impact assessment. For a complete understanding of the program, an evaluation would have to accomplish at least these three tasks: monitoring, impact assessment, and causal explanation".*

(7c) *"A longitudinal approach ... typically involves observing the development of linguistic performance, usually the spontaneous speech of one subject, when the speech data are collected at periodic intervals over a span of time. In a cross-sectional approach the linguistic performance data are usually collected at only one section".*

(8c) *"... the most serious problem is that product-oriented and process-oriented studies are too often viewed as competing rather than complementary approaches to classroom-centred research. As a result, we are left with product-oriented studies without a product component. Clearly, there is a need for both in investigations of teaching and learning in classroom settings".*

(9c) *"... what is important for researchers is not the choice of a priori paradigms or even methodologies, but rather to be clear on what the purpose of the study is and to match that purpose with the attributes most likely to accomplish it. Put another way, the methodological design should be determined by the research question".*

(10c) *"The basic idea of an experiment is a powerful one. If one group of subjects is treated in one fashion*

and another in a different fashion, and there are no other factors influencing the two groups differentially, a cause-effect relationship between treatment and consequence can be determined".

(11c) "Apparently cloze items reflect overall comprehension of a text. Not every item is sensitive to long-range constraints, but enough items apparently are sensitive to such constraints to affect overall performance ... it is difficult to imagine anyone filling in the blanks on a cloze test correctly without understanding the meanings of the text in the sense of mapping it onto extralinguistic context ...".

(12c) "... the principle of randomness needs to be abandoned in favor of the rational selection of deletions, based upon a theory of the nature of language and language processing. The notion of randomness may have been justified when the aim of the procedure was to characterize the difficulty of text, when the selection of difficulties could have given a false impression of the nature of a text. But now that the focus is on the language processor rather than on the language being processed, now that the cloze procedure is being used to produce tests of reading comprehension and language proficiency, the principle would appear to be important no longer".

(13c) "... a graphical description is more easily assimilated and interpreted than a numerical one and can assist all three functions ... summarizing a large mass of numerical data, simplifying the aspect of the data by appealing our natural ability to absorb visual images, and (hopefully) providing a global view of the information, thereby stimulating possible explanations ... rising importance of exploratory data analysis ...".

CAPÍTULO IV

O VOCABULÁRIO E A COMPREENSÃO ENQUANTO PRODUTO DA LEITURA: ANÁLISE E DISCUSSÃO DOS DADOS DE ÊNFASE QUANTITATIVA

Nos Capítulos II e III, procurei construir uma base teórica e metodológica para esta investigação. Neste capítulo, estruturado em três seções principais, apresento e discuto os dados da análise de ênfase quantitativa, ou o estudo com foco no produto da leitura. A Seção 4.1 introduz os procedimentos usados na correção dos testes e na preparação desses dados e dos questionários, para a análise. A Seção 4.2 focaliza a análise propriamente dita, subdividida em duas seções: a Subseção 4.2.1 apresenta a análise intracomponentes, cujo objetivo é cada um dos componentes ou variáveis principais; e a Subseção 4.2.2, que apresenta a análise intercomponentes, cujo objetivo é a interação entre eles. As duas análises são discutidas conjuntamente na seção 4.3.

Embora o objetivo deste estudo seja investigar a relação entre a competência lexical e a compreensão em leitura em LE, a adoção de um modelo interativo sugere que essa relação não pode ignorar a contribuição de outros fatores que interagem nessa relação. Assim, esta investigação, focalizando dois momentos, busca caracterizar:

- Com base nos resultados de testes e questionários, o desempenho dos seus 49 sujeitos não apenas com relação à compreensão em leitura em LE e competência lexical, variáveis dependente e independente, respectivamente, mas também com relação a cada um dos outros componentes ou fatores que, segundo a literatura, parecem contribuir para a compreensão, ou seja, a proficiência em leitura em LM, o conhecimento conceitual, a proficiência em LE assim como as concepções de leitura do leitor. As perguntas que orientam essa parte da pesquisa são explicitadas, quando for o caso, no início da seção referente a cada componente;
- Com base nos resultados desses testes, a natureza dessa relação.

Dessa forma, tem-se uma análise intracomponentes e uma análise intercomponentes, buscando respostas para as seguintes questões:

1. Qual a relação entre a competência lexical e a compreensão enquanto produto da leitura?
2. Qual a contribuição, para essa relação, de fatores como proficiência em LE, diferentes concepções de leitura em LE, conhecimento conceitual e leitura em LM?
3. Há evidências de um nível limiar de competência lexical para a compreensão enquanto produto

da leitura?

4.1 – Preparação dos dados para a análise

Os instrumentos usados para a coleta dos dados, neste estudo, são testes e questionários. Os procedimentos utilizados na correção dos testes são variados, dependendo da natureza de cada instrumento. Esses procedimentos, assim como outros, são apresentados nas próximas seções.

Os resultados dessa correção, juntamente com os dados obtidos através dos questionários passam a constituir bancos de dados para que a análise possa ser feita usando-se um sistema estatístico computacional, no caso, o sistema SAS. São sete os bancos de dados, assim denominados: *Perfil*, *Vocabulário*, *Conhecimento conceitual*, *Proficiência em LE*, *Metacognição/Concepções de leitura em LE*, *Leitura em LM e em LE*. As próximas seções detalham cada um deles.

4.1.1 – Perfil

Os dados obtidos através dos questionários I e II, e parte A do questionário III (vide Apêndices A, Seções A.1, A.2 e A.5) permitem uma caracterização mais completa do grupo de sujeitos que participaram deste estudo, complementando a Seção 3.2.1 do Capítulo III no que diz respeito à idade, ano de ingresso, escola de origem, como também *background* lingüístico em inglês, leitura em português e inglês e dificuldades encontradas na leitura em inglês.

Alguns itens dos questionários eram redundantes, propositalmente repetidos para verificação da consistência na resposta dos sujeitos. Como os resultados mostraram respostas consistentes, esses itens foram, portanto, codificados uma só vez. As respostas em aberto também foram codificadas (vide Apêndice A, Seção A.4).

4.1.2 – Vocabulário

Os testes de vocabulário foram corrigidos usando-se procedimentos variados de acordo com a natureza de cada um. Nos testes de vocabulário I (T1 ou familiaridade) e vocabulário II (T2 ou frequência) (vide Apêndice B, Seções B.5 e B.6), a correção foi objetiva, isto é, consistiu apenas na contagem do número de itens em cada categoria, uma vez que eram itens de múltipla escolha. As categorias, portanto, permaneceram as mesmas dos testes originais. Entretanto, nos outros testes, houve necessidade de recodificações e recategorizações, uma vez que, com exceção do subteste de

vocabulário III-E (T7/T8 ou contexto/classe gramatical), todos eram de respostas abertas.

O teste de vocabulário III (conhecimento *rico* de vocabulário) subdivide-se em seis subtestes (vide Apêndice B, Seções B.7, B.8, B.9, B.10 e B.11, assim Capítulo III). No subteste de vocabulário III-A (T3 ou flexão/derivação), a correção consistiu na contagem do número de acertos de cada sujeito nos itens 1, 2 e 3. O sujeito teve um acerto cada vez que forneceu pelo menos uma palavra derivada (item 1), ou uma palavra primitiva (ou que deu origem a uma derivada (item 2) ou ainda o singular ou plural corretamente (item 3), dependendo do caso.

No subteste de vocabulário III-B (T4 ou definição/significado), foram considerados acertos todas as definições ou explicações e sinônimos em inglês ou português corretos.

Nos subtestes de vocabulário III-C (T5 ou associação) e III-E (T6 ou colocação), contou-se o número de associações corretas para cada palavra.

No subteste de vocabulário III-D (T7/T8 ou contexto/classe gramatical), antes da correção dividiu-se o teste em duas partes, e os resultados foram codificados em termos de contextos e classes gramaticais. A correção teve como critério respostas de um grupo de doze falantes nativos, constituído por professores e alunos de pós-graduação em lingüística, uma vez que divergências são encontradas até mesmo entre falantes nativos. Optou-se, assim, por considerar corretos apenas os contextos apontados corretos por pelo menos metade dos falantes nativos, ou seja, por seis informantes.

No teste de vocabulário IV (vide Apêndice B, Seção B.12) a correção consistiu na contagem das palavras inferidas/não inferidas.

Depois de corrigidos e recodificados, os resultados dos quatro testes principais e seus subtestes passaram para um banco de dados denominado *Vocabulário* (vide Apêndice B, Seção B.13, para os testes categorizados para a análise), conforme mostra a Tabela 4.1 abaixo:

Tabela 4.1 – Vocabulário

Testes/Subtestes	Foco de avaliação
T1	Familiaridade
T2	Frequência
T3, T4, T5, T6, T7 e T8	Conhecimento <i>rico</i> de vocabulário
T9	Inferência em contexto

4.1.3 – Conhecimento conceitual ou prévio

O teste de conhecimento conceitual foi corrigido e as respostas a cada um dos seis itens codificadas. Os resultados passaram a constituir um banco de dados denominado *Conhecimento conceitual*, com seis conceitos (vide Apêndice C, Seção C.2 para teste categorizado para a análise).

4.1.4 – Metacognição/concepções de leitura

Para a verificação das concepções de leitura dos sujeitos foi usado o questionário metacognitivo (Carrell 1989). Além dessa informação, o questionário também mostra evidências relacionadas à percepção de uso de estratégias. Foi usada, neste caso, a codificação original do questionário, ou seja, cada item ou afirmação foi avaliado através de uma escala do tipo Likert de 5 níveis, onde o nível 1 indica *concordo plenamente* e 5 *discordo plenamente*. As 36 afirmações ou itens do questionário (vide Apêndice A, Seções A.5 e A.6) apresentam-se subdivididos em quatro grupos ou categorias, conforme a Tabela 4.2:

Tabela 4.2 – Metacognição/Concepções de leitura

Categorias	Afirmações/estratégias
Confiança	6 (conf1-conf6)
Correção/reparo	5 (rep7-rep11)
Eficiência	17 (efi12-efi20) e (efi20-efi36)
Dificuldade	8 (dif21-dif28)

4.1.5 – Leitura em LM e LE e proficiência em LE

Os testes Cloze I e Cloze II, em português e em inglês, respectivamente, foram corrigidos usando-se procedimentos variados (vide Apêndice D, Seções D.2 e D.4).

No Cloze I em português (LM), constituído por 72 itens/lacunas, adotou-se dois procedimentos de correção, resultando, portanto, em dois escores diferentes, indicativos do número de acertos, e em duas medidas de leitura em LM:

- *Palavra exata*. Considerado correto somente o item que trouxer a palavra exata do texto original. Essa versão, denominada Cloze1A, ignora erros menores de ortografia.
- *Palavra aceitável*. Considerados corretos todos os itens que trouxerem palavras semanticamente aceitáveis naquele contexto, e não necessariamente as palavras do texto original. Essa versão, denominada Cloze1B, também ignora erros menores de ortografia.

Para a correção do Cloze II em inglês, constituído de 36 itens/lacunas, foram usados três procedimentos, resultando em três escores diferentes, indicativos do número de acertos, e em três medidas distintas:

- *Palavra exata* ou Cloze2A, desenhado e corrigido da mesma forma que o cloze em português, constituindo numa medida aproximada de proficiência em LE.
- *Palavra aceitável* em inglês ou Cloze2B, elaborado da mesma forma que o cloze em português.

O critério para definir *palavra aceitável* foi estabelecido a partir das respostas do mesmo grupo de falantes nativos usado para o teste de vocabulário. Esse tipo de procedimento, como já salientei no Capítulo III, chamado de *escore da clozentrópia* (ou *clozentropy score* ou *communality of response score*), é embasado em um grupo que serve de critério ou de referência, e tem sido considerado pela literatura como um critério razoável para se julgar a adequação das respostas dadas por falantes não-nativos. Este teste, com esse critério de correção, consiste, como no anterior, numa medida aproximada de proficiência em inglês.

- *Palavra aceitável* em inglês ou Cloze2C. A diferença entre este teste e o anterior (Cloze2B) é que, neste caso, também foram aceitas palavras em português, desde que o sentido fosse preservado. Erros de colocação e gramática (ausência de *s* na terceira pessoa do singular, por exemplo) foram ignorados. Esse teste, com esse critério de correção, fornece uma medida de proficiência em leitura em inglês. Não foram consideradas respostas certas pela metade, mas somente certas e erradas.

Convém lembrar que o teste de leitura I inclui três partes (I, II e III) enquanto o teste de leitura II inclui apenas duas (II e III) (vide Apêndice D, Seções D.5 e D.6). Na parte I, em que é solicitado aos sujeitos grifarem as palavras desconhecidas, a correção constou na contagem dessas palavras. Na parte II, cujo objetivo é a elaboração de um resumo, foram dadas notas de 0 a 100 em separado (Resumo1) e (Resumo2). A parte III, ou as perguntas de compreensão propriamente ditas foram corrigidas também de acordo com uma escala de notas de 0 a 100. Pesos diferentes foram dados a cada um dos cinco itens ou questões, de acordo com a relevância da informação no teste e o nível de dificuldade, conforme mostra a Tabela 4.3.

Para se contrapor ao resumo (idéia geral), foram dadas notas (de 1 a 25) em separado ao item 5 do teste de leitura, uma vez que o objetivo da questão era verificar a compreensão de detalhes. Assim, mais duas variáveis foram criadas, ou seja: (Detalhe1) e (Detalhe2).

Os resultados dos (Cloze1A) e do (Cloze1B), em português e do (Cloze2C), em inglês, dos testes de (Leit1) e (Leit2) e suas subpartes, assim como os escores do vestibular (FaseV1, LeitV2 e LeitV3) passaram a constituir um banco de dados denominado *Leitura*, conforme a Tabela 4.4.

Tabela 4.3 – Valores dos itens dos testes de leitura I e II

Itens	Valores (pontos)
1	20
2	20
3	20
4	15
5	25

Tabela 4.4 – Leitura

(a) Testes cloze

Teste	Objetivo	Número de itens	Critério de correção	Escore
CLOZE1A	Leitura em LM	72	Palavra exata	0-72
CLOZE1B	Leitura em LM	72	Palavra aceitável	0-72
CLOZE2C	Leitura em LE	36	Palavra aceitável	0-36

(b) Testes de leitura em LE

Teste	Número de itens	Critério de correção	Escore
LEIT1	5	Palavras inventadas	0-100
LEIT2	5	Original	0-100
RESUMO1		Palavras inventadas	0-100
RESUMO2		Original	0-100
DETALHE1	1	Palavras inventadas	0-25
DETALHE2	1	Original	0-25

(c) Escores no exame vestibular

Teste	Objetivos	Escore
FASEV1	Redação em LM e conhecimentos variados	0-100
LEITV1	Leitura e produção escrita em LM	0-100
LEITV3	Leitura em LE	0-100

Tabela 4.5 – Proficiência em LE

Teste	Objetivo	Número de itens	Critério de correção
CLOZE2A	Proficiência em LE	36	Palavra exata
CLOZE2B	Proficiência em LE	36	Palavra falante nativo

Os resultados dos testes Cloze2A e Cloze2B passaram a constituir um banco de dados denominado *Proficiência em LE* (Tabela 4.5). Finalmente, os escores de todos os testes foram transformados em uma escala de 0 a 10 para que pudessem ser comparados.

4.2 – A análise dos dados

A análise dos dados deste estudo foi realizada em duas etapas. Em um primeiro momento, procurou-se analisar cada banco de dados ou as variáveis intracomponentes. Essa análise foi seguida pela análise intercomponentes, que levou em conta todas as variáveis envolvidas no estudo de produto.

4.2.1 – A análise intracomponentes

São sete as variáveis desta análise, como mostrei acima: *Perfil*, *Vocabulário*, *Conhecimento conceitual ou prévio*, *Metacognição/Concepções de leitura*, *Leitura em LM e LE* e *Proficiência em LE*. As subseções a seguir apresentam um detalhamento dessa análise, focalizando cada um dos componentes separadamente.

4.2.1.1 – Perfil

Os resultados do Perfil, subdivididos em três partes, têm por objetivo servir de pano de fundo para o estudo dos demais componentes. Na Tabela 4.6 são apresentados os dados de identificação e de *background* lingüístico do grupo; na Tabela 4.7 basicamente os resultados dos testes de leitura em LM e em LE; e, finalmente, as auto-avaliações e avaliações do pesquisador são exibidas na Tabela 4.8.

Como mostra a Tabela 4.6, os sujeitos deste estudo são 49 estudantes de graduação de uma universidade estadual paulista, com idade variando de 16 a 25 anos. São todos da área de Exatas distribuídos entre diferentes cursos, tendo ingressado na universidade de 1987 a 1990. A grande maioria já estudou inglês há mais de 5 anos, tendo sido esse contato com a língua inglesa, para mais da metade dos sujeitos, realizado também em escolas de línguas.

No que se refere a leitura, observa-se, a partir da Tabela 4.7, que a grande maioria (43) afirma ler em inglês, mas apenas (3) lêem em outras línguas. Cruzando-se esse dados com os das tabelas de auto-avaliação abaixo, nota-se que, embora todos afirmem ler em inglês, somente 7 se consideram bons leitores, enquanto são 25 maus leitores. Tanto a leitura em português como em inglês da maioria dos sujeitos focaliza textos variados, embora apenas 7 façam leituras de textos técnicos em inglês de sua área de especialização. Surpreendentemente, 12 afirmam não ler nada em inglês.

A Tabela 4.8 mostra que, na tarefa de leitura e auto-avaliação da idéia geral realizada antes de responder ao questionário, 29 sujeitos afirmaram ter conseguido uma boa compreensão do texto, considerado de dificuldade média por 33; a avaliação do pesquisador confirma de forma bastante precisa esses dados, mostrando que os leitores realmente têm uma idéia clara de sua compreensão. Quanto à dificuldade do texto, a maioria achou o texto de dificuldade *regular*; para 9, foi *fácil* e para 7, *difícil*.

No item *dificuldades encontradas no texto* observa-se, a partir da Tabela 4.8(c) que o vocabulário foi apontado como o maior problema por 40 dos sujeitos, confirmando a expectativa

deste estudo. Para 3 sujeitos, a gramática e o vocabulário são dificuldades. Para um deles, somente a gramática. Apenas 4 sujeitos afirmaram não terem encontrado nenhum tipo de dificuldade na leitura do texto.

Tabela 4.6 – Perfil dos sujeitos: dados de identificação e *background* linguístico

(a) Idade

Idade	Número de sujeitos
16	1
17	3
18	12
19	14
20	7
21	8
22	2
25	2

(b) Ano de ingresso

Ano de ingresso	Número de sujeitos
87	2
88	10
89	11
90	26

(c) Curso

Curso	Número de sujeitos
Estatística	1
Ciências da Computação	2
Física	1
Engenharia Elétrica	30
Matemática Aplicada	15

(d) Inglês instrumental

Inglês Instrumental	Número de sujeitos
Nível I	17
Nível II	32

(e) Anos estudando inglês

Anos estudando inglês	Número de sujeitos
Mais de 5 anos	39
Menos de 5 anos	10

(f) Curso em escolas de língua

	Já fez	Está fazendo
Sim	29	7
Não	20	42

Tabela 4.7 – Perfil dos sujeitos: leitura

(a) LE		
	Inglês	Outras línguas
Sim	43	3
Não	6	46

(b) Tipo		
	Português	Inglês
Variadas	48	30
Técnicas	1	7
Nenhuma	0	12

Tabela 4.8 – Perfil dos sujeitos: avaliação da leitura

(a) Idéia geral: auto-avaliação/avaliação do pesquisador		
	Auto-avaliação	Avaliação do pesquisador
Bom	29	29
Regular	16	13
Ruim	4	7

(b) Auto-avaliação: nível de dificuldade do texto	
	Número de sujeitos
Fácil	9
Regular	33
Difícil	7

(c) Auto-avaliação: dificuldades encontradas no texto	
	Número de sujeitos
Sem dificuldades	4
Vocabulário	40
Gramática	1
Vocabulário e gramática	3
Outras	1

(d) Auto-avaliação de bom leitor: português e inglês		
	Bom leitor em português	Bom leitor em inglês
Sim	27	7
Regular/ alguns textos	15	17
Não	7	25

4.2.1.2 – Vocabulário

O objetivo dessa análise é verificar, principalmente, como se caracteriza a competência lexical apresentada pelos leitores em LE, medida através de testes de vocabulário.

Essa questão geral se subdivide em questões mais específicas:

1. Existe alguma relação entre familiaridade (T1) e frequência (T2)?
2. Existe alguma relação entre familiaridade (T1) e significado (T4)? E entre familiaridade e os seis subtestes que compõem o conhecimento *rico* de vocabulário (T3, T4, T5, T6, T7 e T8)?
3. Existem relações entre os seis subtestes que compõem o conhecimento *rico* de vocabulário (T3, T4, T5, T6, T7 e T8)?
4. Existe alguma relação entre inferência em contexto (T9) e conhecimento *rico* de vocabulário (T3, T4, T5, T6, T7 e T8)?

Para responder a essas perguntas, os dados coletados a partir dos testes de vocabulário são analisados em três níveis ou seja, considerando-se o grupo de sujeitos como um todo (todos os sujeitos, todas as palavras), cada sujeito (todas as palavras) e finalmente cada palavra.

4.2.1.2.1 – Vocabulário: foco no grupo

O desempenho do grupo de sujeitos como um todo foi o primeiro passo da análise. Para isso, foram construídas tabelas de frequência para cada um dos testes, levando-se em conta os dados de todos os sujeitos em cada teste. Foi aplicada a análise de correspondência (AC) (vide descrição da técnica no Capítulo III) a essas tabelas. Através dessa análise, pode-se observar uma proximidade das categorias:

- *nunca vi a palavra e a palavra existe mas não sei o significado* no teste de familiaridade
- *branco e incorreta* nos testes de significado, flexão/derivação e inferência em contexto
- *1, 2 ou mais de 2 corretas* no teste de associação
- *pelo menos 1 correta* no teste de colocação
- *meio correta e correta* no teste de contexto gramatical
- *meio correta e correta* no teste de classe gramatical

Considerando-se essas proximidades, e que os sujeitos tiveram tempo suficiente para responder as questões em classe, e que, além disso, os testes foram realizados em etapas para que a variável "fadiga", em princípio, fosse eliminada, decidiu-se, para uma simplificação da análise sem

perda de informação, por uma recategorização dos testes (vide Apêndice B, Seção B.14 para testes recategorizados após essa primeira análise).

A proximidade das categorias *nunca vi essa palavra* e *a palavra existe mas não sei seu significado* mostra que uma escala de familiaridade de três níveis ou categorias, ao invés de quatro, pode ser mais apropriada. Essas categorias passam a ser *familiar* (*sei o significado*), *mais ou menos familiar* (*estou em dúvida com relação ao significado*) e *não familiar* (*nunca vi a palavra e a palavra existe mas não sei o significado*).

As Tabelas 4.9 a 4.12, a seguir, apresentam os resultados de cada um dos testes, já recategorizados, considerando-se o grupo como um todo.

Tabela 4.9 – Familiaridade

Familiaridade	Palavras inventadas	Palavras reais	Total
Não familiar	388	436	824
Mais ou menos familiar	58	177	235
Familiar	40	367	407
Total	486	980	1466

- (T1) Teste de familiaridade

Para o teste de familiaridade, o total de freqüências é 1466, ou seja, 49 alunos vs 30 palavras (inventadas e reais) menos quatro respostas em branco. Desse total, há 824 ocorrências na categoria *não familiar*, 407 na *familiar* e 235 na *mais ou menos familiar*. Entretanto, esse resultado se mostra distorcido se forem levadas em conta as palavras criadas. Excluindo-se essas palavras, há 436 ocorrências na categoria *não familiar*, 367 na *familiar* e 177 na *mais ou menos familiar*. Esses resultados mostram que o número de palavras desconhecidas, na percepção dos alunos como um grupo, supera o número de palavras conhecidas.

Tabela 4.10 – Freqüência

Freqüência	Palavras inventadas	Palavras reais	Total
Não freqüente	352	350	682
Mais ou menos freqüente	105	308	413
Freqüente	50	322	372
Total	487	980	1467

- (T2) Teste de Freqüência

O total de freqüências é 1467, uma vez que 3 respostas estão em branco. Desse total, 682 ocorrências na categoria *não freqüente*, 372 na *freqüente* e 413 na *mais ou menos freqüente*. Nota-se que o número de palavras freqüentes, na percepção dos sujeitos, é bastante inferior ao de não freqüentes. Entretanto, esse resultado é ilusório, uma vez que nesse total estão também as 10 palavras

inventadas. Retirando-se essas palavras, há 350 ocorrências na categoria *não freqüente*, 322 na *freqüente* e 308 na *mais ou menos freqüente*, mostrando um resultado mais equilibrado entre o número de palavras desconhecidas e conhecidas. Todas as 487 palavras inventadas teriam que estar na categoria *não freqüente*. Entretanto, das 487, total das palavras inventadas, 155 estão nas categorias *pouco freqüente* (105) e *freqüente* (50), mostrando que houve engano em pelo menos metade, embora o mais sério, sejam as 50 ocorrências na categoria *freqüente*.

- (T3, T4, T5, T6, T7 e T8) Teste de conhecimento *rico* de vocabulário

Esse teste é composto de seis subtestes, com resultados apresentados na Tabela 4.11.

- (a) (T3) Teste de flexão/derivação.

Convém lembrar que nesse teste, assim como nos que o seguem, o total de palavras é 980 (49 sujeitos x 20 palavras reais) e não mais 1470, uma vez que foram eliminadas as palavras inventadas.

O grupo apresentou 553 acertos, contra um número de 427 erros, ou seja, uma porcentagem de erros de 56% de acertos para 44% de erros.

- (b) (T4) Teste de significado/definição

Este teste inclui duas partes. Na parte I, é solicitado apenas um significado ou definição para cada palavra. Na parte II, são solicitados significados adicionais, uma vez que muitas das palavras podem ter mais de um significado. Entretanto, a parte 2 não foi respondida por nenhum sujeito, ficando totalmente em branco, diferindo bastante, como já se previa, dos resultados apresentados pelos falantes nativos. O número de acertos (600) foi maior que o número de erros (380), ou seja, uma porcentagem de 61% de acertos para 39% de erros. Isso quer dizer que 61% das palavras foram "traduzidas" corretamente, uma vez que o tipo de resposta mais comum foi um sinônimo em português.

- (c) (T5) Teste de associação

O número de erros neste teste (702) foi bem maior que o número de acertos (278), ou uma porcentagem de 72% de erros para 28% de acertos.

- (d) (T6) Teste de colocação

Neste teste o número de erros (823) foi, mais uma vez, superior ao de acertos (157), ou uma porcentagem de 84% de erros para apenas 16% de acertos.

- (e) (T7) Teste de contexto gramatical e (f) (T8) Teste de classe gramatical

Os testes de contexto e de classe gramatical apresentam resultados diferentes. Em ambos, o número de acertos superam os de erros. Entretanto, no de contexto, os acertos são 544 contra 621 no de classe, enquanto os erros são 436 contra 359 no de classe. Isso significa uma porcentagem de 56% de acertos contra 44% de erros no teste de contexto e 63% de acertos contra 37% de erros no de classe.

Tabela 4.11 – Teste de conhecimento *rico* de vocabulário

(a) Flexão/derivação (T3)

	Palavras reais
Certo	553
Errado	427
Total	980

(b) Significado/definição (T4)

	Palavras reais
Certo	600
Errado	380
Total	980

(c) Associação (T5)

	Palavras reais
Certo	278
Errado	702
Total	980

(d) Colocação (T6)

	Palavras reais
Certo	157
Errado	823
Total	980

(e) Contexto gramatical (T7)

	Palavras reais
Certo	544
Errado	436
Total	980

(f) Classe gramatical (T8)

	Palavras reais
Certo	621
Errado	359
Total	980

• (T9) Teste de inferência em contexto

Ao considerar os resultados do teste de inferência em contexto (vide Tabelas 4.12 (a) e (b) abaixo) observei a necessidade de se subtrair do total de palavras inferidas, aquelas que já eram conhecidas (acertos no teste T3) para se obter o total de acertos resultante do uso do contexto.

Os acertos desse teste, subtraídas as palavras já conhecidas, é então denominado inferência real em contexto (T10). Esse resultado é um função de T9, ou seja, do teste de inferência em contexto e de T4, ou seja, do teste de significado. No teste de inferência em contexto o total de

palavras inferidas e não inferidas é 604 contra 376 ou 62% de acertos para 38% de erros. Das 604 inferidas, 299 ou 49% já eram conhecidas; portanto, as palavras realmente inferidas somaram 305 ou 51%.

Tabela 4.12 – Inferência em contexto

(a) Inferência em contexto	
	Palavras reais
Inferiu	604
Não inferiu	376
Total	980

(b) Inferência real em contexto (T10)	
	Palavras reais
Já sabia	299
Inferiu	305
Não inferiu	376
Total	980

A partir de todas as tabelas acima, que resumem os dados de Vocabulário, pode-se inferir o nível de dificuldade de cada teste para o grupo como um todo. Deixando-se de lado os resultados dos testes de familiaridade e frequência, que são auto-avaliações, e o de inferência em contexto, que leva em conta as palavras já conhecidas, observa-se um contínuo de dificuldade do mais fácil para o mais difícil, conforme a Tabela 4.13 abaixo:

Tabela 4.13 – Testes de vocabulário em ordem crescente de dificuldade

Ordem	Testes
1	Classe gramatical
2	Significado/definição
3	Flexão/derivação
4	Contexto gramatical
5	Inferência real em contexto
6	Associação
7	Colocação

Comparando-se o nível de dificuldade dos testes de classe gramatical e contexto gramatical, é interessante observar que, embora os dois testes busquem o mesmo tipo de informação, o de classe gramatical, em que se usou metalinguagem, mostrou-se curiosamente mais fácil do que o de contexto, onde contextos possíveis e impossíveis foram apresentados para escolha. O teste de classe gramatical não só pareceu mais fácil do que o de contexto, mas foi o que teve, de todos os testes, um

maior número de acertos. Apesar de apresentarem níveis diferentes de dificuldade, os dois testes apresentam uma grande associação, como demonstra o Quadro F.1.

A análise descrita até agora procurou mostrar o desempenho do grupo em cada um dos testes. Entretanto, através dessa análise, é também possível verificar-se semelhanças ou associações entre os resultados de cada teste, foco das perguntas de pesquisa colocadas no início desta seção, visando assim complementar a caracterização da competência lexical dos sujeitos acima apresentada. Algumas perguntas serão respondidas através da análise que considera os sujeitos enquanto um grupo. Outras serão respondidas ou complementadas através dos dois outros níveis de análise, que focalizam os sujeitos ou as palavras.

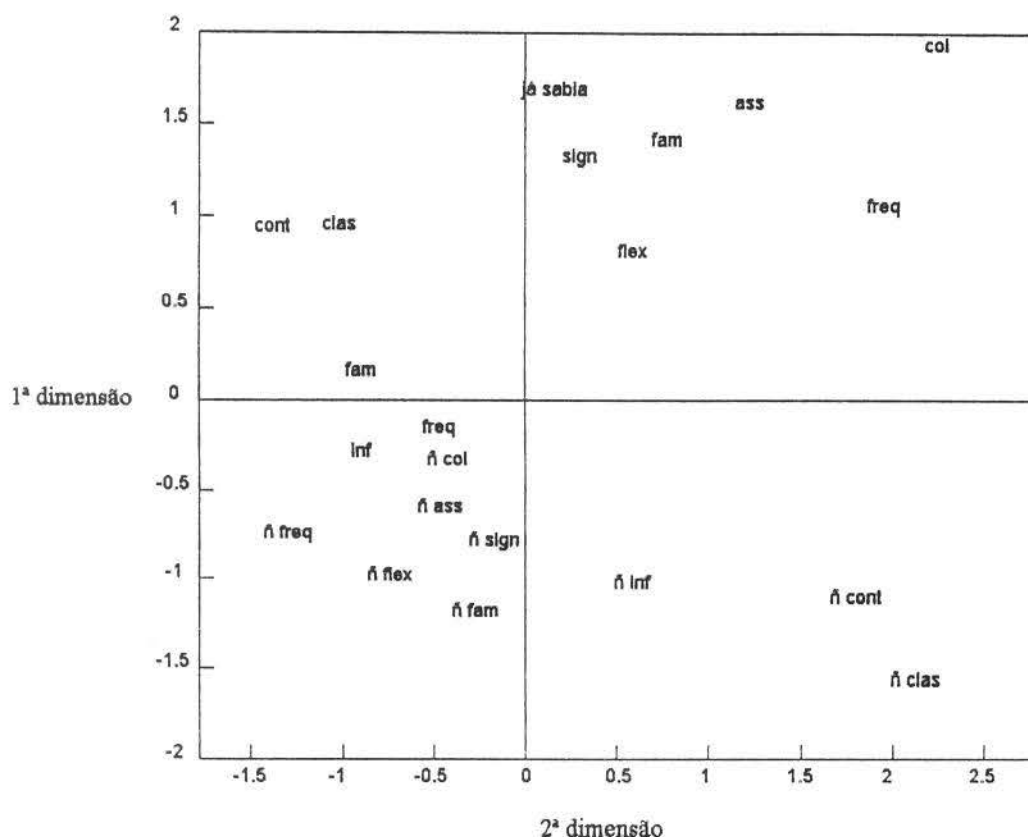


Figura 4.1 – Todos os testes

Aplicando-se a AC ao conjunto das Tabelas acima (4.9-4.12) obteve-se a Figura 4.1. Esse gráfico mostra como os todos os testes se associam entre si. O primeiro eixo desse gráfico explica grande parte da variabilidade e separa claramente as categorias que indicam conhecimento das de não-conhecimento. Pode ser interpretado como um fator de familiaridade-conhecimento. Apesar de frequência e familiaridade estarem associadas, familiaridade é um índice melhor do que frequência, uma vez que separou conhecimento de não conhecimento. A maior contribuição no eixo foi dada por esse fator.

O segundo eixo separa os testes de frequência, significado, associação, flexão/derivação e colocação dos testes de classe, contexto e inferência real em contexto, enquanto que o terceiro eixo

separa basicamente o teste de frequência dos outros testes. Assim, aqueles alunos que souberam colocar, associar, flexionar, que souberam o significado, indicaram maior familiaridade. Aqueles alunos que não souberam classes, contextos, significado, indicaram menor familiaridade.

Para responder à primeira pergunta, ou seja,

1. Existe alguma relação entre familiaridade e frequência?

considero os dados da Tabela 4.14, abaixo:

Tabela 4.14 – Familiaridade vs. frequência

Frequência	Familiaridade			Total
	Familiar	Pouco familiar	Não familiar	
Frequente	240	51	31	322
Pouco frequente	107	82	119	308
Não frequente	20	44	286	350
Total	367	177	436	980

Aplicando-se a AC a esses dados, obtive o Quadro F.2, que mostra uma grande associação entre as duas variáveis. As palavras mais familiares, na percepção dos alunos, ou em sua auto-avaliação, são geralmente também aquelas mais frequentes.

Os resultados mostrados na Figura 4.1, que justapõe todos os testes, corrobora esses resultados. Isso significa dizer que há uma associação positiva entre familiaridade e frequência, ou seja, as palavras frequentes tendem a ser familiares e vice-versa.

Para responder à primeira parte da segunda pergunta,

2. Existe alguma relação entre familiaridade e significado?

considero os dados referentes às variáveis familiaridade e significado da Tabela 4.15 abaixo:

Tabela 4.15 – Familiaridade vs. significado

Significado	Familiaridade			Total
	familiar	pouco familiar	não familiar	
Certo	291	58	31	380
Errado	76	119	405	600
Total	367	177	436	980

Aplicando-se a AC a essa tabela, obtive a Figura F.1, mostrando claramente uma forte associação entre familiaridade e significado, ou seja, os sujeitos têm uma percepção bastante clara

das palavras conhecidas/desconhecidas. Entretanto, convém lembrar que, nessa tabela, são consideradas apenas as palavras reais e não as inventadas, que haviam sido incluídas para uma verificação de consistência e para minimizar as chances de pura adivinhação. Para uma análise das palavras inventadas e reais, retomo os dados da Tabela 4.9.

Como pode-se observar na coluna das palavras inventadas, a grande maioria, ou seja, 388 (de um total de 490, uma vez que são 10 palavras vezes 49 sujeitos) são apontadas pelos sujeitos como não familiares. Isso significa dizer que considerando-se apenas as palavras inventadas, o acerto foi de 80% contra um erro de 20%. Desses 20%, 40 estão na categoria familiar (o que é mais sério) e 58 na pouco familiares, perfazendo um total de 98 palavras (4 respostas foram em branco). Entretanto, considerando-se o número total de palavras, ou seja, 1466 nesta tabela, levando-se em conta as respostas em branco, 98 de erro é aproximadamente 7% e, portanto praticamente irrelevante.

Para responder à segunda parte da segunda pergunta de pesquisa, ou seja, para verificar se existe alguma relação entre familiaridade e os seis subtestes que compõem o conhecimento *rico* de vocabulário (T3, T4, T5, T6, T7 e T8) considero os dados da Tabela 4.11 acima e também a Figura 4.1, gerada através da AC. Esses dados também permitem responder à terceira pergunta de pesquisa, isto é,

3. Existem relações entre os seis aspectos que compõem o conhecimento *rico* de vocabulário (T3, T4, T5, T6, T7 e T8)?

Como já havia salientado acima, o teste de familiaridade foi o que teve uma contribuição maior, separando claramente conhecimento de não conhecimento. Assim, aqueles que reconheceram as palavras como familiares também souberam colocar, associar, flexionar, contextos e classes. Por

outro lado, aqueles que reconheceram palavras como não familiares não souberam colocar, associar, etc. Além da associação desses testes com familiaridade, observa-se uma associação dos testes entre si, ou seja, significado, flexão/derivação, colocação, associação, classe e contexto. Isso significa dizer que o sujeito que sabe o significado também consegue associar, flexionar, por exemplo. Há uma relação entre significado certo e respostas certas nos subtestes e entre significado errado e respostas erradas nos subtestes. Entretanto, convém ressaltar que embora em muitos casos, mesmo sabendo o significado, o sujeito dá respostas incorretas nos outros subtestes, porque há um grau de dificuldade entre eles, como já foi mostrado acima.

A quarta pergunta de pesquisa, isto é,

4. Existe alguma relação entre inferência real em contexto e conhecimento *rico* de vocabulário (T3, T4, T5, T6, T7 e T8) ?

baseia-se na análise dos dados sob o ponto de vista dos sujeitos.

4.2.1.2.2 – Vocabulário: foco nos sujeitos

Aplicando-se a AC à Tabela 4.12 acima, obtive a Figura F.2, em que os 49 sujeitos são representados por letras maiúsculas e minúsculas. As categorias *já sabia*, *não inferiu* e *inferiu* são representadas pelos números 0, 1 e 2, respectivamente. A categoria *inferiu* tem que ser vista relativamente, uma vez que ela é uma função da categoria *já sabia*. Assim, aqueles que mais sabiam menos inferiram. Os sujeitos próximos à categoria *já sabia*, ou acima do eixo horizontal, são os que apresentam um maior conhecimento de vocabulário; os mais próximos a *não inferiu*, e bem embaixo, na figura, os que menos inferiram e também apresentam um menor conhecimento de vocabulário. Os que estão posicionados mais à direita, próximos da categoria *inferiu*, são os que realmente mais inferiram.

Para comparar os dados de inferência com os de conhecimento de vocabulário, classifico os sujeitos com relação a esse conhecimento. Aplicando-se a AC sobre a Tabela 4.11 acima, obtive a Figura F.3, que exhibe os sujeitos, representados por letras, e os testes que compõem um conhecimento *rico* de vocabulário, representados pelos números de 1 a 8. A análise dessa figura permite uma divisão do grupo em três subgrupos, ou seja, os melhores (na parte superior), os médios, ou próximos ao eixo horizontal que divide conhecimento de não conhecimento, e os piores (grupo inferior). Dentro do grupo de maior conhecimento, há sujeitos com níveis variados de conhecimento. Os localizados mais acima são os que mais sabem, e, portanto, os que menos tiveram necessidade de inferir (por exemplo, os sujeitos (T), (f), (l), (p)).

Voltando agora para a Figura F.2, o que chama atenção nesse grupo são os sujeitos (S) e (s). Eles estão em uma posição intermediária como relação ao conhecimento (vide Figura F.3), entretanto, foram os que mais inferiram.

Analizando o grupo de sujeitos com menor conhecimento, os resultados também confirmam o que se esperava, ou seja, que os sujeitos com um conhecimento lexical limitado não são capazes de inferir muitas palavras. Esses sujeitos estão posicionados à esquerda e embaixo da figura, ou seja, as letras (I), (M), (Q) (mesma posição do M), (i), (O), (J), dentre outros. Quanto mais à esquerda estiverem, menos conseguiram inferir. Entretanto, eliminando-se esse subgrupo, há dois sujeitos, representados pelas letras N e e, que, apesar de estarem no grupo de menor conhecimento, conseguem inferir razoavelmente bem (10 e 11 palavras das 20), considerando-se o número de palavras conhecidas, que era muito pequeno, ou seja, 2 e 1, respectivamente. Com exceção desses dois casos, que, no entanto, não parecem muito significativos, levando-se em conta o número de sujeitos (49) o que a análise parece mostrar, de uma maneira geral, é um padrão de inferência ligado

a conhecimento de vocabulário.

Essa análise que focaliza cada sujeito, além de adicionar detalhes importantes, também foi necessária para gerar um escore para cada sujeito em cada um dos testes, a serem usados na análise intracomponentes e na caracterização de um perfil detalhado dos sujeitos que participaram no estudo de ênfase qualitativa. Esse escore foi gerado aplicando-se a AC aos resultados de cada sujeito em cada teste. Foram gerados cinco escores:

- Familiaridade (FAM)
- Frequência (FREQ)
- Conhecimento *rico* de vocabulário (os seis subtestes que compõem o teste conhecimento de vocabulário foram agrupados, gerando um escore único, denominado (TCV)
- Inferência real, deixando de lado as palavras conhecidas, denominado (INFREAL)
- Conhecimento de vocabulário e inferência real (TCVI)

4.2.1.2.3 – O vocabulário: foco nas palavras

Um terceiro e último nível de análise consistiu em focalizar as palavras avaliadas nos diversos testes. Essa análise, além de adicionar detalhes importantes às perguntas já respondidas, é necessária para responder a outras. A análise das palavras mais freqüentes/menos freqüentes é mostrada através da Tabela 4.16.

Aplicando-se a AC sobre essa tabela, obtive a Figura 4.2.

A análise desse gráfico mostra que o grau de freqüência segue um *arco* no espaço. Esse fenômeno, denominado *efeito ferradura* ou *de Guttman*, ocorre quando há uma ordem entre as categorias segundo os dados da tabela, além de mostrar a força dessas categorias. Observa-se, assim, um contínuo de freqüência, que vai do mais freqüente para o menos freqüente. Isso quer dizer que as posições das palavras no espaço indicam o grau de freqüência segundo o grupo de sujeitos que participaram do teste, portanto, uma ordem de freqüência relativa.

Para efeitos de comparação, a freqüência das palavras do grupo de sujeitos não nativos foi comparada com a freqüência baseada nos dados de um grupo de nativos, que haviam sido coletados para servirem de critério para correção, principalmente, do teste de contextos gramaticais (vide Capítulo III). Os contínuos de freqüência dos dois grupos são apresentados na Tabela 4.17.

Como pode-se observar, a freqüência dos dois grupos é diferente. Para os falantes nativos, somente são consideradas não freqüentes a palavra *spawned*, além das palavras inventadas. Para os sujeitos deste estudo, esse grupo inclui outras palavras, tais como *focusing* e *cluster*. Entretanto, a maior diferença está nas palavras *mais ou menos freqüentes* e *freqüentes*. *Clusters*, *loomed*,

fortuitously e *disk-shaped* são as únicas consideradas de frequência média pelos falantes nativos. Todas as outras são consideradas de alta frequência. O conjunto de palavras frequentes dos sujeitos inclui apenas cinco palavras, sendo uma delas uma palavra inventada. Essas diferenças podem ser consideradas como uma amostra das palavras frequentes e não frequentes do léxico como um todo.

Tabela 4.16 – Frequência (por palavra)

Palavra	Frequência			Total
	Não frequente	Pouco frequente	Frequente	
*lodings	41	7	1	49
cluster	34	7	8	49
*skapping	38	9	2	49
spawned	46	2	1	49
shaken	20	18	11	49
*glim	45	3	1	49
loomed	45	3	1	49
*drealled	42	5	2	49
resembling	31	17	1	49
cloud	7	24	18	49
*learn	36	10	2	48
disk-shaped	26	16	7	49
*tark	40	5	3	48
fortuitously	23	17	9	49
fairly	10	24	15	49
*observement	4	24	21	49
massive	23	16	10	49
*trivel	37	10	2	49
picked up	9	26	14	49
focusing	30	11	8	49
spring	7	20	22	49
empty	5	13	31	49
throughout	6	20	23	49
pattern	10	19	20	49
appeared	7	17	25	49
close	1	4	44	49
*adjustion	13	21	14	48
receiving	9	20	20	49
*resks	36	11	2	49
slowly	1	14	34	49
Total	682	413	378	1467

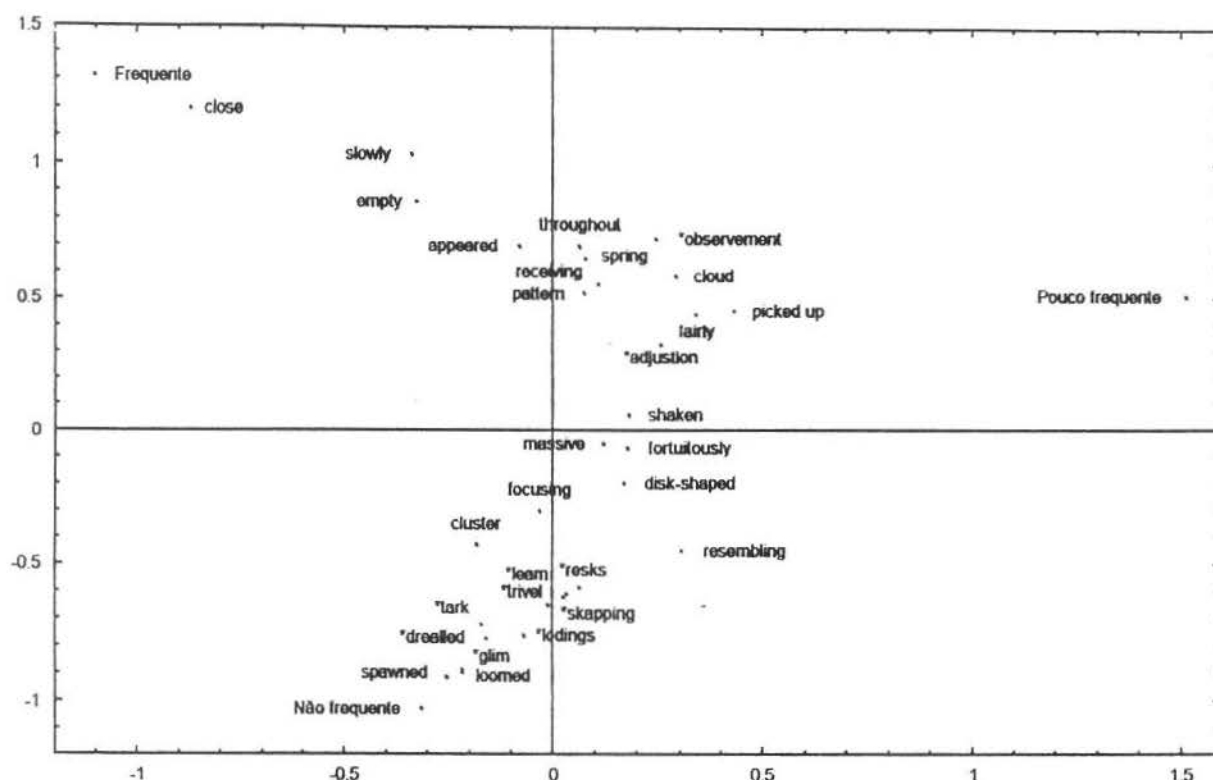


Figura 4.2 – Posição das palavras quanto à frequência

Tabela 4.17 – Palavras avaliadas, classificadas por falantes nativos/não nativos em um contínuo de frequência

Palavras	Falantes	
	Nativos	Não nativos
Frequentes	shaken, resembling, cloud, fairly, massive, picked-up, focusing, spring, empty, throughout, pattern, appeared, close, receiving, slowly	close, slowly, empty, *observement, appeared
Mais ou menos frequentes	clusters, looming, fortuitously, disk-shaped	throughout, spring, cloud, receiving, pattern, disk-shaped, fortuitously, picked-up, fairly, *adjustion, shaken, massive
Não frequente	spawned, *glim, *drealled, *lodings, *tark, *skapping, *trivel, *learn, *resks, *observement, *adjustion	spawned, *glim, looming, *drealled, *lodings, *tark, *skapping, *trivel, *learn, *resks, resembling, cluster, focusing

Considerando o teste de familiaridade, agora por palavra, apresento a Tabela 4.18 a seguir. Aplicando-se a AC a ela, foi gerada a Figura 4.3 abaixo, que classifica as palavras em diferentes graus de familiaridade.

O eixo vertical ou primeiro eixo da Figura 4.3 explica 83% da inércia e o segundo, 17%. Observando-se o gráfico, pode-se verificar que as categorias coluna que têm maior contribuição absoluta sobre o primeiro eixo são: *familiar*, do lado positivo, com uma contribuição de 61% e *não*

familiar do lado negativo com uma contribuição absoluta igual a 38%, de modo que essas duas categorias juntas explicam 99% da inércia do primeiro eixo. Portanto, esse eixo vai no sentido crescente, de um menor a um maior grau de familiaridade e pode ser interpretado como um fator de familiaridade. Observando-se o segundo eixo, verifica-se que a maior contribuição absoluta (83%) é dada pela categoria *mais ou menos familiar* do lado positivo.

Tabela 4.18 – Familiaridade (por palavra)

Palavra	Familiaridade			Total
	Não familiar	Pouco familiar	Familiar	
*lodings	45	3	1	49
cluster	37	7	5	49
*skapping	41	6	2	49
spawned	45	4	0	49
shaken	21	7	21	49
*glim	47	2	0	49
loomed	47	1	1	49
*drealled	48	1	0	49
resembling	40	8	1	49
cloud	18	10	21	49
*learn	44	4	1	49
disk-shaped	27	15	7	49
*tark	47	0	1	48
fortuitously	27	10	12	49
fairly	21	17	11	49
*observement	16	15	16	47
massive	26	7	16	49
*trivel	45	3	0	48
picked up	17	13	19	49
focusing	34	9	6	49
spring	9	5	35	49
empty	6	4	39	49
throughout	15	21	13	49
pattern	17	14	18	49
appeared	14	10	25	49
close	3	2	44	49
*adjustion	15	16	18	49
receiving	6	12	31	49
*resks	40	8	1	49
slowly	6	1	42	49
Total	824	235	407	1466

Conseqüentemente, o lado positivo do eixo pode ser interpretado como familiaridade

média. Esses dois resultados mostram que o grau de familiaridade também mostra o *efeito de Guttman*, da mesma forma que no gráfico da frequência. Isso quer dizer que as posições das palavras no espaço indicam o grau de familiaridade segundo o grupo de alunos que participaram do teste.

A análise o gráfico mostra (em ordem decrescente de familiaridade):

- Palavras mais familiares: *close, slowly, empty, spring, receiving e appeared*.
- Palavras de familiaridade média: *cloud, *adjustion, picked up, pattern, shaken, *observement, throughout, massive, fairly, fortuitously e disk-shaped*
- Palavras menos familiares: **drealled, *glim, *tark, *trivel, loomed, spawned, *lodings, *leam, resembling, *resks, *skapping, cluster e focusing*.

A Figura 4.3 revela a informação da tabela com bastante precisão. As palavras inventadas foram classificadas como não familiares, como esperado, com exceção das palavras **adjustion* e **observement*, que foram classificadas como *mais ou menos familiares* talvez por serem cognatas e se assemelharem com palavras em português.

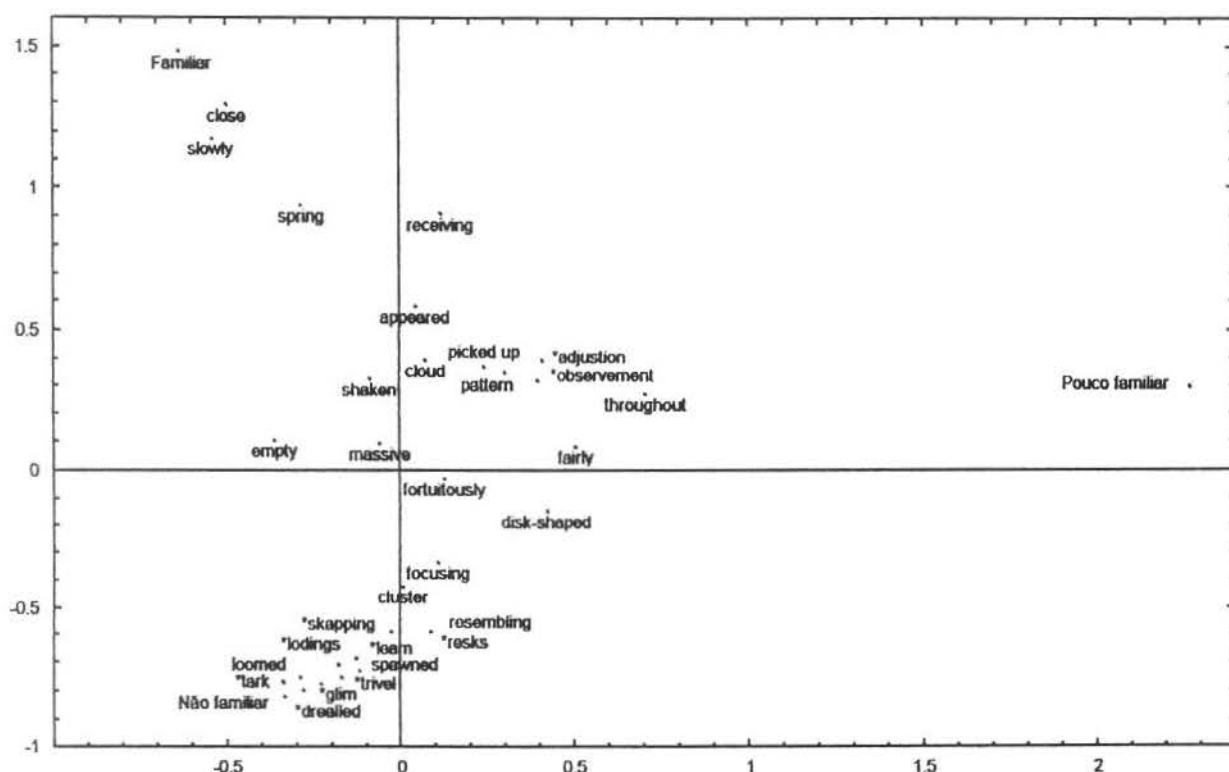


Figura 4.3 – Posição das palavras quanto à familiaridade

A análise conduzida nos resultados dos testes de vocabulário procurou mostrar o desempenho do grupo como um todo nos vários testes, as associações entre eles, além de gerar um escore para cada sujeito em cada um dos testes. Os resultados principais, que serão retomados na discussão dos resultados, são:

- Um contínuo de dificuldade entre os testes, mostrando que o desempenho do grupo foi diferente

em cada um dos testes; nesse sentido, o teste de colocações foi o mais difícil e o de classes gramaticais o mais fácil. No teste de significado, os acertos referem-se apenas à parte I, ou seja, a um significado para cada palavra. A parte II, solicitando significados adicionais, não foi respondida por nenhum sujeito, ficando totalmente em branco. Isso significa que para o grupo como um todo, os significados das palavras se restringem a apenas um.

- Um padrão de inferência ligado a conhecimento de vocabulário.
- Uma associação positiva entre todos os testes entre si. Esses resultados mostram a consistência dos instrumentos de avaliação e principalmente a adequação do conceito de vocabulário operacionalizado através dos testes. Uma associação forte dos testes que compõem o conhecimento *rico* de vocabulário com o teste de familiaridade poderia indicar sua adequação ou de uma escala de verificação enquanto medida de conhecimento de vocabulário.

4.2.1.3 – Conhecimento conceitual ou prévio

Nesta análise, são levados em conta os resultados do teste de conhecimento conceitual. Da mesma forma que nos testes de Vocabulário, ela é realizada em dois níveis. O primeiro considera o grupo como um todo, e, seu objetivo, é uma caracterização do grupo com relação aos conceitos do texto, e sua interação com as outras variáveis. O segundo considera os sujeitos individualmente, e seu objetivo é gerar um escore ou uma medida de conhecimento conceitual para cada sujeito.

O teste de conhecimento conceitual é composto por seis itens ou conceitos a serem definidos pelos sujeitos. A correção foi feita considerando-se três níveis: respostas corretas, meio-corretas e incorretas. A Tabela 4.19 traz os resultados de cada um dos conceitos considerando-se o grupo de sujeitos.

Tabela 4.19 – Conhecimento conceitual (por conceito)

Conceito	Resposta		
	Correta	Meio-correta	Incorreta
Galáxia	45 / 91,8%	3 / 6,1%	1 / 2%
Big Bang	37 / 77,5%	1 / 2%	11 / 22,4%
Pedra de Rosetta	11 / 22,4%	1 / 2%	37 / 75,5%
Embrião galáctico	20 / 40,8%	5 / 10,2%	24 / 49%
Via Láctea	40 / 81,6%	6 / 12,2%	3 / 6,1%
Ano-luz	42 / 85,7%	7 / 14,3%	0 / 0%

É possível, a partir dessa tabela, verificar-se o desempenho do grupo como um todo com relação ao conhecimento conceitual. Esse desempenho leva em conta a dificuldade relativa de cada conceito. Nesse sentido, parece haver um gradiente de dificuldade entre os conceitos. Nos mais

difíceis (*Pedra de Rosetta* e *Embrião Galáctico*) o total de erros superou o total de acertos (22,4% de acerto para 75.5 de erro em *Pedra de Rosetta* e 40.8 de acerto para 49,0% de erro em *Embrião Galáctico*). Nos mais fáceis (*Galáxia* e *Ano-luz*) o total de acertos (91.8% de acertos e 2,0% de erros em *Galáxia* e 85.7% de acertos e 14.3 de erros em *Ano-luz*). Os conceitos, em ordem crescente de dificuldade: *Galáxia*, *Ano-luz*, *Via Láctea*, *Big Bang* e *Embrião Galáctico*.

Com exceção dos erros apresentados nos conceitos *Embrião Galáctico* e *Pedra de Rosetta*, o desempenho do grupo no teste de conhecimento conceitual foi muito bom, apresentando acertos acima de 75%. Também é importante salientar que, em termos do assunto do texto, o conceito *Pedra de Rosetta* (o mais difícil para o grupo) foi o menos relevante, sendo necessário apenas para se entender uma comparação.

Para gerar-se uma medida ou escore de conhecimento conceitual para cada sujeito, aplicou-se a AC a uma tabela com os acertos e erros de cada um. Nessa tabela, optou-se por uma recategorização dos dados, de forma que cada conceito passa a ter apenas duas categorias, ou seja, correto e incorreto, uma vez que havia sido observado, na Tabela 4.19, que a categoria *meio-correta* havia sido escolhida por um número bastante limitado de sujeitos (total de 23, considerando-se todos os conceitos, em um total de 294).

Os resultados principais dessa análise são:

- O grupo apresentou um desempenho muito bom, com acertos acima de 75%.
- Observou-se um gradiente de dificuldade entre os conceitos. Os mais difíceis foram *Pedra de Rosetta* e *Embrião Galáctico* e, os mais fáceis, *Galáxia* e *Ano-luz*.

4.2.1.4 – Metacognição/concepções de leitura

A análise da parte B do questionário III ou metacognitivo teve dois objetivos:

- Caracterizar o grupo como um todo com relação a cada uma das 36 afirmações ou estratégias metacognitivas, subdivididas em quatro grupos ou categorias (confiança, correção ou reparo, eficiência e dificuldade), ou seja, verificar a percepção do grupo como um todo com relação ao uso dessas estratégias, ou sua consciência sobre seus recursos cognitivos, e o que é exigido deles na tarefa de leitura.
- Classificar cada sujeito com relação à sua concepção de leitura, que pode ser *local* ou *global*. Para essa análise, levou-se em conta apenas as categorias eficiência e dificuldade. Essas categorias contém, juntas 26, afirmações que podem ser classificadas em *globais* (que tem como aspectos subjacentes a idéia central, o conhecimento prévio e a organização do texto) ou *locais* (que tem como aspectos subjacentes som/letra, palavra/significado, sintaxe e detalhes de conteúdo).

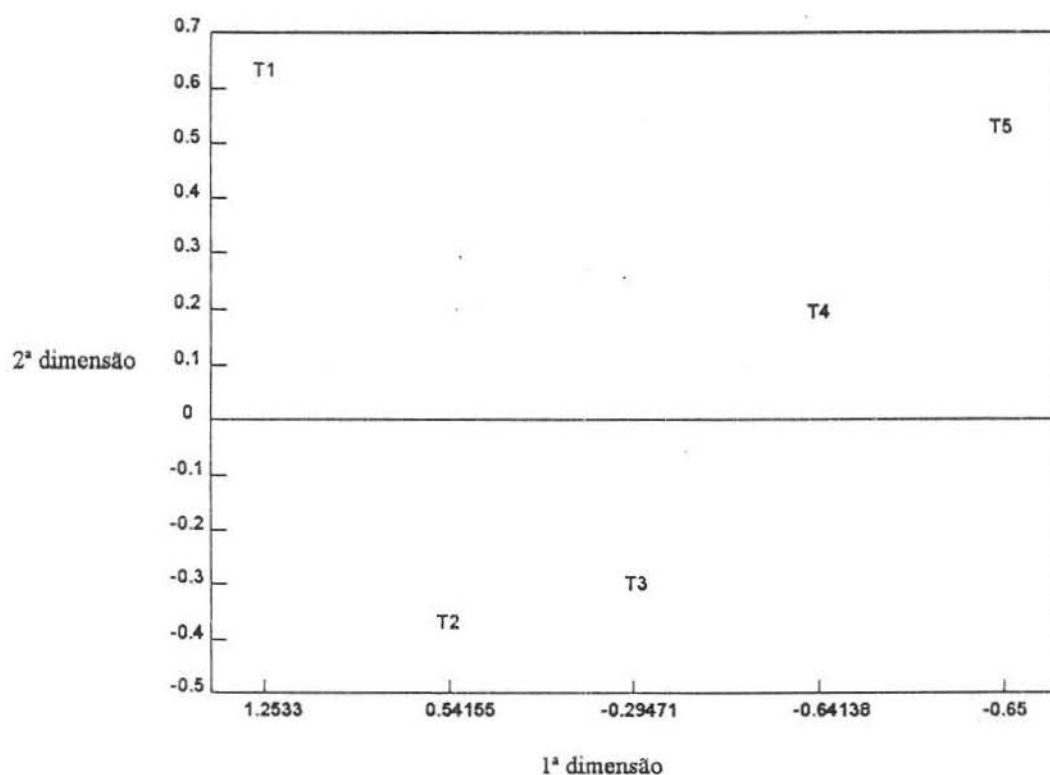


Figura 4.4 – Categorias das estratégias metacognitivas

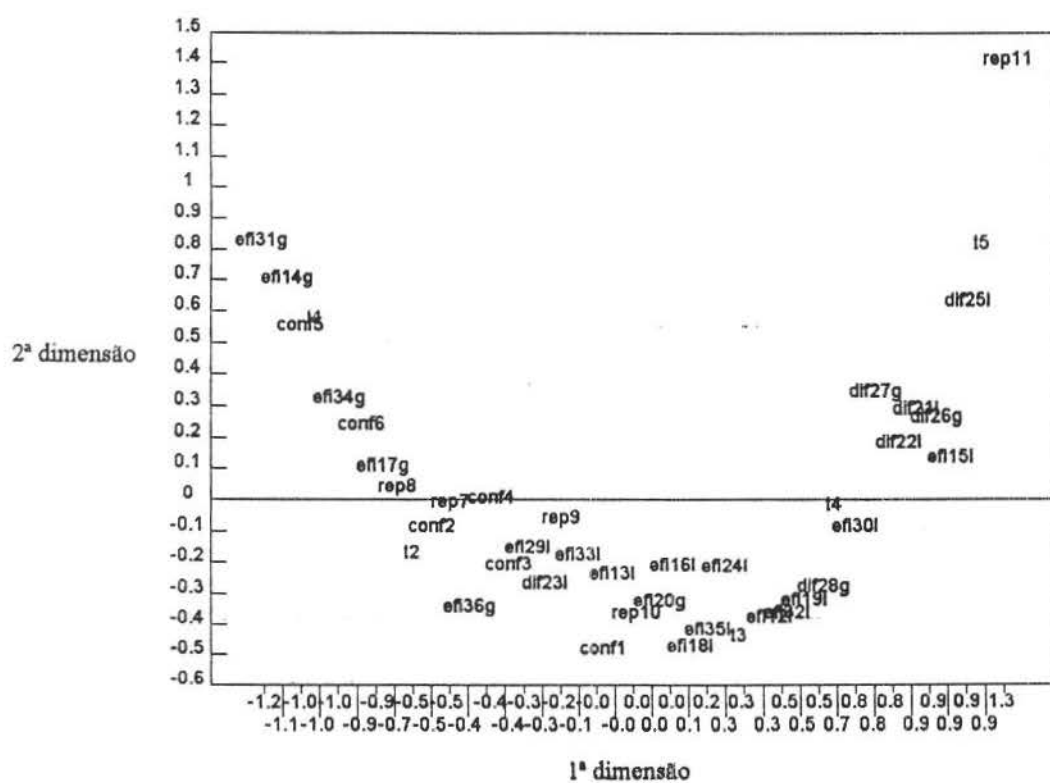


Figura 4.5 – Estratégias metacognitivas

O primeiro passo dessa análise foi a elaboração de uma tabela de frequência com todos os dados de todos os alunos, ou seja, 36 afirmações, subdivididas em quatro grupos, cada uma trazendo uma resposta em uma escala de cinco categorias. Sobre essa tabela aplicou-se a AC, a princípio, sobre as colunas, ou seja, sobre as cinco categorias ou níveis da escala Likert – *concordo plenamente* (1), *concordo* (2), *neutro* (3), *discordo* (4) e *discordo plenamente* (5).

Verifica-se, a partir da figura gerada, ou Figura 4.4, novamente um *efeito ferradura*, mostrando um gradiente ou contínuo entre as categorias. Aplicando-se a AC também às linhas, observou-se o mesmo padrão, ou uma tendência do grupo a associar certos padrões de resposta a cada afirmação, conforme Figura 4.5. Assim, por exemplo, a afirmação 1 do grupo *confiança* ou *conf1* tem associada a ela um padrão de resposta dado pela categoria 3, ou seja, *neutro*. O padrão de respostas é definido pela proximidade, no gráfico, da afirmação à categoria.

Analiso, a seguir, cada um dos quatro grupos de estratégias.

- Estratégias de confiança:

Considerando-se as seis afirmações do grupo *confiança*, relacionadas a vários aspectos da proficiência em leitura em LE percebida pelo leitor, os resultados são mostrados na Tabela 4.20.

Para o grupo como um todo, as afirmações *conf5* e *conf6* são as mais fortes de todas, ou seja, *ser capaz de usar conhecimento prévio e monitorar a compreensão*, uma vez que se encontram mais próximas à categoria 1 ou *concordo plenamente*. São vistas como mais importantes do que as afirmações *conf2* (*ser capaz de reconhecer a diferença entre pontos principais e detalhes*), *conf3* (*ser capaz de relacionar as informações antigas às novas*) e *conf4* (*ser capaz de questionar a veracidade do texto*), que por sua vez são consideradas mais importantes que a afirmação *conf1* (*ser capaz de fazer previsões sobre o que está para vir no texto*). Há, portanto, um contínuo de importância entre esse conjunto de estratégias: $+(conf6\ conf5)\ (conf4\ conf3\ conf2)\ (conf1)-$.

- Estratégias de correção ou reparo:

Com relação as cinco afirmações relacionadas a estratégias de correção, usadas quando há falhas na compreensão, os resultados são mostrados na Tabela 4.21.

Isso significa dizer que, para o grupo como um todo, a estratégia *rep8* (*reler a parte problemática*) é reconhecida como a estratégia de correção ou reparo mais eficiente quando há falhas na compreensão. Também eficiente, embora um pouco menos, é a estratégia *rep7* (*continuar a leitura*). A estratégia *rep9* (*voltar ao ponto imediatamente anterior à parte problemática e reler a partir daí*) também é considerada, embora menos. Não parece haver concordância nem discordância, uma vez que a resposta é na categoria neutro, com relação à estratégia *rep10* (*procurar as palavras*

desconhecidas no dicionário), enquanto a discordância mais forte parece ser com relação à estratégia *repl1* (ao encontrar um problema, desistir e parar de ler).

Tabela 4.20 – Estratégias de confiança

Afirmação/estratégia	Padrão de resposta
conf1	nível 3
conf2	nível 2
conf3	nível 2
conf4	nível 2
conf5	nível 1
conf6	nível 1

Tabela 4.21 – Estratégias de correção ou reparo

Afirmação/estratégia	Padrão de resposta
rep7	nível 2
rep8	nível 1
rep9	entre os níveis 2 e 3
rep10	nível 3
rep11	nível 5

Levando-se em conta todas as estratégias de correção, tem-se também um contínuo, que pode ser representado por $+(rep8) (rep7) (rep9) (rep10) (rep11)-$.

- Estratégias de eficiência e dificuldade

Os dois grupos são considerados conjuntamente uma vez que estão relacionados aos mesmos aspectos. A categoria *eficiência* é constituída por dezessete afirmações relacionadas às estratégias de leitura que, na opinião do leitor, tornam a leitura eficiente. A categoria *dificuldade* é constituída por oito afirmações relacionadas a aspectos que tornam a leitura difícil. Antes de se apresentar os resultados, faz-se necessário uma análise dos vários aspectos subjacentes a cada afirmação, conforme Tabela 4.22.

As estratégias que têm como aspectos subjacentes som/letra, palavra significado, sintaxe e detalhe de conteúdo são consideradas estratégias *locais* enquanto as que têm como aspectos subjacentes a idéia central, o conhecimento prévio e a organização do texto são consideradas estratégias *globais*. Uma concepção de leitura *local* caracteriza-se pela predominância de estratégias *locais*, enquanto uma concepção de leitura *global* pela predominância de estratégias *globais*.

Tabela 4.22 – Aspectos subjacentes às estratégias de eficiência e dificuldade

Aspectos subjacentes	Estratégias	
	Eficiência	Dificuldade
Som/letra	efi12, efi30, efi15	dif21, dif22, dif25
Palavra/significado	efi13, efi29, efi33, efi18, efi32	dif23
Idéia central	efi14, efi31	dif27
Conhecimento prévio	efi17, efi34	dif26
Organização do texto	efi20, efi36	dif28
Sintaxe	efi16	dif24
Detalhes de conteúdo	efi19, efi35	–

Considerando-se em primeiro lugar o grupo *eficiência*, os resultados são exibidos na Tabela 4.23. As estratégias mais importantes para o grupo como um todo são as afirmações *efi14*, *efi31*, *efi20* e *efi34*, que apresentam um padrão de respostas no nível 1, ou *concordo plenamente*. São elas: *entender o significado global do texto, focalizar na organização do texto e integrar a informação do texto com a informação que o leitor já tem*. Uma consulta à Tabela 4.22 mostra que essas estratégias têm como aspectos subjacentes a idéia central e conhecimento prévio.

Tabela 4.23 – Estratégias de eficiência

Afirmação/estratégia	Padrão de resposta
efi12	nível 3
efi13	entre os níveis 2 e 3
efi14	nível 1
efi15	nível 5
efi16	entre os níveis 2 e 3
efi17	entre os níveis 1 e 2
efi18	nível 3
efi19	nível 3
efi20	nível 1
efi29	nível 2
efi30	nível 4
efi31	nível 1
efi32	entre os níveis 3 e 4
efi33	entre os níveis 2 e 3
efi34	nível 1
efi35	nível 3
efi36	nível 2

Ainda importantes, embora um pouco menos, são as estratégias que têm um padrão de respostas entre os níveis 1 e 2 (*efi17*) ou na categoria 2 ou *concordo*, ou seja, um acordo "fraco". São elas as estratégias *efi29* e *efi36* ou *relacionar o texto ao que já sei do tópico, reconhecer o*

significado de palavras e apreender a organização do texto. Essas estratégias têm como aspectos subjacentes o conhecimento prévio, palavra/significado e organização do texto. Com exceção do reconhecimento das palavras, que pode ser considerada uma estratégia *local*, todas as outras estratégias consideradas importantes para o grupo são estratégias mais *globais*, que levam em conta o conhecimento prévio e a organização do texto.

As estratégias menos importantes, ou que têm um padrão de resposta predominantemente nos níveis de discordância, ou seja, 4 e 5, são as *efi30* e *efi15*, ou *pronunciar palavras e ser capaz de pronunciar palavras*, que têm subjacentes os aspectos som/letra.

Todas as outras estratégias apresentam padrões predominantemente no nível 3 ou *neutro* (*efi12*, *efi18*, *efi19* e *efi35*) ou (*falar mentalmente as partes das palavras, procurar palavras nos dicionário, focalizar na organização do texto, focalizar nos detalhes do texto ou de conteúdo*), que tem como aspectos subjacentes som/letra, palavra/significado e detalhes de conteúdo. Há também estratégias num nível intermediário entre 2 e 3 (*efi13*, *efi16*, *efi33*) ou *entender o significado de cada palavra e adivinhar o significado de palavras*, ou entre 3 e 4 (*efi32*) ou *usar um dicionário*. Essas estratégias têm subjacentes os aspectos palavra/significado e sintaxe.

Uma inconsistência do grupo é o fato de considerar a estratégia *apreender a organização do texto* como importante, e logo a seguir, salientar a estratégia *focalizar a organização do texto* como *neutra*. Entretanto, uma explicação para esse padrão de respostas pode estar em uma deficiência do questionário de coleta dos dados. Embora as duas estratégias tenham o mesmo aspecto subjacente, elas poderiam ter sido interpretadas de maneira diferentes, uma vez que uma inclui o termo *apreender* e a outra *focalizar*.

Levando-se em conta todas as estratégias de *eficiência*, tem-se também um contínuo em ordem decrescente, que pode ser representado por +(efi14, efi31, efi20, efi34) (efi17, efi29, efi36) (efi13, efi16, efi33) (efi12, efi18, efi19, efi35) (efi32) (efi15, efi30)-.

A Tabela 4.24 mostra os resultados das estratégias do grupo *dificuldade*. Uma análise dessas afirmações não identifica nenhum problema com um nível de resposta 1, ou seja, *concordo plenamente*. A única estratégia a apresentar um padrão de respostas de nível 2 ou *concordo*, e portanto, a mais problemática, é a *dif23* (*reconhecimento das palavras*), que tem como aspecto subjacente palavra/significado. A *dif24* (*estruturas gramaticais*), que tem como aspecto subjacente a sintaxe, é a única a apresentar um padrão de respostas neutro ou de nível 3. Todas as outras ficam predominantemente entre os níveis 3 e 4 (*dif28*), no nível 4 (*dif26*, *dif27*, *dif21* e *dif22*) ou no 5 (*dif25*), e, portanto, menos problemáticas. Essas estratégias têm como aspectos subjacentes a ideia central, o conhecimento prévio, a organização do texto e som/letra.

Levando-se em conta todas as estratégias de *dificuldade*, tem-se também um contínuo em ordem decrescente, que pode ser representado por +(dif23) (dif24) (dif28) (dif26, dif27, dif21,

dif22) (dif25)-.

Tabela 4.24 – Estratégias de dificuldade

Afirmação/estratégia	Padrão de resposta
dif21	nível 4
dif22	nível 4
dif23	nível 2
dif24	nível 3
dif25	nível 5
dif26	nível 4
dif27	nível 4
dif28	entre os níveis 3 e 4

Após uma análise que considera o grupo como um todo com relação às estratégias metacognitivas, passo a analisar apenas as estratégias que se encontram nas categorias *eficiência* e *dificuldade*, foco propriamente da análise, que tem, como objetivo, inferir as concepções dos sujeito com relação ao seu processo de leitura. Em outras palavras, espera-se poder classificar o sujeito em centrado na palavra ou som (*local*) ou no significado (*global*). Esses resultados, ou o escore de cada sujeito, são usados na caracterização do perfil dos três sujeitos que participaram do estudo de processo da leitura, apresentado no próximo capítulo.

O Quadro 4.1 apresenta as estratégias que caracterizam os leitores em concepção de leitura *local* e *global*. A Figura 4.6, também mostrada abaixo, foi gerada aplicando-se a AC sobre essa tabela.

Essa análise não mostrou ser possível classificar os sujeitos exclusivamente em *locais* ou *globais*, uma vez que todos apresentaram uma combinação de estratégias. Foi, entretanto, possível classificá-los, ou apresentar seus escores, em um contínuo entre *local* e *global*. A Figura 4.6 mostra a posição dos sujeitos com relação ao contínuo *local/global*, que vai desde -2 negativo (extremo *local*) até 7 positivo (extremo *global*). Quanto mais abaixo no gráfico se localizar o sujeito, mais *local* pode ser considerado. Assim, os sujeitos que têm uma concepção de leitura mais *local* são os sujeitos 32, 31, 13, 10, 17, 48; os que têm uma concepção de leitura mais *global* são os sujeitos 45, 29, 44, 7, 11, dentre outros.

Quadro 4.1 – Estratégias/concepções de leitura local e global

Estratégias/Concepções de leitura	
Local	Global
efi12, efi13, efi15, efi16efi18, efi19, efi29, efi30, efi32, efi33, efi35, dif21, dif22, dif23, dif24 e dif25	efi14, efi17, efi20, efi31, efi34, efi36, dif26, dif27 e dif28

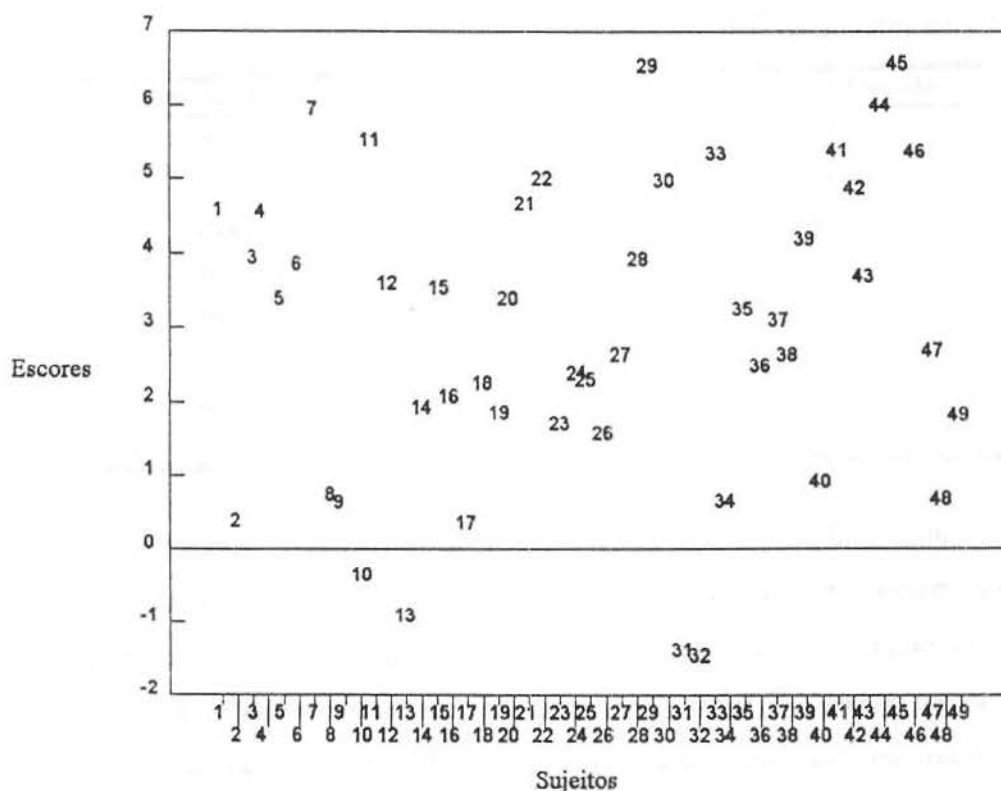


Figura 4.6 – Posição dos sujeitos quanto a concepções de leitura local e global
Contínuo negativo-positivo corresponde a local-global

Os resultados principais dessa análise são:

- Um gradiente entre as afirmações em cada uma das categorias. As mais importantes para o grupo são:
 1. Correção ou reparo: *reler a parte problemática.*
 2. Confiança: *ser capaz de usar conhecimento prévio e monitorar a compreensão.*
 3. Eficiência: *entender o significado global do texto, focalizar na organização do texto e integrar a informação do texto com a informação que o leitor já tem.* Essas estratégias têm como aspectos subjacentes a idéia central e conhecimento prévio.
 4. Dificuldade: *reconhecimento das palavras,* que tem como aspecto subjacente palavra/significado
- No que se refere à concepção de leitura do sujeitos, a análise não mostrou ser possível classificá-los exclusivamente em *locais* ou *globais*, uma vez que todos apresentaram uma combinação de estratégias, mas em um contínuo *local/global*.

4.2.1.5 – Leitura em LM e LE e proficiência em LE

A análise conduzida nos testes de leitura em LM (Cloze1A e Cloze1B) tem como único

objetivo gerar um escore para cada sujeito que seja representativo de sua proficiência em LM. Esse escore foi calculado com base no número de acertos em cada teste, transformados em uma escala de 0 a 10.

A análise conduzida no teste de leitura do tipo cloze em LE (Cloze2C) também tem, da mesma forma que no teste de leitura em LM, o único objetivo de gerar um escore para cada sujeito que represente sua proficiência em LE. Por esse motivo, essa análise não considera, por exemplo, os erros e as lacunas ou itens deixados sem resposta. Embora uma análise desses aspectos pudesse ser útil para se inferir, embora indiretamente, o processo de leitura, considere-se dispensável, uma vez que o processo de leitura já está sendo analisado, e mais diretamente, no estudo apresentado no Capítulo V. Como no teste de LM, os escores foram gerados a partir do número de acertos, transformados em uma escala de 0 a 10.

Além dessa medida de compreensão em leitura, o estudo usa uma outra, elaborada a partir do texto de onde foram extraídas as vinte palavras dos testes de vocabulário. Como já salientei anteriormente, esse teste de leitura consiste em três partes. Na parte I, seu objetivo principal é um controle para as palavras desconhecidas além daquelas 20 previstas e avaliadas pelos testes de vocabulário. Não se poderia deixar de cogitar que o sujeito pudesse não saber outras palavras, além das 20. Um objetivo secundário é, mais uma vez, verificar a consistência do sujeito com relação às palavras inventadas. A expectativa, nesse caso, é de que o sujeito grife as palavras inventadas ou pelo menos parte delas.

Analisando esses resultados, pode-se verificar que grande parte dos sujeitos, ou seja, 35, grifaram apenas palavras que faziam parte da lista de inventadas. Entretanto, 14 sujeitos grifaram, além das inventadas, também outras palavras do texto. São elas: *while*, *birth* (7 ocorrências), *discovered*, *shortly*, *happened* (6), *signals*, *outer*, *Rosetta Stone*, *announced* (5), *gestation*, *announcement*, *calibrate*, *Big Bang stone*, *thought*, *embryo*, *after*, *some*, *ago*, *astrophysicist*, *in order to*, *Milky Way*, *travels*, *contained* (3), *during*, *form*, *earth*, *born* (4), *since* (2).

Como pode-se notar, palavras muito comuns ou de frequência bastante alta fazem parte dessa lista, assim como palavras consideradas cognatas, tais como *announced*, *announcement*, *discovered*, *signals*, *calibrate*, *embryo*, *form*.

Esses resultados são importantes por mostrar que o critério usado para selecionar as palavras problemáticas no teste de vocabulário superestimou o vocabulário dos sujeitos, ou seja, deixou de considerar palavras desconhecidas. Uma análise das palavras escolhidas pelo grupo que serviu de critério para a seleção das palavras (vide Apêndice B, Seções B.2, B.3 e B.4, para critério de seleção usado) mostra que um número maior de palavras, na realidade, tinham sido identificadas como problemáticas. Entretanto, como a inclusão de todas no teste de vocabulário o tornaria muito longo, foi necessário a adoção de um ponto de corte, selecionando assim palavras que tinham sido escolhidas por pelo menos 10% dos leitores.

Na parte II do teste de Leitura I ou Resumo1, os escores mínimos do grupo foram 0. No Resumo1, o escore máximo foi 9.0 e no Resumo2 foi 9.5. Comparando o desempenho dos sujeitos nas versões 1 e 2 do resumo, observa-se em geral um desempenho melhor dos sujeitos na versão 2, ou seja, que contém as palavras reais. Entretanto, a diferença entre os desempenhos é menor do que se esperava. Em outras palavras, escores baixos no Resumo1 corresponderam, em geral, a escores baixos no Resumo2, e vice-versa.

A partir da parte III do teste de leitura têm-se duas medidas. A primeira, Leit1 e Leit2 consiste nas questões propriamente ditas. Os escores mínimos de Leit1 e Leit2 são iguais, ou seja, .5. O escore máximo de Leit1 foi 8.0, enquanto de Leit2 foi 9.5. Também nessa parte do teste observa-se o mesmo padrão de desempenho da parte II. A segunda medida considera em separado a nota dada a uma das questões, mais especificamente a questão de detalhes (Deta1 e Deta2). Os escores mínimos de Deta1 e Deta2 são iguais, ou 0,0. O escore máximo de Deta1 é 9,0 e de Deta2 é 10,0. Nessa parte do teste observa-se um desempenho melhor na versão 2, ou uma contribuição maior da versão com palavras reais.

A análise desse teste de leitura em suas três partes mostrou que os sujeitos que tiveram escores consistentemente baixos nas três medidas, ou seja, resumo, leitura e questão de detalhe foram aqueles que grifaram, na parte I, palavras que não constavam da lista de palavras inventadas. Como mostrei acima, o número de palavras grifadas é bastante significativo. Esses resultados, entretanto, são ainda parciais, uma vez que outros aspectos ainda têm que ser considerados. Assim retomo esta questão mais adiante, depois de analisadas as correlações de todas essas medidas.

A análise das duas medidas de proficiência em LE ou os dois testes do tipo cloze (Cloze2A e Cloze2C) teve como objetivo único gerar escores que representassem a proficiência global em LE dos sujeitos. Convém salientar que o uso do procedimento cloze, nesse caso, fornece uma medida de proficiência aproximada. O escore consistiu no número de acertos transformado numa escala de 0 a 10, como nos testes de leitura em LM e LE acima.

4.2.2 – A análise intercomponentes

Depois de concluída a análise de cada um dos componentes separadamente, onde foram caracterizados o desempenho do grupo como um todo em cada um deles, assim como geradas medidas ou escores para cada sujeito, procedeu-se a um cruzamento desses componentes, para se estabelecer suas associações/relações, ou verificar como se dá a interação entre eles.

A Tabela F.1 apresenta todas as variáveis levadas em conta nessa análise, os escores mínimos/máximos, a média, o desvio padrão assim como também um intervalo de confiança para cada uma das variáveis.

4.2.2.1 – Análise de correlação

Para investigar as associações entre as variáveis foi usada a análise de correlação. Essa análise é feita sobre uma matriz, que leva em conta os escores de cada uma das vinte e uma variáveis acima, gerados a partir da análise intracomponentes (vide descrição da técnica no Capítulo III).

A análise da matriz de correlação foi feita em duas etapas. Em primeiro lugar procurou-se observar as correlações intracomponentes, complementando a análise da seção anterior. Foram consideradas somente aqueles componentes compostos de mais de uma variável, ou seja, vocabulário, leitura em LM e leitura em LE. Devo salientar, além disso, que apenas são apresentadas as correlações resultantes de cruzamentos pertinentes. Não são apresentados, por exemplo, correlações entre frequência (vocabulário em LE) e redação em português. Assim, apresento as correlações referentes ao vocabulário na Tabela 4.25 abaixo..

Esses resultados apenas vêm confirmar os resultados da AC intracomponentes. As correlações mais fortes são apresentadas entre as variáveis familiaridade e inferência real (0,89), familiaridade e conhecimento de vocabulário (0,85), familiaridade e conhecimento de vocabulário e inferência (0,85) e conhecimento de vocabulário e inferência (0,89). Correlações fortes são evidências de que escores altos de uma variável correspondem a escores altos de outra e vice-versa. O que esses resultados parecem sugerir é a existência de um grupo de variáveis, que se apresentam associadas entre si ou que apresentam uma tendência de associação. Entretanto, é necessário salientar que em algumas dessas correlações bastante altas, como por exemplo, entre conhecimento de vocabulário e conhecimento de vocabulário/inferência, a segunda variável já inclui a primeira. Correlações mais baixas entre testes ou variáveis que são distintos, portanto, podem ser consideradas mais importantes.

Tabela 4.25 – Correlações entre variáveis do vocabulário

Variáveis do vocabulário	Correlação
FAM × FREQ	0,60
FAM × TCV	0,85
FAM × INFREAL	0,89
FAM × TCVI	0,85
TCV × INFREAL	0,89
TCV × TCVI	0,85
Legenda: FAM: familiaridade FREQ: frequência TCV: conhecimento de vocabulário INFREAL: inferência real TCVI: conhecimento de vocabulário mais inferência	

O segundo componente analisado foi a leitura em LM. Reproduzo, na Tabela 4.26, essas

variáveis e as correlações resultantes de cruzamentos relevantes.

Tabela 4.26 – Correlações entre variáveis de LM

Variáveis de leitura em LM	Correlação
CLOZE1A × CLOZE1B	0,78
FASEV1 × LETTV2	0,21
FASEV1 × CLOZE1A	0,28
FASEV1 × CLOZE1B	0,46
LEITV2 × CLOZE1A	0,32
LEITV2 × CLOZE1B	0,56
Legenda: CLOZE1A: cloze/palavra exata CLOZE1B: cloze/palavra aceitável FASEV1: redação/outros do vestibular LEITV2: leitura/outros do vestibular	

A correlação mais alta foi aquela resultante do cruzamento dos testes Cloze1A e Cloze1B entre si, ou seja, do mesmo teste corrigido por critérios diferentes (0,78). Essa correlação era esperada, pois ambos estão medindo leitura. O primeiro critério, de palavra exata ou Cloze1A, também é reconhecido, na literatura, como uma medida de proficiência global na língua, extrapolando, portanto, a leitura. Uma correlação relevante (0,56), embora não alta, também pode ser observada entre as variáveis Cloze1B, ou seja, leitura em LM e leitura/outros do vestibular. Os escores do vestibular têm que ser vistos com cuidado, uma vez que são apenas medidas complementares, como já salientei. Entretanto, através de correlações positivas, dão validade às medidas usadas neste estudo. Com exceção dessas variáveis, todas as outras apresentam correlações baixas não sugerindo, portanto, como no componente anterior, a existência de um grupo ou uma tendência de associação.

O terceiro componente analisado é leitura em LE. Esses resultados são apresentados na Tabela 4.27. Nessa análise foram considerados o teste Cloze2C, os testes de leitura, o resumo e as perguntas de detalhes, todos em duas versões, ou seja, I (com palavras inventadas) e II (com as palavras reais). Como uma medida suplementar, foi usado o escore de leitura em inglês do vestibular.

O primeiro bloco de variáveis de leitura em LE compreende os testes de leitura com palavras inventadas (Leit1) e com palavras reais (Leit2). A correlação bastante forte (0,95) entre as duas versões não era esperada, uma vez que a primeira versão inclui as palavras inventadas e a segunda as palavras reais. As correlações entre resumo1 e resumo2 também são altas (0,90). Entretanto, as correlações entre Deta1 e Deta2, apesar de bastante relevantes, são as mais baixas desses cruzamentos (0,71). Esses resultados parecem mostrar uma variação mais forte no desempenho dos sujeitos entre as duas versões da pergunta de detalhes, mas uma variação menos forte no desempenho dos sujeitos entre as duas versões do resumo ou dos testes como um todo. A

partir desses resultados poder-se-ia inferir que mais conhecimento de vocabulário é necessário para responder à pergunta de detalhe do que para a elaboração do resumo. As correlações entre conhecimento *rico* de vocabulário (TCV) e Resumo1, Resumo2, Deta1 e Deta2 apresentadas na Tabela acima confirmam esse aspecto, uma vez que as maiores correlações desse grupo são observadas entre conhecimento de vocabulário e Deta2. As correlações entre Leit1, Deta1 e entre Leit2, Deta2 são relevantes (0,69 e 0,67, respectivamente). Esses resultados podem ser explicados, uma vez que Deta1 é parte de Leit1, da mesma forma que Deta2 faz parte de Leit2. Entretanto, as correlações de Leit1 e Leit2 com Resumo1 e Resumo2 são ainda mais altas (0,74 e 0,75). Esses resultados parecem indicar uma associação maior entre todas as perguntas que compõem a Leit1 e o resumo do que todas as perguntas e a pergunta de detalhe. O mesmo padrão se observa com relação ao Leit2, Resumo2 e Deta2.

Tabela 4.27 – Correlações entre variáveis de leitura em LE

Variáveis de leitura em LE	Correlação
LEIT1 × LEIT2	0,95
DETA1 × DETA2	0,71
RESUMO1 × RESUMO2	0,90
LEIT1 × RESUMO1	0,74
LEIT1 × RESUMO2	0,70
LEIT1 × DETA1	0,69
LEIT1 × DETA2	0,57
LEIT2 × RESUMO1	0,70
LEIT2 × RESUMO2	0,75
LEIT2 × DETA1	0,62
LEIT2 × DETA2	0,67
DETA1 × RESUMO1	0,61
DETA1 × RESUMO2	0,56
DETA2 × RESUMO1	0,56
DETA2 × RESUMO2	0,66
CLOZE2C × LEIT1	0,61
CLOZE2C × LEIT2	
CLOZE2C × LEITV3	0,68
LEITV3 × LEIT1	0,45
LEITV3 × LEIT2	0,48
Legenda: CLOZE2C: cloze/palavra aceitável LEIT1: teste de leitura I (palavras inventadas) LEIT2: teste de leitura II (palavras reais) RESUMO1: resumo I (palavras inventadas) RESUMO2: resumo II (palavras reais) DETA1: perguntas de detalhes (palavras inventadas) DETA2: perguntas de detalhes (palavras reais) LEITV3: leitura em inglês do vestibular	

Com exceção das correlações muito altas entre Leit1 e Leit2 e Resumo1 e Resumo2, todas as outras correlações são mais ou menos semelhantes, sugerindo mais uma vez a existência de um grupo de variáveis.

O segundo bloco de variáveis da Tabela 4.27 compreende as variáveis do tipo cloze cruzadas com os testes de leitura. A maior correlação (0,65) deste bloco é apresentada pelo cruzamento de Cloze2C e Leit2. Uma correlação relevante entre essas duas medidas de leitura é importante, podendo ser usada para validar o teste Leit2 enquanto uma medida de leitura.

O terceiro conjunto de variáveis inclui a variável do tipo cloze cruzada com o escore de leitura em inglês do vestibular. A correlação apresentada também é relevante (0,68), mostrando uma associação entre as duas medidas.

O quarto bloco de variáveis compreende as variáveis do teste de leitura I e 2 cruzadas com o escore de leitura do vestibular. As correlações, apesar de positivas e indicarem tendências de associação, são as mais fracas de todos os cruzamentos dessa tabela.

Os resultados apresentados nos três últimos blocos de variáveis são interessantes, uma vez que parecem sugerir uma associação das medidas Leit2 e Cloze2C, das medidas Cloze2C e LeitV3 das medidas Leit2 e LeitV3. Apesar de difíceis de serem interpretados, pelo menos neste estágio da análise, o que esses resultados parecem mostrar é que todas as três medidas apresentam correlações com pelo menos uma outra, podendo ser consideradas, portanto, medidas válidas.

O quarto e último componente analisado é a proficiência em LE. O resultado é apresentado na Tabela 4.28 a seguir:

Tabela 4.28 – Correlações entre duas medidas de proficiência em LE

Variáveis de proficiência em LE	Correlação
CLOZE2A × CLOZE2B	0,92
Legenda: PROFICIÊNCIA EM LE 1. CLOZE2A: cloze/palavra exata 2. CLOZE2B: cloze/palavra falante nativo	

As duas variáveis desse componente foram cruzadas entre si, apresentando uma correlação bastante forte, ou seja, de 0,92. Mesmo usando critérios diferentes de correção, observa-se uma correlação bastante forte entre as duas medidas. Mais uma vez essas correlações altas devem ser consideradas com cautela, podendo ser enganosas, uma vez que as duas variáveis foram obtidas a partir do mesmo teste, embora corrigido com critérios diferentes, e que ambos são medidas de proficiência em LE.

Depois de concluída a análise das correlações intracomponentes, procedeu-se à análise das correlações intercomponentes propriamente dita.

A Tabela 4.29 abaixo exibe as correlações entre os componentes principais deste estudo,

que compreende as correlações entre conhecimento de vocabulário/inferência e proficiência em LE, assim como entre essas variáveis e os vários grupos de testes de leitura em LE:

Tabela 4.29 – Correlações entre variáveis de conhecimento de vocabulário/inferência e leitura em LE

Variáveis de conhecimento de vocabulário/inferência e leitura em LE	Correlação
TCV × CLOZE2A	0,75
TCV × CLOZE2B	0,78
TCV × CLOZE2A	0,81
INFREAL × CLOZE2A	0,87
TCVI × CLOZE2A	0,77
TCVI × CLOZE2AB	0,80
TCV × LEIT1	0,45
TCV × LEIT2	0,46
TCV × DETA1	0,32
TCV × DETA2	0,50
TCV × RESUMO1	0,35
TCV × RESUMO2	0,38
INFREAL × LEIT1	0,55
INFREAL × LEIT2	0,56
TCVI × LEIT1	0,48
TCVI × LEIT2	0,47
CLOZE2A × LEIT1	0,54
CLOZE2A × LEIT2	0,55
CLOZE2B × LEIT1	0,60
CLOZE2B × LEIT2	0,61
TCV × CLOZE2C	0,76
INFREAL × CLOZE2C	0,83
TCVI × CLOZE2C	0,77
CLOZE2A × CLOZE2C	0,57
CLOZE2B × CLOZE2C	0,85
TCV × LEITV3	0,64
INFREAL × LEITV3	0,71
TCVI × LEITV3	0,65
CLOZE2A × LEITV3	0,68
CLOZE2B × LEITV3	0,74

O primeiro bloco cruza as variáveis de conhecimento de vocabulário/inferência com as variáveis de proficiência em LE. Como era de se esperar, as correlações são altas, uma vez que o conhecimento de vocabulário é medido por um teste abrangente. Dentre essas correlações altas, a mais relevante é aquela encontrada entre INFREAL e Cloze2B, ou seja, entre inferência e a versão corrigida segundo o critério dos falantes nativos.

Observa-se, no segundo bloco, que as correlações entre conhecimento de vocabulário (TCV) e conhecimento de vocabulário mais inferência (TCVIR), e Leit1 e Leit2, apesar de não serem altas, são positivas, podendo indicar tendências de associação (0,45 e 0,46). As correlações com a inferência, entretanto, são mais altas 0,55 e 0,56. Correlações mais altas com a inferência não são apenas identificadas no cruzamento com Leit1 e Leit2, mas também podem ser encontradas em outros cruzamentos nessa tabela. Esses resultados parecem sugerir que, embora inferência e conhecimento de vocabulário sejam muito associados, o escore inferência é um escore mais completo do que apenas o conhecimento de vocabulário, envolvendo outras habilidades.

Mesmo considerando a correlação de 0,50 entre inferência e Leit1 e Leit2, ela não é alta. A expectativa era de uma correlação maior com conhecimento de vocabulário ou inferência, uma vez que as medidas Leit1 e Leit2 foram elaborados a partir do texto de onde foram extraídas as palavras para o teste de vocabulário. É necessário estar ciente de que as correlações muito altas foram aquelas obtidas a partir do cruzamento de variáveis obtidas a partir do mesmo teste ou que incluem a outra, como no caso da inferência e que aspecto seja levado em conta na interpretação de correlações mais baixas entre variáveis distintas, como é o caso do conhecimento de vocabulário e leitura.

Uma análise mais detalhada dos escores de vocabulário e de leitura, entretanto, mostrou que os sujeitos que obtiveram escores fracos de conhecimento de vocabulário foram também aqueles que grifaram como desconhecidas outras palavras no texto. Por outro lado, os sujeitos que apresentaram escores altos de conhecimento de vocabulário foram aqueles que, em geral, não grifaram palavras além daquelas usadas para os testes de vocabulário. Além disso, aqueles que mostraram escores altos de leitura em Leit1, ou seja, com as palavras inventadas, também obtiveram escores altos em Leit2, com as palavras reais. Por outro lado, aqueles que apresentaram escores baixos na versão com palavras inventadas, também tiveram, em geral, escores baixos na versão com palavras reais, mostrando que as palavras do teste não foram decisivas no desempenho, mas sim as outras palavras do texto. Foram poucos os casos, ou apenas seis, que podem ser considerados uma exceção a essa regra. Esses casos serão apresentados e discutidos mais adiante na seção.

Por sua vez, a análise do teste de leitura mostra que aqueles que conseguiram escores acima de cinco no teste de leitura foram também aqueles que faziam parte do grupo que somente apresentou problemas com as palavras do teste, enquanto aqueles que obtiveram escores abaixo de cinco foram os que desconheciam palavras além daquelas avaliadas pelos testes. Convém lembrar que apenas as palavras do teste foram levadas em conta para gerar-se o escore de conhecimento de vocabulário. Se tivessem sido levadas em conta todas as palavras, os escores baixos de conhecimento de vocabulário seriam ainda mais baixos, as correlações entre TCV e Leit2 seriam mais altas e as diferenças entre Leit1 e Leit2 não seriam tão altas.

Além disso, a análise do segundo bloco de variáveis também mostra que as relações do

TCV com Resumo1 e Resumo2 foram fracas, as mais fracas de todo o grupo (0,35 e 0,38). Entretanto, as correlações de TCV e Deta2 (0,50) são maiores. Esses resultados parecem confirmar os resultados apresentados na Tabela 4.30 acima, ou seja, de que parece haver a necessidade de um conhecimento maior de conhecimento de vocabulário para responder a pergunta focalizando detalhe do que para a elaboração do resumo.

Ao analisar-se o terceiro bloco de variáveis da Tabela 4.29, observam-se correlações bem mais relevantes, sendo as mais altas (0,60 e 0,61) aquelas apresentadas pelo cruzamento de Cloze2B e Cloze2C, ou seja, entre proficiência em LE e leitura do tipo cloze. Apesar de relevantes, essas correlações não são muito altas, mostrando que embora se tenha usado o mesmo teste, a correção com critérios diferentes mostrou ser eficiente para caracterizar um teste de leitura e um teste de proficiência em LE.

No quarto grupo, com os resultados do vestibular, embora correlações muito boas sejam observadas com relação a todas as medidas de conhecimento de vocabulário, a maior correlação é apresentada pelo cruzamento de Cloze2B ou proficiência em LE e a leitura em inglês do vestibular (0,74).

A Tabela 4.30 apresenta as correlações entre a leitura em LE e outros componentes, ou seja, conhecimento conceitual ou prévio, concepções de leitura e leitura em LM.

Tabela 4.30 – Correlações entre variáveis de leitura em LE, conhecimento conceitual, concepções de leitura em LE e leitura em LM

Variáveis de leitura em LE, conhecimento conceitual, concepções de leitura em LE e leitura em LM	Correlação
TCC × LEIT1	0,53
TCC × LEIT2	0,49
TCC × RESUMO1	0,28
TCC × RESUMO2	0,24
TCC × DETA1	0,17
TCC × DETA2	0,08
META × LEIT1	0,11
META × LEIT2	0,12
META × CLOZE2C	0,10
META × LEITV3	0,11
META × CLOZE1A	0,09
META × CLOZE1B	0,15
CLOZE1A × CLOZE2C	0,42
CLOZE1B × CLOZE2C	0,58
CLOZE1A × LEIT1	0,28
CLOZE1A × LEIT2	0,35
CLOZE1B × LEIT1	0,36
CLOZE1B × LEIT2	0,42

O primeiro bloco de variáveis exibe as correlações entre o teste de conhecimento conceitual e os escores das medidas Leit1 e Leit2. Convém lembrar que, com relação ao conhecimento conceitual, somente essas correlações são pertinentes, um vez que esse teste foi elaborado considerando-se os conceitos do texto do teste de leitura I e II, e, portanto, só faz sentido com relação a esse teste. As únicas correlações que mostram tendências de associação são aquelas apresentadas pelos Leit1 (0,52) e Leit2 (0,48), mostrando que em Leit1, ou seja, com palavras inventadas, a contribuição do conhecimento conceitual parece ter sido maior, o que era previsível.

O segundo conjunto mostra as correlações entre os resultados do questionário metacognitivo, mais especificamente entre escores que refletem uma concepção de leitura mais *local* ou mais *global* e escores de leitura em LE e LM. Como se pode observar, as correlações entre essas variáveis são bastante fracas. As hipóteses para explicar esses resultados são várias. Em primeiro lugar, eles poderiam significar que tanto uma concepção *local* como *global* pode estar associada ao sucesso na compreensão. Uma segunda explicação é que os escores que mostram a concepção de leitura na realidade não classificam o leitor como apresentando exclusivamente uma concepção ou outra, ou seja, totalmente *global* ou *local*, mas mostram um contínuo, isto é, os leitores na realidade são ora *globais* ora *locais*, combinando os dois tipos de estratégias. Ainda uma terceira explicação poderia estar em uma limitação do instrumento de coleta, ou seja, no questionário metacognitivo, ou ainda na amostragem para a coleta.

O terceiro grupo de variáveis da Tabela 4.30 apresenta todas as correlações possíveis entre os vários testes de leitura em LM e os de leitura em LE. Correlações positivas podem ser observadas entre a leitura em LM, medida pelo teste do tipo Cloze1B e os testes de leitura do tipo cloze (0,58). Curiosamente, a medida Leit2 apresenta as correlações mais baixas (0,42), principalmente quando comparadas com as correlações de 0,80 e 0,90 apresentadas por algumas das variáveis. Dos três grupos de testes de leitura em LE, as correlações mais baixas com os testes de proficiência em LM são apresentadas por Leit1 e Leit2 (0,41).

A análise conduzida na matriz de correlações parece indicar a existência de alguns agrupamentos de certas variáveis vs. o não agrupamento de outras, o que poder ser mais bem investigado através da análise fatorial, apresentada na Seção 4.2.2.2. Antes, entretanto, de focalizar essas análise, considero a seguir, mais detalhadamente, a relação entre conhecimento de vocabulário e leitura.

Para isso, foi elaborado um gráfico que cruza os escores do teste de conhecimento *rico* de vocabulário ou TCV (escore dos subtestes T3, T4, T5, T6, T7 e T8) e os escores de Leit2 ou seja, o teste com as palavras reais. O objetivo dessa análise é uma investigação mais detalhada dos escores de leitura e de conhecimento de vocabulário, uma vez que os resultados da análise das correlações mostrou não serem muito altas as correlações entre essas duas variáveis.

Esses resultados são apresentados na Figura 4.7 a seguir.

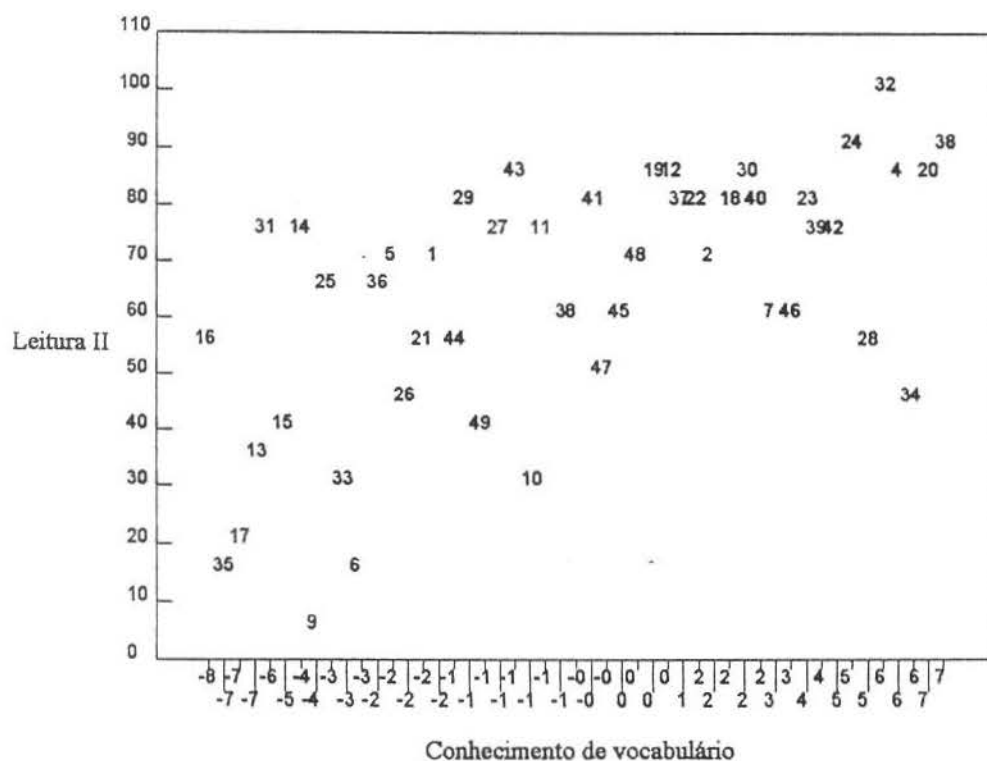


Figura 4.7 – Escores de leitura vs. escores de conhecimento de vocabulário

Esse gráfico exibe no seu eixo vertical os escores dos sujeitos no teste de leitura II (de 5 a 100) ou Leit2, e, no horizontal, seus escores de conhecimento de vocabulário (de -4 a 7).

A análise dessa figura parece mostrar uma correspondência entre escores de conhecimento de vocabulário e escores de leitura. Escores baixos em leitura parecem corresponder a escores baixos em conhecimento de vocabulário e vice-versa. Pode-se observar, por exemplo, que os sujeitos que apresentaram escores mais altos em conhecimento de vocabulário, ou aqueles que figuram em posições superiores nos gráficos de vocabulário, isto é, os sujeitos 32, 38, 42, 20, 4, também são os que se posicionam na parte superior direita da figura em questão. Os sujeitos mais fracos em vocabulário, ou seja, 9, 6, 35, 17, 10, também aparecem na porção inferior esquerda. Entretanto, alguns sujeitos parecem se desviar desse padrão. É o que se observa principalmente com os sujeitos 31, 14, 25, 16 e também 28 e 34. Embora apresentando escores de conhecimento de vocabulário baixos, observa-se, para os quatro primeiros, um desempenho em leitura que lhes garantiu escores entre 55 e 75. Para os sujeitos 28 e 34, a situação é inversa. Embora seus escores de vocabulário estejam entre os mais altos, principalmente o 34, verifica-se um escore de leitura de aproximadamente 40, o que é baixo. Em princípio, esses sujeitos poderiam explicar as correlações mais baixas, uma vez que seus escores seriam responsáveis pela introdução de "ruído".

Uma análise mais cuidadosa do comportamento desses sujeitos é necessária. Essa análise

é conduzida em duas etapas. Na primeira delas apresento os escores de leitura e conhecimento de vocabulário que definiram as posições dos sujeitos nos gráficos. Em segundo lugar, tento explicar essas posições através dos desempenhos ou do perfil dos sujeitos considerando-se as outras variáveis.

Considero em primeiro lugar os sujeitos que tiveram escores baixos em conhecimento de vocabulário mas que conseguiram escores acima de 5 em leitura, dos quais se salientam principalmente os sujeitos 16, 31, 14 e 25.

A Tabela 4.31 abaixo apresenta os escores de cada um desses sujeitos em Leit1 e Leit2 e em conhecimento de vocabulário/inferência lexical, assim como os escores do grupo nessas duas variáveis, de forma que se possa avaliar esses resultados.

Tabela 4.31 – Sujeitos com escores baixos em conhecimento de vocabulário e escores altos em leitura

Sujeitos	TCV	Escore		
		LEIT1	LEIT2	INFREAL
16	-0,77	7,0	7,0	-0,63
31	-0,55	7,0	8,5	-0,44
14	-0,39	7,5	8,5	-0,36
25	-0,34	5,0	6,5	-0,42
Grupo				
Mínimo/máximo	-0,77/0,73	0,5/8,5	0,5/10,0	-0,77/0,91
Média	-0,002	5,63	6,59	-0,008
Intervalo de confiança	-0,10 – 0,10	5,20 – 6,45	5,96 – 7,21	-0,13 – 0,11

O sujeito 16, por exemplo, obteve um escore em vocabulário de -0,77 e de inferência -0,63. Comparando-se o primeiro escore com a média dos escores do grupo, pode-se concluir que esse é o escore mais baixo do grupo. Em leitura, por outro lado, seu escore foi 7,0 nos dois testes, mostrando um desempenho ligeiramente acima da média. O que merece ser salientado, entretanto, é que o seu desempenho em Leit1 e em Leit2 é exatamente o mesmo, apesar da diferença entre os dois, ou seja, as palavras inventadas no primeiro e as palavras reais no segundo. Mais ou menos o mesmo padrão de comportamento pode ser observado com os demais sujeitos da tabela, embora, nesses casos, haja uma diferença de desempenho, embora pequena, entre Leit1 e Leit2.

Uma possível explicação para esses escores pode ser encontrada através de uma análise mais cuidadosa do teste de leitura I e leitura II do sujeito. Essa análise mostra que o leitor indica desconhecer apenas três palavras além daquelas avaliadas pelos testes de vocabulário. O fato de conhecer outras palavras no texto pode ter sido suficiente para o entendimento do texto como um todo. A pergunta focalizando detalhes é a que ficou mais prejudicada pela falta das palavras avaliadas.

Ainda com relação ao leitor 16, observa-se que os escores de proficiência em LE (Cloze2A

e Cloze2B) assim como os outros escores de leitura em LE (Cloze2C) e vestibular (LeitV3) são muito baixos (1,1; 1,1; 1,7; 2,1), mostrando uma consistência na avaliação da leitura pelas três medidas. O único escore que realmente chama a atenção é o escore obtido no teste de conhecimento conceitual, ou seja, -0,62, que é o escore máximo, e que, juntamente com as outras palavras do texto, poderia ter sido usado para compensar seu vocabulário deficiente. Entretanto, esse escore máximo, que significa que os seis conceitos eram conhecidos, não o difere muito do resto do grupo como um todo, uma vez que o desempenho do grupo nesse teste foi considerado muito bom. Seu escore de inferência lexical também pode ser considerado bastante baixo (-0,74) ou seja, o mais baixo do grupo.

Com relação ao leitor 31, os outros escores de leitura em LE (Cloze2C e LeitV3) mostram desempenhos abaixo da média. Os escores em LM do vestibular, ou seja, FaseV1 e principalmente LeitV2, entretanto, foram acima da média, ou seja, 6,44 e 7,7, respectivamente, uma vez que o mínimo e máximo do grupo foram 3,81 a 8,56 para FaseV1 e 1,2 a 8,0 para LeitV2, com intervalos de confiança entre 5,53 e 6,0 para FaseV1 e 4,37 a 5,40 para LeitV2. Além disso, dois fatores chamam a atenção no perfil desse sujeito. Um deles é o seu escore de leitura em LM, mais especificamente o Cloze1B, que foi 9,0, quando o escore máximo foi 9,6, a média 7,80, e o intervalo de confiança entre 7,44 e 8,15, portanto, um escore bem acima da média. Um outro, sua habilidade de inferência. Esse leitor, juntamente com o leitor 14, o próximo a ser analisado, destacam-se por sua habilidade de inferência. Esses dois sujeitos apresentaram, na inferência de palavras, um comportamento totalmente imprevisível e difícil de se explicar, uma vez que, apesar de apresentarem um escore de conhecimento de vocabulário baixo, conseguiram inferir um número razoável de palavras (10 e 11, respectivamente), levando-se em conta o número de palavras conhecidas, que era muito pequeno, ou seja, 2 e 1, respectivamente. Esses resultados poderiam sugerir uma compensação, nesse caso, de habilidades de LM. Entretanto, uma análise mais cuidadosa do seu teste de leitura não mostra outras palavras grifadas além daquelas avaliadas no teste de vocabulário e além das inventadas, o que estaria indicando que, apesar de desconhecer as palavras de frequência baixa e média do teste de vocabulário, parece conhecer palavras de frequência alta, e, através delas, pode ter inferido as outras, exatamente como o leitor 16.

Com respeito ao leitor 14, além de seu escore de inferência, não se observa nenhum outro que mereça ser salientado. O mesmo comportamento dos dois leitores acima se observa com relação a esse leitor, ou seja, apenas desconhece as palavras do teste.

O comportamento do leitor 25 é o que menos chama a atenção nesse grupo, uma vez que seu conhecimento de vocabulário não é tão baixo e seu escore de leitura é o mais baixo entre os quatro sujeitos. Seu escore em LM do vestibular ou FaseV1 foi 5,69, quando o intervalo de confiança da média foi entre 5,53 e 6,0, portanto, se classificando, portanto, exatamente na média. O escore de LeitV2, entretanto, foi um escore bastante baixo, ou seja, 2,3, uma vez que o mínimo

do grupo foi 1,2 e o intervalo de confiança da média foi entre 4,37 e 5,40. Apesar desses resultados, um escore que se salienta no perfil desse sujeito é o de leitura em LM ou Cloze1B, que é 8,9, quando o escore máximo foi 9,6 e o intervalo de confiança entre 7,44 e 8,15, portanto, acima da média. Uma análise do teste de leitura também mostra que esse leitor não desconhecia outras palavras, além daquelas avaliadas no teste de vocabulário e, aparentemente através dessas palavras, pode ter chegado a uma compreensão global do texto. Sua resposta à pergunta de detalhes ficou prejudicada pela falta de conhecimento das palavras do teste de vocabulário.

Considero, a seguir, os sujeitos cuja situação se apresenta inversa a dos sujeitos acima, ou seja, embora tendo alcançado escores altos de conhecimento de vocabulário, seus escores de leitura são relativamente baixos. É o que acontece principalmente com os leitores 34 e 28. A Tabela abaixo 4.32 exhibe esses escores, e também os do grupo, para comparação.

Tabela 4.32 – Sujeitos com escores altos de conhecimento de vocabulário e escores baixos em leitura

Sujeitos	Escore			
	TCV	LEIT1	LEIT2	INFREAL
34	0,63	4,0	4,5	0,74
28	-0,49	5,0	5,5	0,26
Grupo				
Mínimo/máximo	-0,77/0,73	0,5/8,5	0,5/10,0	-0,77/0,91
Média	-0,002	5,63	6,59	-0,008
Intervalo de confiança	-0,10 – 0,10	5,20 – 6,45	5,96 – 7,21	-0,13 – 0,11

Como pode-se observar, o leitor 34 é o que chama mais atenção, uma vez que seu escore de vocabulário está muito acima da média, quase um dos mais altos, e seu escore de leitura está abaixo da média. A situação do leitor 28 é similar, embora seu escore de vocabulário seja um pouco mais baixo, mas ainda bem acima da média, e o de leitura um pouco mais alto, exatamente na média.

Analisando os escores do sujeito 34 em outras variáveis, observa-se, que as outras medidas de leitura em LE e LM são altas (7,2 no Cloze2C, 8,3 no vestibular e 8,2 no Cloze1B, 8,0 no LeitV2). Para o leitor 28, a situação é bastante similar. Também apresenta escores altos em outras medidas de LE (6,9 no Cloze2C, 6,0 no vestibular) e em leitura em LM (8,9 no Cloze1B, 6,5 na FaseV1 e 6,5 na LeitV2). Esse desempenho diferenciado apenas em Leit1 e Leit2 poderia ser indicativo de que algum problema externo à avaliação pode ter ocorrido, como por exemplo, falta de motivação do sujeito para realizar esse teste.

4.2.2.2 – Análise fatorial

Como salientei na seção anterior, a análise da matriz de correlações indicou a existência de agrupamentos de certas variáveis vs. o não agrupamento de outras, o que pode ser melhor investigado através do uso da análise fatorial. Como também mostrei no Capítulo III, a análise fatorial é uma técnica para análise de escores em termos de fatores subjacentes. Basicamente, ela se baseia na pressuposição de que, em qualquer teste ou escore, há provavelmente um ou mais traços subjacentes que estão sendo avaliados. O objetivo é construir esses fatores subjacentes e decompor as variâncias do escore em termos da correlação dos fatores e dos escores observados.

O método de rotação utilizado neste estudo foi o Varimax. Apresento, na Tabela F.2 o padrão dos fatores, que mostra a contribuição de cada variável, e as Figuras 4.8 e 4.9, resultantes da análise, exibindo as variáveis distribuídas entre os Fatores 1 e 2 e entre os Fatores 2 e 3, respectivamente.

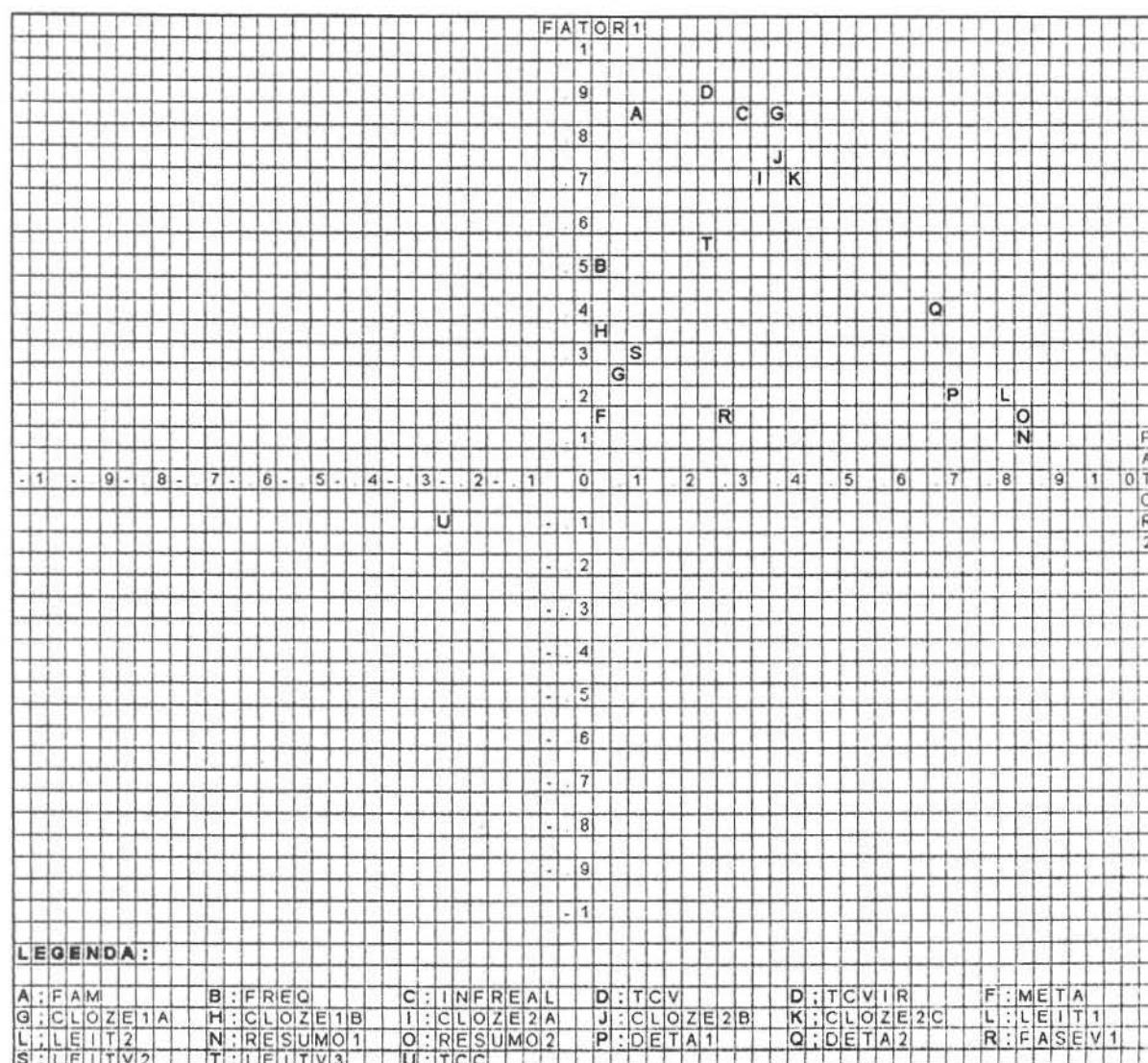


Figura 4.8 – Análise fatorial: Fator 1 × Fator 2

Como pode-se observar, tanto pelo padrão dos fatores como pela Figura 4.8, três grupos de variáveis se salientam, embora dois grupos apresentem uma maior "explicação". Analisando-se a coluna do Fator 1, observa-se que as maiores contribuições são dadas pelas variáveis TCVI, TCV, FAM, INFREAL, Cloze2B, Cloze2C, Cloze2A, Leitiv3, TCC e FREQ. Essas variáveis, portanto, carregam no Fator 1. A única variável que aparece na parte de baixo (em princípio estaria contraposta às demais) do gráfico é a variável TCC, ou conhecimento conceitual. Seu posicionamento, entretanto, deve ser deslocado para a parte inferior do gráfico porque essa variável foi incluída na análise fatorial com o sinal invertido. Da maneira como está apresentada ela se constitui uma medida de não conhecimento-conceitual. Corrigindo-se o sinal, automaticamente ela volta para a parte de cima do gráfico, juntamente com as variáveis de conhecimento de vocabulário.

Analisando-se a coluna do Fator 2, observa-se que as maiores contribuições são dadas pelas variáveis Resumo1 e Resumo2, Leit1 e Leit2, Deta1 e Deta2.

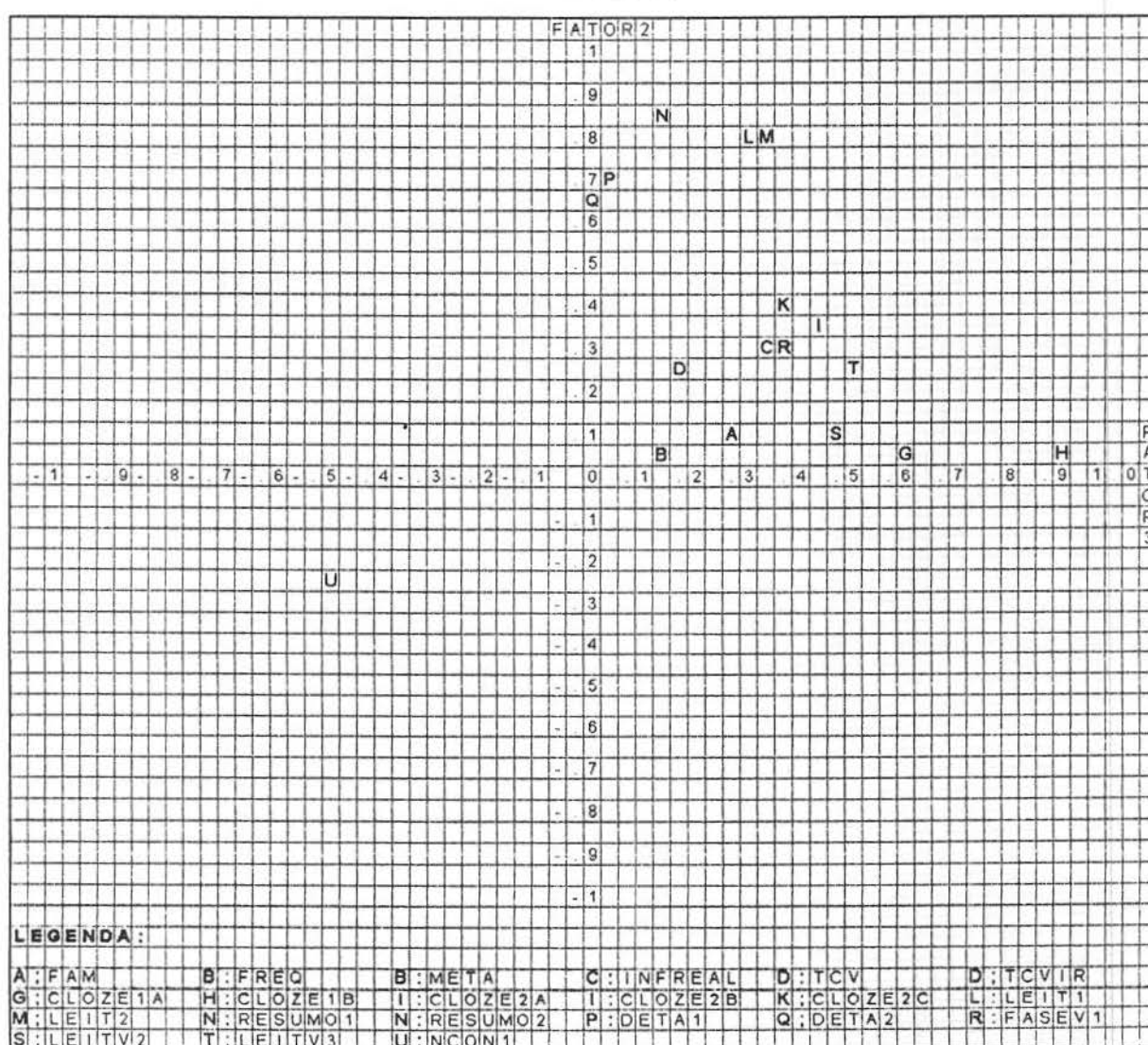


Figura 4.9 – Análise fatorial: Fator 2 × Fator 3

Analiso, a seguir, a Figura 4.9, acima, que apresenta os dados sob um outro ângulo, ou

seja, as variáveis distribuídas entre os Fatores 2 e 3. A existência de três grupos de variáveis parece se confirmar também por esse gráfico. Analisando-se a coluna do Fator 3, nota-se que as maiores contribuições são dadas pelas variáveis Cloze1B, Cloze1A, LeitV2 e FaseV1, sendo a maior, ou a que mais carrega no Fator 3 a variável Cloze1B.

É importante salientar, nesse caso, que apesar da existência de subagrupamentos entre as variáveis, todas as variáveis formam um único grupo, contidas no lado direito do gráfico, entre os eixos vertical e horizontal. Isso mostra que embora se possa identificar pelo menos três subgrupos dentro desse grupo, eles não se contrapõem em nenhum momento, mas se somam. A existência de três grupos de variáveis, portanto, poderia ser indicativa da existência de três fatores comuns às 21 variáveis, ou seja,

1. Fator 1, que apresenta um componente de vocabulário e de proficiência lingüística fortes, e que poderia ser denominado de componente lingüístico. Todos os testes cloze, mesmo o de leitura, parecem carregar mais nesse fator. Entretanto, juntamente com esse fator, há a variável TCC ou conhecimento conceitual, com uma contribuição baixa (0,16).
2. Fator 2, composto por Leit1 e Leit2 e suas subpartes, que poderia ser denominado de componente de leitura. Ele envolveria também a variável Concepções de leitura ou Meta. Apesar de baixa, a contribuição maior da variável Meta ainda é para esse fator (0,12).
3. Fator 3, ou seja, composto pelos testes de leitura em LM, incluindo os do vestibular, e que poderia constituir o componente de LM.

4.3 – Discussão dos resultados

Esta seção discute os resultados da análise, conduzida em duas etapas, apresentada na Seção 4.2 deste capítulo. Na primeira delas ou intracomponentes buscou-se uma caracterização do grupo de sujeitos deste estudo e também individualmente em cada um dos componentes para um posterior cruzamento na análise intercomponentes. Discuto, a seguir, os resultados das duas análises conjuntamente, usando as perguntas de pesquisa como elementos direcionadores dessa discussão.

A primeira pergunta de pesquisa, ou seja,

1. Como se caracteriza a relação entre a competência lexical e a compreensão enquanto produto da leitura?

investiga a natureza da relação entre competência lexical e compreensão em leitura. Para responder essa questão apresento uma caracterização do grupo de sujeitos desta investigação com relação à sua competência lexical, embora esses dados focalizem essa competência enquanto conhecimento,

estático, sob o ponto de vista de produto. Essa caracterização é complementada pelos resultados do estudo que focaliza o processo da leitura, na discussão que apresento no Capítulo VI ou na discussão comparativa.

Apesar de se afirmar frequentemente que o conhecimento lingüístico ou mais especificamente lexical do leitor em LE é limitado, não se tem notícias na literatura de estudos que apresentem caracterizações do vocabulário de leitores em LE. Para que um parâmetro aproximado de comparação com os resultados deste estudo pudesse ser estabelecido, usei os dados do grupo de falantes nativos que serviram de critério para alguns dos testes de vocabulário. Esses dados, entretanto, não foram analisados estatisticamente como os dos sujeitos, mas apenas qualitativamente. Embora ciente de que a proficiência em LE seja distinta daquela apresentada pelo falante nativo, considero esses dados importantes na medida em que dão consistência ao conceito *rico* de vocabulário proposto por Richards (1976), que, como o autor mesmo reconhece, é baseado em pressuposições sobre o conhecimento do falante nativo e não parece ter sido "testado" em situações reais, além de dar validade à bateria de testes que são sua operacionalização. Embora os resultados dos falantes nativos mostrem diferenças entre si com relação a algumas palavras, ou seja, que o conhecimento de algumas palavras é mais completo do que de outras, pode-se afirmar, pelo menos para as palavras avaliadas, que um conceito *rico* se mostrou bastante adequado. Nos subtestes que compõem esse conhecimento, por exemplo, flexão/derivação, significado, colocações, dentre outros, os falantes nativos raramente deixam itens sem resposta e as respostas apresentam uma profusão de detalhes. As muitas lacunas apresentadas pelos falantes não nativos em seus testes não impedem que se considere adequado um conceito *rico* de vocabulário também para uma situação de LE, uma vez que fortes associações entre os vários subtestes de conhecimento de vocabulário foram identificadas através da AC, conforme se verá mais adiante. Essas lacunas, dessa forma, poderão ser consideradas, com mais segurança, como insuficiências no conhecimento desses sujeitos e não limitações do instrumento.

A análise dos resultados dos testes de vocabulário revelou uma forte associação não apenas entre os testes de familiaridade e frequência, mas principalmente entre os vários subtestes que compõem o conhecimento de vocabulário e entre esses subtestes entre si. Em outras palavras, isso significa dizer que os sujeitos que identificaram a palavra como familiar em geral conheciam seu significado, além de saber também associar, colocar, flexionar. Apesar dessa associação, alguns dos testes se mostraram mais difíceis que outros, ou seja, parece haver um contínuo de dificuldade entre os testes.

O teste que teve um maior número de acertos foi o teste de classe gramatical, podendo, portanto, ser considerado o mais fácil. Esse resultado é curioso, uma vez que esse teste exigia a classificação das palavras em suas classes gramaticais, e, portanto, exigia o uso de metalinguagem. O teste de contexto gramatical, onde contextos possíveis e impossíveis eram apresentados para

escolha, foi considerado mais difícil pelos sujeitos, embora não exigisse o conhecimento de metalinguagem. Apesar de apresentarem níveis diferentes de dificuldade, os dois testes apresentam uma grande associação, como pode ser demonstrado. O teste de significado foi o segundo mais fácil, seguido pelo de derivação/flexão. As palavras não foram definidas, mas sim "traduzidas", uma vez que o tipo de resposta mais comum foi um sinônimo em português. Os testes mais difíceis foram os de associação e de colocação.

Uma explicação para a relativa facilidade de alguns testes e dificuldade de outros poderia ser relacionada ao tipo de abordagem de ensino de LE a que esses sujeitos têm sido submetidos. No caso da derivação/flexão, a facilidade poderia estar associada à abordagem instrumental, que enfatiza prefixos e sufixos. No caso do significado, esse resultado já podia ser previsto, uma vez que as abordagens em geral focalizam o significado como o único aspecto relevante. A dificuldade maior nos testes de associação e colocação também poderia ser explicada por essas abordagens, que em geral são atomistas, focalizando as palavras isoladamente, sem a consideração de suas redes semânticas e possibilidades de colocações.

A relação entre os aspectos familiaridade e frequência tem sido geralmente identificada na literatura, uma vez que, para falantes nativos, as palavras mais frequentes tendem a ser mais conhecidas. A análise conduzida nos dados dos sujeitos mostra uma grande associação entre os dois aspectos, ou seja, que as palavras mais familiares, na percepção dos sujeitos, ou sua auto-avaliação, também são as mais frequentes. Entretanto, muitas das palavras consideradas frequentes pelos nativos não o são para os não nativos, uma vez que a escala de frequência dos dois grupos é diferente. Para os falantes nativos, as palavras de alta frequência são aquelas que compõem seu vocabulário básico. Essas palavras têm sido denominadas de palavras básicas (*core vocabulary*) ou ainda de vocabulário procedimental (*procedural vocabulary*) (vide Robinson 1989); são palavras de uso generalizável, não pertencendo a nenhum esquema em particular. Elas são, portanto, mais difíceis de serem inferidas, uma vez que são usadas na construção de outros significados.

Neste estudo, portanto, o que realmente parece distinguir o vocabulário dos dois grupos é a frequência das palavras conhecidas/desconhecidas. Enquanto o falante nativo conhece palavras de alta frequência, podendo apresentar problemas com palavras de baixa frequência, o falante não nativo parece apresentar problemas especialmente no reconhecimento de palavras de alta frequência, além da dificuldade com as palavras de baixa frequência.

Esse aspecto é mais uma vez observado e parece se confirmar através dos resultados da Tarefa I do teste de leitura, ou seja, a tarefa de grifar palavras. Além das palavras selecionadas para o teste de vocabulário, muitos leitores (pelo menos 14) identificaram como desconhecidas outras palavras. Uma análise dessas palavras mostra que são, em sua maioria, palavras de alta frequência. Esse resultado é importante na caracterização da competência lexical do grupo, mostrando que outras palavras, além daquelas selecionadas para os testes de vocabulário, são também desconhecidas. Isso

significa, em termos do grupo estudado, que se pode identificar, na realidade, dois sub-grupos: um deles, que tem seus problemas de vocabulário limitados às palavras do teste de vocabulário, e o outro, que além das palavras do teste, ainda desconhece outras palavras no texto.

Assim, para que uma caracterização mais completa dos sujeitos deste estudo possa ser feita, haveria a necessidade de se levar em conta não apenas as palavras avaliadas pelos testes de vocabulário mas principalmente as outras palavras grifadas como desconhecidas. Embora essas palavras terem sido deixadas de lado no teste de conhecimento de vocabulário, não significa que os escores de vocabulário obtidos não representem o conhecimento dos sujeitos, uma vez que se observou que aqueles que obtiveram escores baixos nos testes de vocabulário também, de uma maneira geral, apesar de algumas poucas exceções, também foram os que grifaram um número maior de palavras desconhecidas.

Os resultados dos cruzamentos entre as medidas de vocabulário geradas a partir dessa análise e as medidas de compreensão em leitura mostram uma correlação forte entre a leitura e o vocabulário. Algumas observações, entretanto, merecem ser feitas para se caracterizar melhor essa relação.

São três as medidas de conhecimento de vocabulário usadas na análise. A primeira considera apenas o conhecimento de vocabulário, a segunda apenas a inferência, enquanto a terceira considera as duas primeiras conjuntamente. Essas medidas foram geradas através da AC que considerou o componente vocabulário, e confirmadas posteriormente pela análise de correlações assim como pela análise fatorial. Além dessas medidas específicas de conhecimento lexical, tem-se também duas medidas de proficiência lingüística mais global, resultantes dos testes do tipo cloze em suas versões A e B. Essas duas medidas foram obtidas corrigindo-se as duas versões segundo critérios diferentes. Na primeira versão usou-se o critério de palavra exata e na segunda um critério que levava em conta as respostas de falantes nativos.

São três as medidas de compreensão usadas neste estudo. A primeira, em duas versões, é constituída pelo teste de leitura I e II. A primeira versão inclui as palavras inventadas e a segunda as palavras reais. Uma outra medida é o tipo cloze em sua versão C, corrigida segundo um critério de palavra aceitável. Assim, um mesmo teste, corrigido por critérios diferentes, fornece medidas diferentes.

A existência de uma relação significativa entre vocabulário e compreensão em leitura pode ser inferida a partir das correlações positivas entre as medidas de proficiência em LE, conhecimento de vocabulário, conhecimento de vocabulário/inferência, e três medidas de compreensão em leitura em LE, confirmadas posteriormente pela análise fatorial e análise de correspondência.

Considerando em primeiro lugar apenas as medidas de proficiência (Cloze2A e Cloze2B) e conhecimento de vocabulário/inferência, essas correlações são bastante fortes. As correlações mais fortes são observadas, portanto, entre o conhecimento de vocabulário/inferência e a versão B do

cloze, isto é, aquela que teve como critério de correção os testes de falantes nativos. Entretanto, se for considerada apenas a inferência, observa-se que a correlação é ainda mais alta (0,87 entre inferência e versão B do cloze ou proficiência).

Cruzando-se as três variáveis conhecimento de vocabulário e inferência com as medidas de leitura também resulta, em geral, em correlações bastante relevantes. Essas correlações são 0,77 entre o teste de leitura do tipo cloze (Cloze2C) e o conhecimento de vocabulário e inferência; 0,65 entre o teste de leitura do vestibular (Leitv3) e o conhecimento de vocabulário e inferência e finalmente 0,48 e 0,47 entre os testes de leitura I e II e conhecimento de vocabulário e inferência. Entretanto, se considerar apenas a inferência, observa-se correlações ainda mais fortes (0,83 entre inferência e o teste de leitura do tipo cloze, 0,71 entre inferência e leitura do vestibular e finalmente 0,55 entre inferência e leitura I e 0,56 entre inferência e leitura II).

Pelo menos três observações precisam ser feitas com relação a esses resultados. Em primeiro lugar, do cruzamento das medidas de conhecimento de vocabulário/inferência e compreensão em leitura, todas as correlações apresentadas podem ser consideradas relevantes. Muitas delas são fortes. As menores correlações são observadas entre o teste de leitura I e II e as medidas de vocabulário (0,48 e 0,47 respectivamente). As correlações com as medidas de proficiência se mostram um pouco mais altas (0,60 e 0,61). Quando se considera apenas a inferência, as correlações também são mais fortes (0,55 e 0,56). Esse resultado é difícil de ser explicado, à primeira vista, uma vez que os testes de leitura I e II foram elaborados com base no mesmo texto de onde foram extraídas as palavras para os testes de vocabulário. Além disso, observa-se também correlações muito fortes entre as duas versões do teste de leitura, ou seja, com palavras inventadas e palavras reais, mostrando que os sujeitos que se saíram bem em uma versão também se saíram bem na outra.

Entretanto, uma análise mais detalhada dos escores de leitura e dos escores de vocabulário mostra que os sujeitos que tiveram escores fracos de conhecimento de vocabulário foram também aqueles que grifaram como desconhecidas outras palavras no texto. Por outro lado, os sujeitos que tiveram escores altos de conhecimento de vocabulário foram aqueles que não grifaram palavras além daquelas usadas para os testes de vocabulário. Além disso, aqueles que obtiveram escores altos de leitura no teste de leitura I, ou seja, com as palavras inventadas, também apresentaram escores altos no teste de leitura II, com as palavras reais. Por outro lado, aqueles que mostraram escores baixos na versão com palavras inventadas, também apresentaram escores baixos na versão com palavras reais, mostrando que as palavras do teste não foram decisivas no desempenho, mas sim as outras palavras do texto. Por sua vez, a análise do teste de leitura mostra que aqueles que conseguiram escores acima de cinco no teste de leitura foram também aqueles que faziam parte do grupo que somente tiveram problemas com as palavras do teste, enquanto aqueles que tiveram escores abaixo de cinco foram aqueles que desconheciam palavras além daquelas avaliadas pelos testes. Convém

lembrar que apenas as palavras do teste foram levadas em conta para gerar-se o escore de conhecimento de vocabulário. Se tivessem sido levadas em conta todas as palavras, os escores baixos de conhecimento de vocabulário seriam ainda mais baixos, as correlações entre esse conhecimento e o teste de leitura II seriam mais altas e as diferenças entre leitura I e leitura II não seriam tão altas.

O que esses resultados parecem sugerir é que o desconhecimento das palavras do teste juntamente com outras palavras é o que está levando a desempenhos fracos na compreensão. Entretanto, apenas o desconhecimento das palavras do teste pode não ser suficiente para explicar o desempenho fraco. O desempenho muito bom do grupo no teste de conhecimento conceitual mostra que o conhecimento do assunto do texto era bastante bom e que esse conhecimento poderia ter sido usado para uma compensação do vocabulário desconhecido. Entretanto, essa compensação não se observou com a maioria dos sujeitos. Aqueles que tiveram escores abaixo de cinco nos testes de leitura não conseguiram compensar seu conhecimento de vocabulário. Como os conceitos foram apresentados em português, pode-se dizer que sua ativação dependia de um nível de conhecimento de vocabulário maior do que aquele apresentado por esses sujeitos. Isso significa dizer que a grande diferença entre os leitores com escores abaixo de cinco e entre leitores com escores acima de cinco está principalmente nas outras palavras do texto, além daquelas dos testes de vocabulário e que o desconhecimento dessas palavras também impediu a compensação com o conhecimento prévio. Esses resultados parecem ser indicativos da existência de um nível limiar de vocabulário, foco da pergunta de pesquisa 3, e que será discutido mais detalhadamente mais adiante.

Uma segunda observação que pode ser feita a partir dos resultados das correlações acima apresentadas é que, mesmo se considerar apenas as palavras dos testes, verifica-se que as medidas de inferência, quando cruzadas com os testes de leitura, parecem apresentar correlações mais fortes do que aquelas apresentadas pelo cruzamento do conhecimento de vocabulário/inferência ou apenas conhecimento de vocabulário com as medidas de leitura. Esses resultados são consistentes entre as três medidas, podendo sugerir que a medida inferência é mais abrangente do que o conhecimento de vocabulário, refletindo melhor a competência lexical do sujeito. Isso significa afirmar que, embora a correspondência entre inferência e conhecimento de vocabulário seja bastante grande (mostrada através da AC) e também as correlações sejam bastante fortes (0,88), ou seja, sujeitos que têm um maior conhecimento de vocabulário conseguem inferir mais, ela poderia incluir outros fatores ou habilidades, que poderão ser mais bem examinadas no estudo de processo apresentado no Capítulo V.

Uma terceira e última observação a ser feita a respeito dos resultados acima é que a versão B do cloze ou o teste de proficiência em sua versão B é a que parece apresentar as correlações mais fortes, ou seja, aquela corrigida segundo o critério de palavra aceitável dos falantes nativos. Essa observação parece sugerir que essa versão é mais apropriada do que a versão A, corrigida pelo critério de palavra exata. Resultados de outras pesquisas também parecem apontar para a mesma

direção. Entretanto, sua maior correlação é com o teste de leitura Cloze2C, sugerindo talvez o que tem sido afirmado na literatura, ou seja, de que o teste do tipo cloze tem um viés para habilidades de nível baixo, ou seja, está mais ligada a vocabulário e inferência lexical. Devo lembrar, no entanto, que os testes do tipo cloze usados neste estudo não foram elaborados segundo um critério de aleatoriedade, mas levaram em conta o processamento da linguagem. Portanto, estudos confirmatórios posteriores seriam necessários para se verificar se essa mudança teve um resultado positivo.

Apesar da existência de uma relação bastante forte entre competência lexical e compreensão em leitura em LE, foco da pergunta de pesquisa 1, entretanto, as correlações acima apresentadas, principalmente sobre a inferência e também sobre o conhecimento conceitual parecem mostrar que outros fatores podem ser importantes nessa relação. Uma avaliação mais detalhada da contribuição de outros fatores, tais como a leitura em LM e conhecimento conceitual e diferentes concepções em leitura em LE é o foco da segunda pergunta de pesquisa, isto é,

2. Qual a contribuição, para essa relação, de fatores como proficiência em LE, diferentes concepções de leitura em LE, conhecimento conceitual e leitura em LM?

Para responder a essa pergunta de pesquisa, além das medidas de conhecimento de vocabulário/inferência e medidas de leitura, este estudo leva em consideração outras variáveis tais como conhecimento conceitual, proficiência em leitura em LM e concepção de leitura. Com base nas correlações obtidas entre essas variáveis e as três medidas de leitura, pode-se afirmar que, com exceção da última variável, todas as outras apresentam correlações relevantes.

Especificamente com relação ao conhecimento conceitual, é necessário lembrar que somente o cruzamento dessa variável com as medidas de leitura I e II são pertinentes, um vez que o teste de conhecimento conceitual foi elaborado considerando-se os conceitos do texto usado para esse teste, e, portanto, só faz sentido com relação a esse teste. Embora não tão fortes quanto as correlações entre conhecimento de vocabulário e leitura, correlações relevantes também podem ser observadas entre conhecimento conceitual e os testes de leitura I (0,52) e II (0,48), mostrando que no primeiro, com palavras inventadas, a contribuição do conhecimento conceitual pode ter sido maior. Uma outra explicação para essas correlações é o fato de o grupo ser muito homogêneo com relação ao conhecimento conceitual. Apesar de alguns sujeitos estarem nos extremos desse conhecimento, ou seja, terem acertado todos os seis conceitos (seis sujeitos) ou errarem todos (1), esses casos são poucos (7) em relação ao grupo como um todo. Os acertos do grupo como um todo foram acima de 75%. Entretanto, outros resultados dessa análise mostram que uma possível compensação através desse conhecimento não foi possível para todos os sujeitos, mas apenas para aqueles que tinham um conhecimento lexical mais desenvolvido.

Considero, a seguir, as relações entre metacognição e leitura, ou mais especificamente

entre os escores que refletem uma concepção de leitura mais *local* ou mais *global*, inferidas a partir do uso de estratégias metacognitivas e as três medidas de leitura em LE. Como pode-se observar, as correlações entre essas variáveis são bastante baixas, não mostrando nenhuma correlação significativa ou associação direta com nenhuma variável.

Antes de apresentar possíveis explicações para esses resultados, entretanto, considero importante caracterizar mais detalhadamente o grupo com relação a esses fatores metacognitivos e às suas concepções de leitura.

As estratégias metacognitivas, subdivididas em quatro subgrupos, isto é, confiança, reparo ou correção, eficiência e dificuldade foram avaliadas pelos sujeitos numa escala do tipo Likert, de cinco categorias. Para o grupo em geral, certas respostas são mais comuns, o que não significa dizer que todos os sujeitos respondem da mesma maneira.

As seis estratégias de confiança revelam a confiança do sujeito enquanto leitor. Para o grupo como um todo, as afirmações ou estratégias *ser capaz de usar conhecimento prévio e monitorar a compreensão* são vistas como mais importantes do que as afirmações *ser capaz de reconhecer a diferença entre pontos principais e detalhes*, *ser capaz de relacionar as informações antigas às novas* e *ser capaz de questionar a veracidade do texto*, que por sua vez são consideradas mais importantes que a afirmação *ser capaz de fazer previsões sobre o que está para vir no texto*.

Nas cinco estratégias de correção ou reparo, usadas quando há falhas compreensão os resultados parecem mostrar também um contínuo de importância, que dá destaque à estratégia *reler a parte problemática*. Essa seria, na opinião do grupo, a estratégia de correção mais eficiente. As outras vêm na seguinte ordem: *continuar a leitura, voltar ao ponto imediatamente anterior à parte problemática e reler a partir daí*. A estratégia *procurar as palavras desconhecidas no dicionário* tem um padrão de respostas na categoria *neutro*. Várias interpretações poderia ser dadas a esse tipo de resposta: o leitor pode ser indiferente com relação a essa estratégia, ou não ter uma opinião a respeito. Essa resposta ainda pode significar que a estratégia pode ser usada algumas vezes, dependendo do texto, de sua dificuldade e do propósito da leitura. Entretanto, o grupo como um todo discorda com a afirmação *ao encontrar um problema, desistir e parar de ler*.

Esses resultados parecem bastante interessantes, coincidindo com a maneira como essas estratégias são vistas pela literatura. As três primeiras são consideradas estratégias eficientes e as duas últimas, mais especialmente a 5, geralmente caracterizam o leitor ineficiente.

Levando-se em conta as 17 estratégias de eficiência, que, na opinião do leitor, tornam a leitura eficiente, tem-se também um contínuo em ordem decrescente de importância, sendo os aspectos mais importantes a idéia central e conhecimento prévio, seguidos de palavra/significado e organização do texto, que por sua vez são seguidos bem de perto por sintaxe e detalhes de conteúdo. Os menos importantes são os de som/letra.

A ênfase dada pelo grupo a estratégias mais *globais* pode ser explicada, uma vez que esses

alunos estão sendo expostos, em suas aulas de instrumental, a estratégias mais *globais* em detrimento de estratégias mais *locais*. Esse fato também poderia explicar a não saliência dada aos aspectos palavra/significado e som/letra. Uma análise mais detalhada dessa categoria mostra que das cinco estratégias da categoria, apenas a estratégia *reconhecer o significado das palavras* recebeu um padrão de respostas no nível 2, ou seja, *concordo*. As estratégias *entender o significado de cada palavra* e *adivinhar o significado de palavras* apresentaram padrões de resposta entre os níveis 2 e 3, mostrando uma indecisão entre um *concordo* e um *neutro*. Uma interpretação para esses resultados poderia estar em uma deficiência do próprio questionário, uma vez que em nenhum momento se usa o termo *vocabulário*, mas sim *reconhecer o significado de palavras*, *adivinhar o significado de palavras*. O sujeito poderia não considerar eficiente *inferir palavras* mas já conhecer essas palavras. Além disso, o termo *reconhecimento de palavras* pode ser considerado um termo mais específico ou mais especializado, podendo não significar ao sujeito a associação entre palavra e significado. O uso, no questionário, de *entender o significado*, de todos os termos, o mais objetivo, é complementado por *de cada palavra*, distorcendo a afirmação. O sujeito pode achar que, embora o significado de palavras seja importante, não seja necessário *entender o significado de cada palavra*.

Com relação ao uso do dicionário, o padrão de respostas foi *neutro*, o que poderia ser interpretado como um reconhecimento de sua necessidade apenas em algumas situações, mas especificamente quando o texto é difícil e o propósito da leitura exige níveis mais precisos de significação. O mesmo resultado, ou seja, nível 3, foi obtido para a estratégia de correção acima, que também refere-se ao uso do dicionário.

A análise das estratégias de dificuldade mostra, como o nome diz, as dificuldades dos sujeitos, enquanto grupo, na leitura. Levando-se em conta todas essas estratégias, tem-se, também, um contínuo em ordem decrescente, que mostra maior dificuldade no reconhecimento de palavras, sendo a sintaxe neutra. Todas as outras afirmações são menos problemáticas para o grupo.

Embora não haja nenhuma problema com nível de resposta 1, ou seja, *concordo plenamente*, a mais problemática é o reconhecimento das palavras. As estruturas gramaticais, ou a sintaxe, é de dificuldade neutra. Várias interpretações poderiam ser dadas a esse tipo de resposta: o leitor pode ser indiferente ou indeciso com relação a essa estratégia, ou ainda não ter uma opinião a respeito. Essa resposta ainda pode significar que a estratégia pode ser usada algumas vezes, dependendo do texto, de sua dificuldade e do propósito da leitura.

Todas as outras estratégias ficam predominantemente entre os níveis de desacordo, e, portanto, podem ser interpretadas como menos problemáticas. Essas estratégias são *a organização do texto*, *entender o significado global do texto*, *os sons das palavras* e *a pronúncia das palavras*, tendo como aspectos subjacentes a idéia central, o conhecimento prévio, a organização do texto e

som/letra.

Um cruzamento entre os dados da categoria dificuldade com os da categoria eficiência revela uma coerência do grupo em vários aspectos. Em primeiro lugar, com relação aos aspectos de som/letra, eficiência e dificuldade têm o mesmo padrão de resposta, ou seja, o grupo não considera importantes as estratégias de som/letra e também não apresenta problemas com essas estratégias. O mesmo acontece com o reconhecimento de palavras. O grupo considera eficiente reconhecer palavras e também identifica essa estratégia como problemática.

Entretanto, no que se refere às estratégias que têm como subjacentes a organização do texto, o conhecimento prévio e a idéia central, eficiência e dificuldade estão claramente marcadas mas estão em extrema oposição. Embora o grupo considere eficientes estratégias como *a organização do texto, entender o significado global do texto*, não as considera problemáticas.

Também é interessante notar que as categorias som/letra e sintaxe, que no que se refere à eficiência são identificadas pelo nível 2 e predominantemente 3, mantendo esses mesmos padrões quando se refere à dificuldade.

Várias poderiam ser as explicações para esses resultados. Em primeiro lugar, eles poderiam significar que tanto uma concepção *local* como *global* pode estar associada ao sucesso na compreensão e não apenas uma concepção *global*, como a literatura parece reconhecer. Uma segunda explicação é que os escores que mostram as concepções de leitura na realidade não classificam o leitor como apresentando exclusivamente uma concepção ou outra, ou seja, totalmente *global* ou *local*, mas mostram um contínuo, ou seja, os leitores na realidade são ora *globais* ora *locais*, combinando os dois tipos de estratégias. Ainda uma terceira explicação poderia estar em uma limitação do instrumento de coleta, ou seja, no questionário metacognitivo ou ainda na amostragem para a coleta. Com relação ao questionário, observa-se que muitas das respostas, principalmente aquelas na categoria *neutro*, parecem ser resultantes do fato de que as estratégias de leitura podem variar, dependendo do propósito da leitura. Entretanto, o questionário não é sensível a esse aspecto, uma vez que não considera essas situações diferentes da leitura. Com relação à amostra coletada, convém ressaltar que quando os dados foram coletados, todos os sujeitos (os de nível II de Inglês Instrumental mais, os de nível I menos), já haviam sido expostos às aulas de instrumental e a uma abordagem que enfatizava o uso de estratégias, principalmente estratégias mais globais, como já afirmei antes. Dessa forma, os dados já poderiam estar refletindo a influência dessas aulas e não mais apenas as concepções de leitura ou tendências originais ou as estratégias realmente usadas por esses sujeitos.

A última variável considerada é a proficiência em leitura em LM. Essa variável é representada por duas medidas, ou seja, pelos dois testes do tipo cloze em suas versões ou Cloze1A e Cloze1B. No primeiro caso, o critério de correção é de palavra exata. Esse tipo de teste com esse critério de correção tem sido identificado na literatura como uma medida que extrapola a leitura, ou

seja, uma medida de proficiência mais global. Correlações relevantes podem ser observadas entre a leitura em LM, medida pela versão Cloze1B e o teste de leitura Cloze2C (0,58) e principalmente entre a leitura em LM e o escore de leitura em LE do vestibular (0,69). Curiosamente, mais uma vez o teste de leitura II apresenta as correlações mais baixas (0,42). As correlações mais relevantes apresentadas pela versão B do cloze (0,58 contra 0,42 da versão Cloze1A), corrigida pelo critério de palavra aceitável, parecem mostrar ser essa versão mais apropriada enquanto medida de leitura.

Embora as correlações das medidas de leitura em LM não sejam tão fortes como as de conhecimento de vocabulário/inferência ou medidas de inferência, elas parecem mostrar que a proficiência em leitura em LM pode ter uma influência na leitura em LE.

Todos os resultados discutidos até agora são confirmados pela análise fatorial, que considera o grupo como o todo, e pela análise dos escores de leitura e vocabulário, que considera indivíduos. Uma discussão dos resultados dessas duas análises, vêm, portanto, não apenas complementar as respostas às perguntas de pesquisa 1 e 2 mas também responder à pergunta 3.

A análise fatorial parece revelar três fatores comuns às 21 variáveis, embora os Fatores 1 e 2 apresentem maior "explicação". O primeiro deles, ou Fator 1, é o componente lingüístico, uma vez que apresenta um componente forte de conhecimento de vocabulário e proficiência lingüística. O teste de leitura em LE Cloze2C também parece carregar mais nesse fator. O Fator 2, representado pelos testes de leitura I e II e suas subpartes, é o componente de leitura. Ele envolveria também as variáveis concepções de leitura e conhecimento conceitual. Apesar de baixa, a contribuição maior da variável concepções de leitura ainda é para esse fator (0,12), assim como da variável conhecimento conceitual (0,16). O Fator 3, representado pelos testes de leitura em LM incluindo os escores em LM do vestibular, é o componente de LM. Esse fator é o mais fraco deles, podendo ser considerado conjuntamente com o Fator 2. Os Fatores 2 e 3 formam, praticamente, um único subgrupo, formando um único Fator 2, ou seja, um componente de leitura que engloba proficiência em leitura em LM, assim como conhecimento prévio e diferentes concepções de leitura. Com base nesses resultados, pode-se concluir que a compreensão em leitura é resultante de dois fatores, que contribuem com pesos diferentes: um fator lingüístico e um fator de leitura, que engloba proficiência em LM, concepções de leitura e conhecimento conceitual.

Pelo menos duas observações adicionais precisam ser feitas com relação a esses resultados, sendo as duas referentes a uma questão acima levantada acima relacionada aos testes do tipo cloze. Analisando-se o Fator 1, observa-se que dentre as medidas de proficiência e conhecimento de vocabulário inclui-se também o teste de leitura do tipo cloze. Essa evidência, mais uma vez, pode ser indicativa de que esse tipo de teste é mais uma medida de habilidades de baixo nível do que propriamente de leitura. Entretanto, como o teste tipo cloze deste estudo foi desenvolvido de maneira diferenciada, isto é, levando-se em conta o processamento da linguagem, será necessário um outro estudo para confirmar esse aspecto. Uma outra observação pode ser feita a partir da análise da coluna

do Fator 3. A maior contribuição ou a que mais "carrega" no Fator 3 é a variável Cloze1B, mostrando que das duas versões, essa é a que melhor representa a proficiência de leitura em LM.

Os resultados discutidos até agora, apesar de importantes, ainda não esclarecem um aspecto da relação entre o conhecimento de vocabulário e a leitura em LE, foco da terceira pergunta de pesquisa:

3. Há evidências de um nível limiar de competência lexical para a compreensão enquanto processo e produto da leitura?

ou um nível de competência lexical mínima para que o desempenho em leitura seja eficiente. A investigação dessa questão, em termos absolutos, pressupõe a avaliação do número de palavras ou extensão total do vocabulário do indivíduo, como salientei no Capítulo II. A possível existência de um nível mínimo de competência lexical é, neste estudo, investigada em termos relativos, ou seja, considerando a competência do sujeito com relação às palavras do texto do teste de leitura; essa competência é usada para generalizações sobre o seu léxico total. Para essa análise, foram cruzados os escores do teste de conhecimento de vocabulário e os escores do teste de leitura II, ou seja, o teste com as palavras reais.

Embora essa análise mostre de maneira bastante satisfatória a correspondência entre escores de conhecimento de vocabulário e escores de leitura, ou seja, escores baixos em leitura parecem corresponder a escores baixos em conhecimento de vocabulário e vice-versa, não há, como salientado por Scott (1991) nenhuma evidência de um *salto* ou de uma *saliência* (*bump*) que, para Scott (1990), parece necessária para caracterizar um nível limiar. Além disso, apesar de um padrão de correspondência entre escores altos de leitura e de conhecimento de vocabulário e vice-versa, observa-se que pelo menos para 6 sujeitos dos 49, esse padrão não se observa. Embora apresentando escores de vocabulário baixos, observa-se, para quatro deles, um desempenho em leitura que lhes garantiu escores entre 0,55 e 0,75. Para dois sujeitos, a situação é inversa. Embora seus escores de vocabulário sejam entre os mais altos, verifica-se escores de leitura baixos.

Esses resultados não são evidência contrária a um nível limiar de vocabulário para a compreensão, mas apenas indicativos de que esse limiar pode não ser absoluto ou estático, mas sim relativo ou dinâmico. Aspectos como a dificuldade do texto, propósito da leitura, dentre outros, parecem interagir exigindo um nível mais alto ou mais baixo de competência lexical. A análise do perfil de cada um dos seis sujeitos que se comportam de maneira não previsível no gráfico que cruza escores de leitura com de conhecimento lexical mostra que todos os casos podem ser explicados, através de uma possível compensação desse conhecimento com outros tipos ou habilidades, tais como conhecimento conceitual, proficiência de leitura em LM e habilidade de inferência. Os resultados do estudo de processo poderão ser mais esclarecedores neste sentido.

Mais dois aspectos merecem ser salientados nesta discussão. Um deles refere-se à

adequação dos instrumentos de medida de vocabulário usados nesta investigação, enquanto o outro, aos resultados dos questionários que serviram para a elaboração do perfil do grupo.

Apesar de longa e um tanto cansativa para o avaliado, conforme depoimentos tanto de falantes nativos como de não nativos, a bateria de testes de vocabulário se mostrou bastante consistente, permitindo uma amostra não apenas quantitativa, ou seja, de número de palavras conhecidas, mas principalmente qualitativa, ou da profundidade desse conhecimento.

Um resultado bastante interessante da análise do vocabulário foi uma associação alta entre o teste de familiaridade e o teste de significado, assim como entre os subtestes que compõem o conhecimento de vocabulário. Em outras palavras, ficou evidenciado que os sujeitos que souberam colocar, associar, que souberam o significado, indicaram maior familiaridade com as palavras. Os sujeitos que não souberam frequência, classes, contextos, significado, indicaram menor familiaridade. A familiaridade foi, na realidade, o fator que teve a maior contribuição na AC, separando conhecimento de não conhecimento. Essa associação da familiaridade com os outros aspectos pode ser verificada não apenas no nível da análise do grupo como um todo mas também a nível de palavra.

Esses resultados, portanto, parecem confirmar a consistência ou a validade de uma lista de verificação (*checklist*) como uma medida aproximada de conhecimento de vocabulário, contrariando o que tem sido por vezes afirmado na literatura, ou que esse tipo de instrumento não pode ser considerado uma medida de conhecimento de vocabulário mas apenas de reconhecimento de palavras. A grande vantagem de seu uso está claramente em sua praticidade, trazendo uma simplificação bastante grande para a tarefa de avaliação. O fato de se associar com uma medida de vocabulário mais qualitativa ou que se baseia num conceito *rico* de vocabulário não significa, entretanto, que possa substituí-la totalmente, sem perda de informação. Em alguns casos, essa informação mais qualitativa pode ser de muita importância, principalmente quando o ensino está envolvido. Entretanto, em outros casos, como por exemplo como medidas aproximadas para seleção de sujeitos para pesquisas pode ser suficiente. A percentagem de erro na lista de verificação deste estudo, ou a identificação de palavras inventadas como reais ou uma superestimativa do vocabulário, é de aproximadamente 7% e, portanto praticamente irrelevante. A tendência em se superestimar o vocabulário tem sido identificada em estudos em geral que usam listas de verificação. Entretanto, pode ser minimizada desde que se use palavras inventadas e mecanismos que podem compensar esses erros ou "punir" o sujeito por essa sua tendência, como mostrei no Capítulo II.

Embora as listas de verificação usadas na literatura em geral sejam do tipo *sim/não*, isto é, com apenas dois níveis de conhecimento, aquela utilizada neste estudo apresenta três níveis, isto é, *familiar, mais ou menos familiar e não familiar*. Originalmente, o instrumento tinha um contínuo de quatro níveis, tendo suas categorias reduzidas a três, depois de uma análise preliminar. Essa análise havia revelado uma proximidade das categorias 1 e 2, respectivamente *nunca vi essa palavra* e *a palavra existe mas não sei seu significado* mostrando, portanto, que uma escala de familiaridade

de três níveis ao invés de quatro pode ser mais apropriada.

A análise dos dados que compõem o perfil do grupo mostra que os resultados das avaliações do pesquisador confirmam de forma precisa as auto-avaliações, mostrando que os leitores realmente têm uma idéia clara de sua compreensão, pelo menos a nível de idéia geral. Essa observação, confirmando outras evidências apresentadas na literatura, permitiria sugerir a auto-avaliação como uma medida adicional, para ser usada juntamente com outras medidas, na avaliação da compreensão em leitura.

Foi o objetivo deste capítulo apresentar a análise dos dados, assim como discutir os resultados da investigação que focalizou o vocabulário e a compreensão enquanto produto da leitura. Essa análise foi realizada em duas etapas: a primeira, em que se buscou caracterizar o grupo e cada sujeito em cada um dos componentes da leitura e, a segunda, com base nesses resultados, em que o foco da investigação foi a relação entre eles. Os resultados do próximo capítulo, focalizando o vocabulário e a compreensão enquanto processo da leitura, vem complementar estes resultados.

CAPÍTULO V

O VOCABULÁRIO E A COMPREENSÃO ENQUANTO PROCESSO DA LEITURA: ANÁLISE E DISCUSSÃO DOS DADOS DE ÊNFASE QUALITATIVA

No capítulo anterior, apresentei e discuti os resultados do estudo que focaliza o vocabulário e a compreensão enquanto produto da leitura. Neste capítulo, apresento e discuto os resultados da investigação que considera o papel do vocabulário na construção de significado em leitura, através, principalmente, da análise de protocolos de pausa, o seu foco estando, portanto, no processo de compreensão. Complementando o estudo anterior de produto, introduz-se uma dimensão qualitativa, vertical, focalizada no universo de três leitores de níveis diferentes de proficiência em leitura em LE, selecionados a partir da análise quantitativa.

O capítulo está organizado em três seções principais. A Seção 5.1 apresenta um perfil detalhado dos três sujeitos participantes e a 5.2, a análise de seus protocolos. O resultados da análise, trazendo dados do perfil dos sujeitos, quando relevante, são discutidos na terceira e última seção (5.3).

5.1 – Perfil dos sujeitos

Apresento, a seguir, o perfil detalhado de cada um dos três leitores, elaborado com dados coletados no estudo de produto. Esse perfil, subdividido em duas partes, é apresentado nas Tabelas 5.1 e 5.2 abaixo. O Perfil I é constituído, basicamente, por dados de identificação, *background* lingüístico, isto é, variáveis de perfil do estudo anterior, coletados a partir dos questionários. O Perfil II inclui os escores das variáveis ou componentes de leitura, obtidos através dos testes do estudo de produto. Nas tabelas, o leitor intermediário avançado é identificado por LIA, o leitor intermediário por LIO e o leitor iniciante por LIE.

O leitor intermediário avançado tem 18 anos, tendo ingressado na universidade em 1990. Quando os dados foram coletados, cursava Engenharia Elétrica e a disciplina Inglês Instrumental no nível II. Estudou inglês durante 5 anos, tendo freqüentado escolas de línguas. Na época da coleta, não fazia nenhum outro curso de inglês, além de Inglês Instrumental. Consegue ler em inglês e suas leituras, assim como em português, são variadas. Depois da leitura do texto em inglês, sua auto-avaliação da idéia principal foi *boa*, coincidindo com a avaliação posterior do pesquisador. Entretanto, curiosamente, não se considera um bom leitor em inglês nem em português. Para esse

leitor, o texto em inglês foi de dificuldade *regular*, tendo identificado o vocabulário como seu maior problema.

Tabela 5.1 – Perfil I dos sujeitos

	LIA	LIO	LIE
Idade	18	21	18
Ano de ingresso	1990	1988	1989
Curso	Engenharia Elétrica	Matemática Aplicada	Matemática Aplicada
Inglês Instrumental	Nível II	Nível I	Nível I
Anos estudando inglês	5	2	3
Fez curso de inglês em escola de língua	sim	sim	não
Faz curso de inglês em escola de língua	não	não	não
Lê em inglês	sim	sim	não
Lê em outras línguas	não	sim	não
Tipos de leitura em inglês	variadas	variadas	nenhuma
Tipos de leitura em português	variadas	variadas	variadas
Auto-avaliação/idéia geral	bom	regular	ruim
Avaliação pelo pesquisador	bom	regular	ruim
Auto-avaliação/nível de dificuldade do texto	regular	regular	difícil
Auto-avaliação das dificuldades do texto	vocabulário	vocabulário	vocabulário
Auto-avaliação/bom leitor em português	fraco	médio	fraco
Auto-avaliação/bom leitor em inglês	fraco	fraco	fraco

O leitor intermediário, por sua vez, tem 22 anos, tendo ingressado na universidade em 1988. Na ocasião em que os dados foram coletados, cursava Matemática Aplicada e Computacional e apenas a disciplina Inglês Instrumental no nível I. Estudou inglês durante 2 anos, tendo freqüentado escolas de línguas. Consegue ler em inglês e também em japonês. Suas leituras em inglês, assim como em português, são variadas. Depois da leitura do texto em inglês, sua auto-avaliação da idéia principal foi *regular*, coincidindo com a avaliação posterior do pesquisador. Não se considera um bom leitor em inglês mas se considera um leitor médio em português. Para esse leitor, o texto lido em inglês foi de dificuldade *regular*, tendo também identificado o vocabulário como sua maior dificuldade.

O leitor iniciante tem 18 anos, tendo ingressado na universidade em 1989. Cursava Matemática Aplicada e Computacional e a disciplina Inglês Instrumental no nível I, na época da coleta dos dados. Estudou inglês durante 3 anos, mas nunca freqüentou escolas de línguas ou qualquer outro curso de inglês, além de Inglês Instrumental. Não consegue ler em inglês mas suas leituras em português são variadas. Depois da leitura do texto em inglês, sua auto-avaliação da idéia principal foi *ruim*, coincidindo com a avaliação posterior do pesquisador. Não se considera um bom leitor em inglês nem em português. Para esse leitor, o texto foi difícil e o vocabulário, seu maior

problema, da mesma forma que para os dois leitores anteriores.

A Tabela 5.2 abaixo exhibe os dados de Perfil II dos três sujeitos. Para facilitar a leitura e interpretar esses dados, apresento a Tabela F.1 com as variáveis que fazem parte do Perfil II dos sujeitos, assim como com os escores mínimos/máximos do grupo, as médias e os desvios-padrão para cada uma das variáveis. Devo lembrar que os escores dessa tabela foram gerados através da análise de correspondência (AC).

Tabela 5.2 Perfil II dos sujeitos

Escore	LIA	LIO	LIE
1. FAM	0,74	-0,14	-0,64
2. FREQ	0,99	-0,17	-0,71
3. TCV	0,45	-0,23	-0,36
4. INFREAL	0,86	-0,37	-0,77
5. TCVIR	0,48	-0,25	-0,39
6. TCC	0,11	-0,25	1,60
7. META	4,7	3,2	0,5
8. CLOZE1A	3,6	2,9	2,8
9. CLOZE1B	8,3	5,1	5,4
10. FASEV1	6,5	5,8	4,5
11. LEITV2	6,3	2,5	2,6
12. CLOZE2A	3,9	0,8	0,3
13. CLOZE2B	6,7	1,4	0,3
14. LEIT1	6,5	5,0	0,5
15. LEIT2	7,5	6,0	0,5
16. RESUMO1	7,0	6,0	0
17. RESUMO2	7,5	6,0	0
18. DETA1	6,0	4,0	2,0
19. DETA2	8,0	5,0	2,0
20. LEITV3	8,6	1,7	0,5
21. CLOZE2C	8,9	7,2	0,3

Focalizo, em primeiro lugar, o leitor intermediário avançado e seus escores de proficiência e vocabulário em LE. Tanto no teste de conhecimento de vocabulário, como no de proficiência em LE ou Cloze2A e Cloze2B, os escores apresentados estão acima da média. Em inferência, o escore foi 0,86, um dos maiores, levando-se em conta o grupo como um todo. No teste de conhecimento conceitual, o escore foi 0,11, portanto, dentro da média. Com relação à concepção de leitura, o escore foi 4,7, que num contínuo de -1,0 (*local*) a 6,4 (*global*), significa uma tendência mais *global*.

Em LM, os escores apresentados tanto no teste Cloze1A e Cloze1B, assim como no Vestibular (Fasev1 e Leitiv2) estão acima da média. Comparando esses dados com os do Perfil I, principalmente no que se refere à auto-avaliação da leitura em português, verifica-se que o leitor é

bastante exigente e subestima seu desempenho.

Na leitura em LE, o leitor intermediário avançado apresentou um escore bastante alto no teste Cloze2C, ou seja, 8,9, considerando-se o desempenho do grupo como um todo. Nos testes de leitura I e leitura II, seu desempenho ficou na média, tendo alcançado um escore bastante alto (8,6) no Vestibular. A auto-avaliação em LE, da mesma forma que em LM (Perfil I), também mostra um leitor exigente com seu próprio desempenho.

Considerando-se todos esses escores, nota-se que o LIA é um leitor cujo desempenho, não apenas em leitura mas em outros componentes, é estável e sempre acima da média.

Ainda na Tabela 5.2 focalizo os escores de proficiência e vocabulário em LE do leitor intermediário. No teste de conhecimento de vocabulário, o escore foi 0,23, pelo menos metade daquele apresentado pelo leitor intermediário avançado. Entretanto, ainda pode ser considerado acima da média, levando-se em conta que o maior escore do grupo foi 0,73 e o intervalo de confiança foi da média foi -0,10 – 0,10. Em inferência, seu escore foi -0,37, um escore abaixo da média, considerando que o intervalo de confiança da média foi -0,13 – 0,11. Os escores em proficiência em LE ou no Cloze2A e Cloze2C são abaixo da média. No teste de conhecimento conceitual, seu escore foi -0,25, um escore acima da média, uma vez que o sinal negativo, neste caso, está invertido, significando, portanto, maior conhecimento. O conhecimento conceitual desse leitor é ligeiramente mais alto do que o do leitor intermediário avançado, que apresentou um escore de 0,11, ou na média. Com relação à concepção de leitura, seu escore foi 3,2, que num contínuo de -1,0 (*local*) a 6,4 (*global*), significa uma tendência levemente *global*.

Em LM, ou nos Cloze1A e Cloze1B, os escores do leitor intermediário estão abaixo da média nos dois testes. No vestibular, seus escores foram 5,8 na Fasev1 e 2,5 na Leitiv2, escores estes que, quando comparados com os do grupo, coloca esse leitor em uma posição na média no primeiro, e bem abaixo da média no segundo. O escore 2,5 apresentado na Leitiv2 é bastante baixo, principalmente quando comparado com o do leitor intermediário avançado (6,3). Isso significa que sua auto-avaliação, enquanto *leitor médio* em português, parece se confirmar.

Na leitura em LE apresentou um escore acima da média no Cloze2C (7,2). Nos testes de leitura I e leitura II seu desempenho ficou na média, como o leitor avançado, tendo alcançado um escore bastante baixo no vestibular, ou seja, 1,7. O que é interessante observar é que enquanto o leitor avançado apresenta um desempenho constante, o do leitor intermediário parece oscilar bastante, variando conforme a natureza e nível de dificuldade do teste.

Considerando os escores de proficiência e vocabulário em LE do leitor iniciante, observa-se que foram bem abaixo da média no teste de conhecimento de vocabulário (o maior escore do grupo foi 0,73) e também em proficiência ou Cloze2A e Cloze2B. Nesses últimos, os escores apresentados foram 0,3 nos dois, que são escores são bastante baixos, uma vez que os intervalos de confiança são 1,97 – 2,56 e 3,09 – 4,06. Em inferência, seu escore foi -0,77, um escore bem abaixo da média, se

for levado em conta que o intervalo de confiança da média está entre -0,13 – 0,11. No teste de conhecimento conceitual, o escore foi 1,60. Esse escore não é apenas o mais baixo dentre os três sujeitos, mas o mais baixo do grupo, uma vez que o sinal positivo, nesse caso, encontra-se invertido, significando, portanto, menor conhecimento. Com relação à concepção de leitura, o escore foi 0,5, que num contínuo de -1,0 (*local*) a 6,4 (*global*), significa uma tendência mais *local*, ou seja, a mais *local* dos três sujeitos.

Em LM, os escores do leitor iniciante foram 2,8 e 5,4, nos testes Cloze1A e 1B, respectivamente. Como se observa através da Tabela 5.G1, os intervalos de confiança foram 3,99 – 4,46 para o Cloze1A e 7,44 – 8,15 para o Cloze1B, e, portanto, os escores do leitor iniciante estão abaixo da média nos dois testes. Nesse caso, os escores se equiparam aos do leitor intermediário, que foram 2,9 e 5,4. No vestibular, os escores do LIE foram 4,5 na Fasev1 e 2,6 na Leitiv2, escores que quando comparados com os do grupo, coloca esse leitor em uma posição na média no primeiro e bem abaixo da média no segundo. O escore 2,6 na Leitiv2 é, como o do leitor intermediário, bastante baixo, principalmente quando comparado com o do leitor avançado (6,3). Os resultados dos testes de leitura em LM confirmam sua auto-avaliação como *leitor fraco*.

Na leitura em LE, diferentemente do leitor intermediário, apresentou um escore bastante baixo (0,3), quando o intervalo de confiança da média desses escores foi e 5,0 – 6,29. Nos testes de Leitura I e Leitura II, seu desempenho foi bastante fraco, apresentando o escore mínimo de grupo nos dois testes (0,5). No vestibular, seu escore (0,5) pode ser considerado baixíssimo, correspondendo também à nota mínima do grupo. Sua auto-avaliação como *leitor fraco* em LE parece, mais uma vez, se confirmar.

A análise dos dados até agora procurou fornecer um perfil completo de cada um dos três sujeitos participantes neste estudo. Esse perfil servirá de pano de fundo, tanto para a análise dos protocolos dos sujeitos, como para discussão de seus resultados, apresentados a seguir.

5.2 – Análise dos protocolos

A análise dos dados deste estudo objetiva responder à seguinte pergunta de pesquisa:

- Como se caracteriza a relação entre competência lexical e compreensão enquanto processo da leitura em LE? Ou seja, qual o papel do vocabulário na construção do sentido do texto?

Essa questão envolve outras mais específicas:

1. Tomando por base a categorização de recursos bilaterais do leitor proposta no Capítulo II (onde recurso é definido como a interação entre fontes/níveis de conhecimento (incluindo o prévio ou de mundo) e as restrições do texto, resultando no estabelecimento do contexto do texto e do contexto geral), quais são os recursos efetivamente utilizados pelos leitores no reconhecimento de palavras e construção de um sentido para o texto ?

2. Com base nos recursos utilizados, como se caracteriza o padrão de leitura apresentado por cada um dos sujeitos?
3. Qual a relação entre o padrão de leitura de cada sujeito, sua competência lexical e sua compreensão enquanto processo e produto?
4. Há evidências de um possível nível limiar de competência lexical para a inferência de palavras e para o processo de compreensão?

Como mostrei no Capítulo II, que traz a revisão da literatura, os trabalhos que discutem o papel do léxico na compreensão, seguindo a tendência de segmentação entre conhecimento lingüístico e cognitivo, polarizam-se entre a importância ou a contribuição de um conhecimento sólido e efetivo, que permite o reconhecimento automático de palavras e a inferência através do contexto. Nesses estudos, as duas questões são abordadas como se fossem independentes, isto é, como se a inferência de palavras, através do contexto, fosse possível sem a necessidade de competência lexical, e como se um vocabulário sólido eliminasse a totalmente a necessidade de contexto. Mostro, através desta análise, que a construção do sentido na leitura depende de ambos, embora, em cada caso, a contribuição do contexto seja diferente.

Neste estudo, o foco da análise são as pausas. Elas são consideradas rupturas na construção da compreensão ou momentos de desautomatização desse processo, causadas, em sua maioria, por um problema ou dificuldade. Devo salientar que os estudos que focalizam o processo, em sua maioria, limitam-se a focalizar as estratégias ou o aspecto cognitivo das soluções. Cavalcanti (1989:151-2), entretanto, observa que a identificação e análise dos problemas são fontes mais ricas de informação sobre o processo de leitura do que as estratégias, que dependem de como o problema é abordado. Nesse caso, *"... muitas vezes a heurística idiossincrática é ativada para agilizar o processo de interpretação"*. Além disso, embora a literatura em geral associe o uso de estratégias apenas a problemas ou situações de pausa, nota-se que também situações de fluência (vide Bialystok & Sharwood-Smith 1985) exigem seleção de palavras e essa seleção, embora automática, demanda habilidades ou capacidade de uso, além do conhecimento, e, conseqüentemente, também envolve estratégias.

Neste trabalho, portanto, através de uma análise das pausas, procuro identificar os problemas de três leitores de níveis diferentes de competência lexical e de leitura em LE em sua interação com o texto, assim como inferir os recursos utilizados ou que deixaram de ser utilizados, não apenas no reconhecimento de palavras, mas, principalmente, na construção de um sentido geral para o texto. Em outras palavras, busco investigar como esses leitores chegam até uma interpretação pragmática (para Cavalcanti 1983, 1989 e Moita Lopes 1986). Dessa forma, a análise não será baseada exclusivamente nas soluções efetivamente apresentadas pelos sujeitos para os problemas identificados nas pausas, isto é, soluções bem sucedidas, que permitiram o acesso a significados e a

compreensão, mas também nas mal-sucedidas, ou aquelas que poderiam ter sido usadas, ou recursos em potencial. Nesse sentido, embora as soluções sejam focalizadas, não será apenas considerado seu aspecto cognitivo ou as estratégias usadas, mas principalmente o conhecimento envolvido nessas soluções. Esse processo também envolve, portanto, a análise do problema, que nessa perspectiva, é visto como decorrente da falta de conhecimento e de capacidade de uso desse conhecimento.

Seguindo o arcabouço teórico proposto no Capítulo II, a construção do sentido envolve uma inter-relação de componentes, mais especificamente, a competência de leitura do sujeito, definida, neste trabalho, como uma competência lexical composta de vários níveis (sintático, morfológico, semântico, pragmático-discursivo, além de conhecimento de mundo) e as restrições do texto ou elementos do co-texto. Entretanto, essa interação ainda é mais ampla, uma vez que também considera os fatores pragmáticos e interacionais que constituem as restrições do contexto de leitura, envolvendo o propósito comunicativo do leitor, suas expectativas, motivações, crenças, valores, atitudes, concepções de leitura, por um lado, e as restrições do texto, sua organização, tipo de texto e função comunicativa, de outro. A combinação de todos esses elementos forneceriam ao leitor recursos, que lhe permitiriam estabelecer ou construir o contexto local e o contexto geral, utilizados tanto para a inferência de palavras como para a negociação de um sentido para o texto. Esses contextos de uso, portanto, que estão num nível pragmático-discursivo, vão constituir o foco desta investigação.

Minha hipótese inicial era que, nos trechos dos protocolos em que não ocorreram pausas, o reconhecimento das palavras e a construção do sentido havia ocorrido de forma automática. Entretanto, verifiquei que isso nem sempre pode ser observado. Uma comparação dos protocolos dos leitores com o texto-base revela outros problemas, não identificados por pausas. Pelo menos duas hipóteses poderiam ser levantadas para explicar não explicitação dessas questões, e que poderiam estar relacionadas a fatores pragmáticos ou da própria situação de leitura. A primeira é que realmente não se confirmaram como problemas, uma vez que um sentido foi construído. A segunda é que embora fossem problemas, os leitores, em face de tantos outros, resolvem ignorá-los. Essa estratégia, por sua vez, poderia ter motivações diferentes. Uma maneira consciente ou inconsciente de não se expor demais frente ao pesquisador, que também é seu professor, seria uma possibilidade. Outra seria o resultado de uma influência das molduras de interpretação ou guias (iniciais) (Cavalcanti op. cit. e Moita Lopes op. cit.) que, nesse caso, foram a *leitura de um artigo de assunto geral mas não desconhecido, leitura em inglês como LE, leitura para idéia principal, leitura para coleta de dados de pesquisa*.

A análise das pausas/outros problemas, visando a construção do sentido do texto, é complementada pela análise dos resumos que os sujeitos são solicitados a elaborar no final de cada parágrafo. A análise do produto da compreensão foi duplamente importante: não apenas como forma de validação dos protocolos, como mostrei no Capítulo III, mas principalmente por ter revelado que, embora os problemas sejam identificados nas pausas, é muitas vezes durante os resumos que os

leitores tentam resolver muitos deles. Isso significa que os dados desta investigação são resultado, não apenas da introspecção, mas também da retrospecção. A tendência de solucionar problemas no resumo pode ser vista como facilitadora, fornecendo ao leitor uma noção mais abrangente de todas as dificuldades do parágrafo.

A análise dos dados foi conduzida de acordo com os seguintes procedimentos:

- Transcrição dos protocolos dos três sujeitos (vide Apêndices H, Seções G.1, G.2 e G.3; vide também abreviações e símbolos nas transcrições). Esses dados incluem a Tarefa I (protocolos e resumos) e Tarefa II (grifar palavras desconhecidas);
- Análise do texto base (vide Apêndice E, Seção E.2) em sua estrutura problema-solução (vide Apêndice E, Seção E.3) que vai servir de base para a identificação de itens lexicais chaves (vide Cavalcanti 1989 para procedimentos de identificação desses itens).
- Identificação das pausas feitas durante a leitura e resumos no final de cada parágrafo e comparação das pausas/problemas dos três leitores.
- Análise dos protocolos resultantes da Tarefa I, para identificação dos recursos usados pelos leitores nas soluções/buscas de soluções dadas a palavras e à construção do sentido.
- Confirmação das inferências bem sucedidas através dos dados da Tarefa II.
- Confrontação dos itens lexicais chaves do texto-base com os itens contextualmente relevantes para os leitores.

Antes de iniciar a análise dos protocolos, apresento, no Quadro 5.1 a seguir, a análise do texto em sua estrutura problema-solução, base para a identificação dos itens lexicais chaves, listados no Quadro 5.2, apresentado mais adiante. Embora essa análise seja retomada apenas no final deste parágrafo, quando, então, será confrontada com os itens relevantes dos leitores, sua inclusão neste ponto do texto se justifica na medida em que serve de pano de fundo para os protocolos, mostrando como foram examinados os problemas de leitura e as soluções apresentados pelos sujeitos.

O texto apresenta, como fios condutores, três itens lexicais chaves, isto é, *capital cure*, *ecological ill*s e *Soviet Union*, que recorrem em várias partes do plano principal do discurso: no título, na situação/avaliação da situação e na solução. *Capital cure* reaparece na situação na oração *only a change in the way the economy is run can solve the crisis*, e também no resultado, como *economic reform* e *these reforms*. *Ecological ill*s aparece, mais uma vez, na situação, em forma de um sinônimo, ou seja, *ecological crisis*, além de recorrer também no plano secundário do discurso (*environmental problems*). *Soviet Union* recorre no problema. Esses itens chaves, através de cadeias coesivas, formam fios condutores semânticos que mantêm a unidade temática e a continuidade do sentido do texto. Enquanto *ecological ill*s (resultado da crise econômica) sinaliza o problema, *Soviet Union* o restringe e localiza, sendo que a solução proposta vai depender de uma mudança na economia, ou *capital cure*.

Quadro 5.1 – Estrutura problema/solução do texto-base

Situação/avaliação da situação

- A crise ecológica na União soviética está se tornando séria, a ponto de colocar em perigo a vida das pessoas.
- A crise ecológica é um resultado da crise econômica
- A proposta de solução é mudar os rumos da economia

Problema

- O monopólio do estado estimula o consumo exagerado de recursos naturais
- Há falta de mecanismos de mercado para um uso racional dos recursos naturais;
- Há falta de incentivo para a indústria se modernizar e introduzir tecnologia limpa e segura

Solução

- Cobrar das indústrias uma taxa por recursos naturais
- Privatizar como um meio de se introduzir competição
- Estimular novas empresas e *joint ventures*
- Tornar a moeda conversível, fazendo do país um membro do sistema monetário internacional
- Passar o controle dos recursos naturais para os governos locais

Resultado (plano russo)

- Fechamento de 100 das 500 indústrias
 - Entrada de dinheiro extra
 - Problemas em outras partes da União Soviética
-

Capital cure reaparece na situação em forma na oração *only a change in the way the economy is run can solve the crisis* e também no resultado como *economic reform e these reforms*. *Ecological ills* aparece mais uma vez na situação em forma de um sinônimo, ou seja, *ecological crisis*, além de recorrer também no plano secundário do discurso (*environmental problems*). *Soviet Union* recorre no problema. Esses itens-chaves, através de cadeias coesivas, formam fios condutores semânticos que mantêm a unidade temática e a continuidade do sentido do texto. Enquanto *ecological ills* (resultado da crise econômica) sinaliza o problema, *Soviet Union* o restringe e localiza, sendo que a solução proposta vai depender de uma mudança na economia, ou *capital cure*.

As cadeias coesivas, formadas a partir dos itens acima, são mostradas no Quadro 5.2.

A análise focaliza a busca da compreensão pelo leitor intermediário avançado. Esse processo, entretanto, é complementado com dados dos outros dois leitores, quando necessário. Esse entrelaçamento dos dados vai permitir uma análise comparativa e a caracterização do padrão de leitura dos três sujeitos. Os dados são apresentados de forma linear, reproduzindo a sequência de construção do sentido do leitor focalizado, através de suas idas e vindas ao texto. Propositamente detalhada, a análise mostra que as rupturas na compreensão, sinalizadas por pausas, assim como os recursos que os leitores mobilizam para solucioná-las, não são comportamentos acidentais e isolados, mas reincidentes, que refletem diferentes padrões de leitura e níveis distintos de competência lexical.

Antes de apresentar a análise, devo lembrar que o primeiro passo da Tarefa I consistia na leitura do título do texto, para a verificação das hipóteses do leitor a respeito do tópico e de medida do conhecimento prévio do leitor sobre o assunto.

Quadro 5.2 – Itens lexicais chaves e cadeias coesivas

- **Capital cure**

(economic) crisis, lack of market mechanisms, current economic system, state monopoly, (profligate) consumption, without reducing productivity, to cure, wastefulness, charge the going rate, ecological supplement to the price, surcharge, scarcity, old (economic) system, profitable, privatize, state factories, introduce competition, setting, up new enterprises, joint ventures, move away from heavy industry, step up production of consumer goods, join the international monetary system, rouble becomes convertible, industrial concerns, extra money, export, reduce, unemployment, providing funds, entrepreneurial activity.

- **Ecological ills**

life in danger, health is declining, polluted, lack of market mechanisms for the rational use of natural resources, scarcity of the resource, damage done to the environment, stagnated research to clean up industry, not profitable to introduce clean, safe technology, pollution, ecological standards and rules, 'polluter pays' policy, protection of the environment, raw materials for virtually nothing.

- **Soviet Union**

Soviet people, in the country, the country, Swansea, lake Ladoga, people of Leningrad, local governments, in their region, local environment, local governments, local people, local authorities, Russian republic, government of the republic, within its boundaries, Leningrad region, Russian authorities, coming into the republic, both abroad and to other parts of Soviet Union, Russian Federation, other parts of the countries, in the countries of Eastern Europe.

O título pode ser considerado bastante transparente com relação ao assunto discutido, como mostra o texto-base a seguir:

Texto base: Título

A capital cure for the Soviet Union's ecological ills

Para o LIA, a única pausa no título foi causada pelo desconhecimento da palavra *ills*, conforme o protocolo (1) abaixo:

Protocolo LIA:

Vou começá lê o título .. Bom aí o texto provavelmente vai .. dizê né talvez é .. vai apontá as curas ou soluções pra solucioná os .. *ills* talvez problemas ecológicos que a União soviética vem enfrentando *A capital cure for the Soviet Union's ecological ills* *Capital* talvez esteja relacionado aí com dinheiro economia ..

A partir desse protocolo, pode-se observar que o sujeito chega a um significado aproximado para essa palavra, ou seja, *problema*. Embora não explicita o recurso utilizado, pode-se inferir, pela construção do sentido que apresenta, que parece se utilizar de recursos sintático e semântico: a percepção da palavra enquanto núcleo de um sintagma nominal, que tem *ecological* como adjetivo, de um outro sintagma nominal (*capital cure*), assim como a relação semântica de causa-efeito entre eles. Portanto, a partir do título, o leitor já consegue identificar corretamente o tipo de texto, que envolve um problema e uma proposta de solução.

Uma comparação com o protocolo dos outros dois leitores mostra uma grande diferença de desempenho:

Protocolo LIO:

Eu já deparei com um problema aqui Eu não sei o que é /KəR/ *cure* e .. essa palavra aqui eu não sei o que é (...) [referindo-se a *ills*] Bem acho que vai falar sobre alguma coisa em relação à .. União Soviética mas eu não sei bem o que que é porque acho que a palavra .. /KəR/ *cure* ou *ills* .. tá? tem que saber o que que é Mas deve falar alguma coisa sobre .. é .. ecologia na União Soviética Certo ou alguma coisa nesse sentido Tem a ver com o verde ..

Protocolo LIE:

Vou vê se entendo esse .. o título primeiro .. vê o assunto .. se eu sei alguma coisa do assunto (... ..) *capital* .. *capital* /KuR/ *cure* para a União Soviética .. então *capital* para a União Soviética *ecological ills* .. *ills* .. *ills* .. não sei o que é nem /Kure/ *cure* .. também não sei .. *Soviet Union's capital* então *capital* a mais né para a União soviética dinheiro a mais pra ajudá a ecologia na União Soviética .. Esse subtítulo aqui .. *Stephanie Pains* ah! é a mulher que escreveu Então o que eu sei do título .. a ecologia (...) tá muito (...) a ecologia .. é muito .. dia a dia .. tal tem gente falando de ecologia .. então acho que deve ser alguma coisa que 'tão fazendo lá que 'tá .. acabando com alguma mata alguma coisa é .. uma poluição .. poluição fábrica .. deve sê .. alguma coisa .. poluição nas cidades nas florestas vamo vê o que que é então ..

No caso da palavra *ills*, os LIO e LIE apenas se limitam a identificar a palavra como desconhecida, sem tentativa de inferência. Nota-se, no protocolo do primeiro, a saliência ou importância dada a *cure* e *ills*, embora não consiga fazer uso dessa informação ou construir um sentido com ela, relacionando-as às informações de outros parágrafos. O LIE também infere *capital*, embora também não tenha estabelecido sua ligação com o resto do texto; ele tenta construir um significado para o trecho em questão, ignorando essa palavra chave, que aparece como núcleo do sintagma nominal formado pelo adjetivo *capital*. Dessa forma, o leitor acaba por salientar *capital*, fazendo dele o núcleo ou o item contextualmente relevante, e apagando a palavra *ills*. Uma análise das outras palavras desconhecidas do título mostra, também, que os dois leitores desconheciam e não inferiram a outra palavra desconhecida, ou seja, *cure*, extremamente importante enquanto pista para a inferência de *ills* e construção do sentido do título, como mostrei no protocolo do LIA acima. Um recurso que poderia ter sido usado para acessar *cure* seria lexical, baseado na forma da própria palavra, intermediado por um recurso interlingüístico, ou seja, baseado no conhecimento da LM. Através desses elementos, a palavra poderia ter sido percebida como cognata. Entretanto, suas pronúncias (/Kure/, /Kər/ e /KuR/) podem ter eclipsado sua grafia.

Portanto, *Soviet Union* e *ecologia* foram os itens contextualmente relevantes para os LIO e LIE, usados para a identificação da estrutura do texto, assim como seu foco ou assunto, além da

ativação de conhecimento de mundo. Não há necessidade de mencionar que esses elementos foram, por isso, prejudicados. Essa limitação parece ainda ter impedido a identificação das cadeias coesivas encabeçadas por esses itens chaves e percepção da função dessas cadeias na estruturação do texto (vide parte final da análise para uma discussão mais detalhada desta questão).

Retomando o texto-base, no primeiro parágrafo é explicitado o tópico, isto é, o problema ecológico na União Soviética, resultante da crise econômica, assim como a proposta de solução, que seria uma mudança nos rumos da economia, conforme demonstra o trecho a seguir:

Texto-base: parágrafo 1

The ecological crisis in the Soviet Union is reaching the point where "the very life of the Soviet people is in danger", and only a change in the way the economy is run can solve the crisis. According to Mikhail Krotov, an economist at the Leningrad Mechanical Institute, the ecological crisis is a direct result of the Soviet Union's economic crisis.

No protocolo do LIA não se observam pausas referentes a esse parágrafo. O leitor parece reconhecer de forma automática os significados das palavras, assim como outras pistas que ela traz, conseguindo estabelecer, através delas, as relações com as outras palavras. Esse processamento ocorre de forma automática, confirmando a hipótese que o leitor tinha levantado a partir do título, como fica evidente a partir do protocolo abaixo:

Protocolo LIA:

.. Bom vou começá lê o texto ((LB)) .. *The ecological /Kɹʌz/ crisis in the Soviet Union is reaching the point where the very life of the Soviet people is in danger .. and only a change in the way the economy is run can solve the /Kɹʌz/ crisis. According to /Maukal/ Mikhail /mikaɪ/ Mikhail Krotov an economist at the Leningrad Mechanical Institute the ecological crisis is a direct result of the Soviet Union's economic crisis* .. Bom, esse parágrafo basicamente fala que a crise ecológica na União Soviética 'tá atingindo pontos já .. é .. críticos né que coloca .. a a a vida das pessoas que lá moram em perigo e apenas é .. e somente uma uma mudança em .. nos rumos da economia é que pode pode assim salvá pode solucioná esse problema da crise ecológica de acordo com um economista de um de um Instituto de Leningrado Essa crise ecológica é um di .. é um resultado direto da crise econômica né existente na União Soviética.

O LIE não consegue criar redundância através da oração *only a change in the way the economy is run can solve the crisis*, que retoma o título. O LIO, aparentemente, não consegue estabelecer a relação entre *capital* e *capital* (português), que poderia ser um cognato "perfeito". O longo protocolo do LIE abaixo mostra suas idas e vindas ao texto, assim como sua incerteza e tentativa de adivinhação das palavras *run* (chuva?), *change* (troca?), dentre outras:

Protocolo LIE:

The /ɛKɒlɒɡɪkəl/ ecological /KRɪzɪs/ crisis in the Soviet Union é .. ecologia .. ecologia .. /KRɪzɪs/ crisis .. não sei não sei o que que é /KRɪzɪs/ crisis não sei .. na União Soviética .. is .. /SoM/ some .. um ponto .. /ʃeɪŋɡ/ change não sei /KRɪzɪs/ também não na União Soviética .. vou voltá né esse texto 'tá difícil *The /ɛKɒlɒɡɪkəl/ ecological* .. a ecologia na União Soviética .. a um ponto .. como que .. foi muito .. a vida do povo na União Soviética .. a um ponto .. vão falar de ecologia .. a União Soviética é um ponto que tem muito a ver com a vida .. do povo soviético certo? *In danger danger* .. deve sê o perigo deve sê o perigo o perigo o perigo é .. está em perigo O povo está em perigo o povo soviético está em perigo Então um ponto .. é muito .. o soviético povo está em perigo .. o povo soviético está em perigo e .. a troca não .. homem .. próprio não sei .. uma troca não .. 'pera aí economia correndo? /HɛN/ run /HɛN/ run chuva chuva *can* pode /SoLv/ solve solve não sei .. é isso resolver resolver a /KRɪzɪs/ crisis .. crisis é chave .. você tem que descobrir o que é crisis .. crisis é palavra chave .. aqui volta lá /ɛKɒlɒɡɪkəl/ ecological /KRɪzɪs/ crisis .. crisis ah crisis .. crisis de crise ahn crise .. crise política crise ecológica crise econômica Ah agora sim ecological crisis .. ecologia a crise na ecologia na União Soviética Ah se descobri o que é crise agora .. a um ponto .. onde .. ela .. eles .. muito .. o povo .. perigo Então o povo soviético em perigo .. e cada troca de seus .. economia chuva .. pode resolver a crise Então econômica o modo econômico pode diminuir a crise né pode resolver a crise .. Ah eu estou aqui crise ..

A palavra *crisis*, apesar de aparecer em mais de um contexto no mesmo parágrafo, e apesar de sua semelhança com o português, é inferida com muita dificuldade pelo LIE, que reconhece sua importância no parágrafo. Esse protocolo ilustra claramente que o problema do leitor não está apenas relacionado ao significado, mas é mais amplo, pois o desconhecimento da palavra também mostra desconhecimento de outros níveis, que impedem a construção do sentido no texto. Isso significa dizer que as lacunas causadas pela falta de palavras não são lacunas apenas de significados, mas lacunas de blocos de informação, uma vez que as palavras trazem consigo "instruções" de como se relacionam às demais. A falta desses blocos impede o leitor de construir um sentido para o trecho, que lhe serviria de contexto para a inferência de outras palavras e, dessa forma, o conhecimento acumulado do parágrafo torna-se completamente vago.

No segundo parágrafo, o texto retoma o anterior, trazendo informações adicionais sobre a crise ecológica salientada no título, assim como de seus efeitos na saúde das pessoas, incluindo um exemplo da cidade de Leningrado, poluída por laticínios, fábricas de polpa e de papel e indústrias químicas. Nesse parágrafo, embora o LIA interrompa sua leitura e retome o trecho onde está inserida a palavra *drawn*, ela não parece ter sido um problema, uma vez que seu significado foi explicitado rapidamente. O texto-base e o protocolo do LIA, a seguir, ilustram esse aspecto:

Texto-base: parágrafo 2

Environmental problems in the country have "reached the stage where people's health is declining", Krotov said in an unscheduled talk in Swansea. The people

of Leningrad, for example, drink water drawn from Lake Ladoga, which is polluted by the effluent from nine pulp and paper plants, two dairy complexes and a series of chemical factories on its shores.

Protocolo LIA:

((LB)) *Environmental problems in the country have reached the stage where people's health is declining* Krotov said in an unscheduled talk in Swansea *The people of Leningrad for example drink water drawn from Lake Ladoga which is polluted by the effluent from nine pulp and paper plants two dairy complexes and a series of chemical factories on its shores .. Hum .. drink water drawn .. from Lake Ladoga .. polluted by the effluent* *afluente /PuLP/ pulp .. nine pulp and paper plants .. não sei o que é /PuLP/ pulp plants fábricas indústrias indústrias de papel? .. dairy complexes laticínios deve ser alguma coisa relacionada a isso laticínios and a series of chemical factories on its shores Shores shores baía costa costa costa Bom o segundo parágrafo .. problemas com relação ao meio ambiente na União Soviética atingiram um estágio .. crítico né como eu havia falado no primeiro parágrafo A saúde das pessoas 'tá em declínio Um exemplo é citado aí: as pessoas né de Leningrado bebem água né que .. é poluída de .. bebem água drawn from the lake .. bebem a água do lago .. que é poluída pelo seu afluente (...) é é poluída por um afluente quer dizer por um rio que deságua nesse lago que é poluído por várias indústrias de papel laticínios uma série de fábricas ou indústrias químicas que estão instaladas na costa né desse rio que aflui no lago.*

O que parece ter ocorrido, nesse caso, foi a construção de um contexto, mais especificamente, através da relação de origem ou procedência sinalizada pela preposição *from*, juntamente com a transparência da palavra *lake*, tornando praticamente desnecessária a palavra *drawn*. Esse recurso sintático, aparentemente, parece ter sido usado pelo LIA, em um processamento semi-controlado, para ativar ou confirmar um significado de *drawn* já existente. O fato de a palavra não ter sido salientada pelo LIA como desconhecida na Tarefa II (grifar palavras) poderia ser visto como mais uma evidência de que ela já era conhecida.

O mesmo tipo de recurso pode ter sido usado pelo LIO e LIE na inferência da palavra *drawn* e na construção de um sentido para o trecho *water drawn from de lake*, como mostram trechos de seu protocolo abaixo (vide protocolo completo no Apêndice H.3). Entretanto, para o LIE, por exemplo, esse processo, caracterizado por idas e vindas no texto, não foi tão simples como para os outros dois leitores, uma vez que esse recurso somente foi usado depois de tentar acessar o significado através da ativação de conhecimento gramatical ou sintático das partes do verbo, em um exercício, bastante freqüente no protocolo desse leitor, totalmente mecânico (vide também outros exemplos com os verbos *become*, *get*, *take*, nos parágrafos 6 e 8). Esse procedimento, entretanto, mostra que a palavra já tinha sido encontrada antes. Pode-se observar, que não apenas o conhecimento sintático é importante para o leitor, mas especialmente o som das três partes do verbo (presente, passado, particípio passado), numa tentativa de recuperação, possivelmente, da situação

em que foi exposto a palavra.

Protocolo LIE:

.. povo de Leningrado por exemplo bebe água /DRauN/ *drawn* (...) /DrauN/ *drawn* (...) não sei deixa eu ver .. problemas no meio ambiente no país .. os problemas no meio ambiente no país tem reagido /Hɪʃɛd/ *reached* /Hɪʃɛd/ *reached* tem reagido um estágio onde pessoas quentes estão declinando ..

O povo de Leningrado por exemplo bebe água .. /DRɪN/ *drawn* para Lake Ladoga .. bebe água .. /DRɛuN/ *drawn* /DRɪN/ *drawn* dirigida né que vem acho que vem do Lago .. Ladoga Ladoga deve ser o nome do lago 'tá maiúscula com sua poluição ah .. tá! .. poluição do rio aqui .. então não vai ter água Com sua poluição .. pelo afluente .. pelo seu afluente ..

Então ele dá um exemplo Por exemplo o povo de Leningrado bebe água /DRɛuN/ *drawn* rio dirigida de um rio que vem dum rio não de uma lagoa aliás de uma lagoa Ladoga ...

Protocolo LIO:

Deparei com a mesma .. no segundo parágrafo deparei com o mesmo problema .. certo? a palavra /Hɪʃɛd/ *reach* eu não sei que que é (...) Bom, já deparei também com outras palavras /ʌnʃɛdɪd/ *unscheduled* e /Sænsə/ *Swansea* .. certo? São duas palavras também que eu não sei (...) e /Pʊlp/ *pulp* também é .. outra palavra que eu não sei .. e /ʃɔres/ *shores* também não sei o que que é certo? Bom o segundo parágrafo .. fala sobre os problemas do meio ambiente no país certo? .. É' .. parece-me aqui que ele fala que os problemas econômicos no país .. chegam num estágio onde a população é .. certo? tá .. alguma coisa os problemas os problemas os problemas .. sobre o meio ambiente estão aumentando e .. alguma coisa nesse sentido né? .. Certo? que os problemas do meio ambiente estão aumentando e alguma coisa relacionada às pessoas vem declinando O que eu não sei o que que é devido a palavra /Hɪʃɛd/ *reached* e /Hɛlθ/ *health* tá? .. Bom, segundo (...) Depois ele diz aqui *Krotov said in an unsheduled talk in Swansea* não entendi o que que está escrito aqui devido a palavra *unsheduled* e *Swansea* certo? Depois ele relata que as pessoas de Leningrado por exemplo é .. bebem água de um do lago Ladoga tá? é .. local é poluído por um por .. o local é poluído .. por .. alguma coisa .. tá? Deve ser deve ser fábrica de papel É isso mesmo tá? Então ele .. é o seguinte No lago Ladoga são jogados poluentes certo? vindos tá de nove .. /Pʊlp/ *pulp* eu não sei o que é de nove fábricas de papel que .. no caso essas são as fábricas de papel e as fábricas de algum .. alguma outra coisa que é o /Pʊlp/ *pulp* no caso certo? Essas esses dois complexos tá? de fábricas jogam diariamente uma série de .. fatores químicos certo? e o finalzinho do parágrafo que eu não sei o que que é [referindo-se a *shores*] (...)

Retomando o protocolo do LIA, nota-se que seu processo de construção do sentido é mais uma vez interrompido por uma pausa, detonada pela palavra *pulp*. Apesar da associação semântica entre *pulp* e *paper*, não consegue usar esse recurso para estabelecer uma relação entre as palavras. Também não relaciona *pulp* à *polpa* em português, talvez por ter sido pronunciada incorretamente (não só por esse leitor, mas pelos três) ou seja, /Pʊlp/ e /PʊLP/. A pronúncia correta do primeiro

som vocálico (/e/ ao invés de /u/) poderia ter mostrado a semelhança com a palavra em português, ou ter sido utilizada como um recurso fonológico interlingüístico. Entretanto, o fato de o leitor desconhecer a palavra, não causa rupturas em sua compreensão, pois ele percebe que não se trata de uma palavra essencial ao sentido. Dessa forma, ela é apagada no resumo. O LIA parece perceber, entretanto, que *pulp* está sendo usada como adjetivo para *plant*, isto é, *fábrica*, e que, portanto, *pulp and paper plants* são *fábricas de papel e de alguma outra coisa*. A decisão de apagamento, portanto, baseou-se em um recurso sintático, relacionado à classe de palavras e à sua posição na oração. Essa palavra, devo salientar, também não foi inferida pelos outros dois leitores.

Em alguns casos, como por exemplo, no da palavra *reach* no protocolo acima, observa-se que o LIO, embora desconhecesse a palavra, conseguiu construir um sentido aparentemente por ter percebido a relação de causa-efeito que ela estabelece, embora sem explicitar que tenha sido esse o recurso usado. Este poderia ser um exemplo de *contexto potencialmente rico* ou *potencialmente redundante*, ou um contexto que oferece muitas pistas ou possibilidades de acesso ao significado. Entretanto, o leitor não se dá conta que inferiu a palavra, continuando a citá-la como desconhecida até o final do texto.

A segunda pausa do LIA no parágrafo 2 é detonada pela palavra *effluent*. Nesse caso, para a inferência da palavra, o leitor parece ter-se utilizado também de um recurso lexical, mas desta vez, baseado na forma da palavra juntamente com o conhecimento lexical da LM. Sua semelhança com o português *afluente*, assim como o contexto imediato à palavra, mais especificamente a percepção de sua função gramatical e relação sintática com a palavra *plant* no sintagma *pulp plant*, levam o sujeito a um sentido que, embora incorreto, não causa rupturas à compreensão. Foi possível para esse leitor criar uma situação que faz sentido: *um afluente que deságua no lago ou um rio que aflui no lago*. O contexto, criado dessa forma, foi distorcido para acomodar o significado da palavra. A construção de um sentido, entretanto, pode tê-lo impedido de verificar outros níveis de conhecimento, como por exemplo, conhecimento de mundo, que eliminaria essa hipótese, uma vez que *lago não tem afluente*; ou ainda análise da própria palavra, que poderia identificar o prefixo *e*, indicativo de procedência. A palavra *effluent* também foi inferida como *afluente* pelo LIE; o LIO a ignorou.

A próxima pausa observada no protocolo do LIA é causada pela palavra *shores*, que é repetida em voz alta algumas vezes. O som da palavra parece ativar, para esse leitor, de alguma forma, seu significado. Observa-se, nesse caso, novamente a utilização de um recurso fonológico, baseado no som do item detonador, da mesma forma que ocorreu com o LIE acima. O primeiro significado ativado, entretanto, é *baía*, logo substituído pelo mais apropriado, *costa*. O significado de *shores* é incorporado facilmente ao sentido em construção pelo LIA. O mesmo tipo de procedimento também parece ter sido usado pelo LIE na inferência da mesma palavra, assim como também em inúmeras outras, mas nenhuma vez pelo LIO. Entretanto, enquanto a tentativa do LIA

foi bem sucedida, a do LIE foi mal sucedida, uma vez que, nesse caso, o significado era totalmente desconhecido. O mais próximo que conseguiu chegar foi *lugar*. Essa ênfase no som, neste caso, em detrimento da forma, também observada, por exemplo, no parágrafo 8 com a palavra *boundaries*, poderia sugerir um papel facilitador para a subvocalização em LE, contrariamente do que se observa na leitura em LM. Para que esse recurso possa ser utilizado, entretanto, de maneira eficiente, na identificação de cognatos ou de semelhanças com a LM parece ser necessária a pronúncia correta.

Retornando ao protocolo do LIO, observa-se que esse leitor limitou-se a identificar a palavra *shores* como desconhecida, abandonando a construção do sentido no trecho que incluía a palavra. O LIE, pelo menos consegue identificar que *shores* refere-se a um lugar, talvez pela utilização de um recurso sintático, semântico ou de relação *de lugar* marcada pela preposição *on*, ilustrado pelo trecho de seu protocolo a seguir:

Protocolo LIE:

.. e uma série de químicos ah .. pode ser alguma coisa poluente .. *dairy*
complexos poluentes uma série de chemical fatores .. no setor nos /jərs/ *shores*
lá no lugar /jər/ *shore* .. /jər/ *shore* eu não sei /jər/ *shore* não sei /jərs/ *shores* eu
não sei.

Apesar de *drawn*, *shores* e *affluent* terem sido palavras desconhecidas comuns aos três sujeitos, observa-se que o desempenho do LIA na construção do sentido desse parágrafo difere bastante dos outros dois leitores (vide protocolo completo do LIE no Apêndice H.3). Esse desempenho diferenciado poderia ser atribuído a dois fatores, relacionados entre si: densidade de palavras desconhecidas e falta de conhecimento acumulado. Um número grande de outras palavras desconhecidas, além daquelas acima salientadas, tornam o parágrafo denso para os LIO e LIE, impondo, portanto, restrições à construção de um contexto e percepção das pistas. Embora algumas dessas palavras tenham sido inferidas pelo LIO, foram inferidas incorretamente. Nesse sentido, não se pode esquecer que o contexto é dinamicamente reformulado a partir do inferência ou não inferência de uma palavra. Aquelas que constituem o *foco* ou *evento focal* (Goodwin & Duranti 1991) são, depois de inferidas, *plano de fundo* para outros eventos focais. Juntamente com a falta de contexto local, causada pelo número grande de palavras desconhecidas e, portanto, por uma competência lexical insuficiente, notam-se também problemas no conhecimento acumulado do texto até a ocorrência da palavra, uma vez que para esses leitores, o título, assim como o parágrafo 1, não haviam ficado claros. Em outras palavras, a influência da criação de sentidos distorcidos é muito mais ampla, uma vez que não se restringe apenas a nível local. Os resumos que o LIA faz no final de cada parágrafo, assim como a explicitação da ligação entre os dois parágrafos mostram que seu conhecimento acumulado até esse ponto do texto é bastante bom, contrapondo-se ao dos outros

leitores, que além, de não terem construído um sentido para o parágrafo anterior, não conseguiram estabelecer a relação entre os parágrafos, dificultada principalmente pelo desconhecimento da palavra *health*.

A dificuldade de construção do sentido pelo número grande de palavras desconhecidas é observada muito freqüentemente nos protocolos dos leitores LIO e LIE. Nesse sentido, observa-se, principalmente no protocolo do LIO, uma tendência a identificar apenas as palavras-problemas durante a leitura, sem tentativa de inferência (24 vezes), criando lacunas que vão aparecer no nível de significado e também em outros níveis. Embora o sujeito afirme desconhecer o significado da palavra, não apresenta nenhuma hipótese a respeito. Em nenhum momento, o LIO relê o texto ou tenta reprocessá-lo, construir um sentido sem essas palavras ou a partir das poucas que consegue inferir, ou ainda com o conhecimento acumulado, mesmo incompleto, do parágrafo. No resumo, a tentativa de construção de um sentido poderia ser considerada uma "tradução"(3). Esse procedimento é bastante ilustrativo de um comportamento que é típico desse leitor, mas que também poderia estar relacionado aos propósitos da tarefa de leitura. O leitor estaria, nesse caso, apenas cumprindo uma tarefa. O LIE, apesar também de apresentar grandes problemas em sua competência lexical, salientando como desconhecidas muitas palavras, tem um comportamento distinto, uma vez que tenta reprocessar esses trechos densos.

Retornando ao texto-base, observa-se que o parágrafo 3 complementa e amplia o parágrafo 2, na medida em que explicita como a crise econômica está afetando a ecologia; que a culpa não é das indústrias mas sim, especificamente, da falta de mecanismos de mercado para um uso racional de recursos naturais.

Nesse parágrafo foram identificadas três pausas no protocolo do LIA: *blame*, *narrow-minded* e *lack*. O texto-base, assim como o protocolo do LIA, são apresentados a seguir:

Texto: parágrafo 3

Krotov insisted that industry is not to blame. "Narrow-minded people blame (the crisis) on technological progress itself, and this has led to many industrial activities being banned for ecological reasons." The real problem, he explained, is the "lack of market mechanisms for the rational use of natural resources".

Protocolo LIA:

Krotov insisted that industry is not to blame (...) insisted that industry is not to blame .. Narrow-minded people blame (the crisis) on technological progress itself .. ah .. tá quer dizer ele insistiu que não se deve culpá a indústria .. narrow-minded narrow-minded people .. blame pessoas com .. cabeça pequena .. culpam (...) on technological progress itself and this has led to many industrial activities being banned for ecological reasons. The real problem he explained is the lack of market mechanisms for the rational use of natural resources .. O problema real .. lack .. is the lack .. of market mechanisms .. Não entendi direito

essa frase .. Vou relê .. ((LB)) *Krotov insisted that industry is not to blame*
Narrow-minded people blame (the crisis) on technological progress itself and
this has led to many industrial activities being banned for ecological reasons
 Bem até aí .. *The real problem he explained is the .. lack* já ouvi essa
 palavra antes .. *lack of market* .. falhas? .. nos mecanismos de mercado .. para
 o uso racional dos recursos naturais .. Bom esse terceiro parágrafo disse que ..
 não se deve culpá as indústrias né Pessoas de pensamento curto culpam né o
 progresso tecnológico né pela crise ecológica e isso levou já a .. ao banimento de
 muitas atividades industriais por razões ecológicas. Ele fala que o problema real
 é a falha falhas ou a falha nos mecanismos de mercado .. que o uso racional ..
 para o uso racional dos recursos naturais.

Depois de identificar o problema (*blame*), o LIA faz uma tentativa de inferência da palavra através da releitura do trecho onde ela está inserida, aparentemente, com o objetivo de construir um contexto baseado nas restrições do contexto imediato à palavra ou da oração, e, mais especificamente, na relação semântica e sintática da palavra *blame* com *industry*. Entretanto, o que parece ter sido realmente decisivo, para a inferência bem sucedida, foi o fato de a palavra ter aparecido mais uma vez na oração seguinte, em um outro contexto mais simples e portanto, mais transparente, que lhe possibilitou confirmar uma hipótese levantada a partir do primeiro contexto. Depois de inferir *culpar*, o leitor incorpora esse significado ao sentido em construção, que é retomado, logo em seguida, para inferir *narrow-minded*, que detona a próxima pausa. Esse comportamento é, portanto, distinto do comportamento do LIO que salientei acima, uma vez que o contexto é constantemente reformulado; nesse caso, ele incorpora o significado *culpar* antes de iniciar a próxima inferência.

Não parece claro, através do protocolo do LIA, se a palavra *narrow-minded* foi inferida através do conhecimento das duas partes dessa palavra composta, ou se esse significado já era conhecido e foi apenas confirmado ou ativado através de um recurso lexical, baseado no próprio item detonador. As evidências parecem sugerir a segunda opção, uma vez que o significado atribuído a palavra pelo leitor, isto é, *cabeça pequena* ou *pensamento curto*, embora corretos sob o ponto de vista do significado, são traduções diretas das palavras *narrow* e *mind(ed)*. A expressão *mente*, de certa forma até mais semelhante com a palavra *mind*, teria sido mais adequada a esse contexto. A reformulação para *mente estreita* ou algo semelhante, contudo, não ocorre, o que não impede a construção de um sentido. Os dados da Tarefa II não fornecem nenhuma informação adicional. A percepção da relação sintática entre *narrow-minded*, um adjetivo no sintagma nominal, assim como a relação desse sintagma com *blame*, também mostraram-se importantes.

O LIO, embora tenha inferido *blame*, entretanto, não processa o trecho que contém essa palavra. O LIE, por sua vez, não infere a palavra, embora reconheça sua importância no parágrafo. Em um processo de tentativa e erro, que chamaria de *pura adivinhação* (*wild guessing*), tenta adivinhar um significado para ela, reprocessando todo o trecho para acomodar esse significado.

Entretanto, suas tentativas são mal-sucedidas, uma vez que o trecho é muito denso para esse leitor, contendo muitas palavras desconhecidas (vide Apêndice H2 e H.2, para esses protocolos).

Uma outra pausa, observada no protocolo do LIA, refere-se à palavra *lack*. O leitor consegue, nesse caso, um significado bastante satisfatório (*falha*), que lhe permitiu construir o significado do trecho em questão. Mais uma vez, procurou reler várias vezes não apenas a porção que continha a palavra desconhecida, mas também a anterior a ela, ou seja, todo o parágrafo, repetindo a palavra várias vezes em voz alta, como no caso de *shores* acima. Portanto, o que parecem ter servido como recursos para o estabelecimento de contexto foram o conhecimento acumulado do parágrafo, assim como a pronúncia da palavra em voz alta. O procedimento de leitura e releitura parece característico do LIA, se repetindo inúmeras vezes no seu protocolo. Ele parece depender de conhecimento não apenas local, que envolve o contexto imediato à palavra, mas que também de processamento de porções maiores do texto. A ligação entre as duas partes do parágrafo, isto é, entre *a indústria não deve ser culpada e o problema real parece ter sido fundamental para a inferência bem sucedida*. O protocolo do LIA na Tarefa II mostra, entretanto, que a palavra não era totalmente desconhecida, de forma que o conhecimento acumulado, juntamente com a pronúncia da palavra em voz alta, podem ter sido mobilizados, mais uma vez, apenas para ativar um significado já conhecido, embora não completamente estabelecido.

Protocolo LIA (Tarefa II):

.. *lack is the lack* Ai eu pensei talvez como falha deficiência né mais ou menos assim acho que deu algum sentido pelo contexto Eu já tinha tido contato com essa palavra antes acho que em algum texto que eu já tinha lido nas aulas de .. instrumental Mas não lembro ..

Os protocolos dos LIO e LIE mostram que não conseguiram inferir *lack* e que sem a palavra, que marca um aspecto chave do texto, ou seja, a causa da crise ecológica, foi difícil construir um significado para o trecho, como mostram trechos de seus protocolos a seguir:

Protocolo LIO:

.. o Krotov explica que o problema real certo? é que o /læk/ *lack* eu não sei o que é /læk/ *lack* certo? o /læk/ *lack* não sei os mecanismos de mercado fariam um uso racional de .. é.. fontes naturais certo? Deve ser alguma coisa relacionado tá? que ou seja a indústria ou a atividade industrial aumente certo? mas que ela use o .. as fontes naturais racionalmente não poluindo alguma coisa ..

Protocolo LIE:

o lago de .. é o grande .. /læk/ *lack* /læk/ *lack* não é lago não /læk/ *lack* de mercado .. mecanismo meca é o grande mercado mecanismo para a reação para usar o recurso natural .. Então o real problema ele explanou explanou .. ele quem

o Krotov lá Krotov explanou .. é o .. é seu .. o .. por causa .. é que .. tão poluindo .. os os mercados os mecanismos de mercados para .. para a racionalização o uso dos recursos naturais .. e .. então ele falou que o problema é que .. tem que começar racionar .. os recursos naturais certo? então começar racionar os recursos naturais ..

É necessário lembrar que esses leitores não haviam conseguido processar o primeiro trecho do parágrafo, e a ligação entre as duas partes era importante. O protocolo do LIE, nesse trecho, ilustra muito bem o padrão de comportamento que apresenta durante a leitura. Embora consiga inferir palavras claramente cognatas tais como *ecological*, *mechanisms*, dentre muitas, faz confusões com outras, tais como *reasons* (inferida como *reações*), que também poderiam ser consideradas cognatas. O texto se mostra muito difícil para esse leitor, e provavelmente, se estivesse em outras circunstâncias, já teria desistido da tarefa. Em suas tentativas de inferência, muitas das primeiras hipóteses são, em geral, baseadas na forma ou som da palavra, como por exemplo *lack* = *lago*, ou em um recurso lexical, juntamente com conhecimento da LM ou da LE. Essas hipóteses, entretanto, não são confirmadas ou verificadas através de outros níveis de conhecimento (vide as tentativas do LIE de inferir a palavra *health* (quente?), aparentemente, através de uma semelhança com *heat*)

O parágrafo 4 oferece informação adicional, relacionada com o sistema de monopólio dos recursos naturais em vigor na União Soviética, justificando a afirmação do parágrafo anterior, isto é, porque a falta de mecanismos de mercado levou a uma crise ecológica.

São duas as pausas observadas no protocolo do LIA referentes a esse parágrafo, mais uma vez detonadas pelas palavras *timber* e *profligate*, conforme texto-base e protocolo abaixo:

Texto base: parágrafo 4

Under the current economic system, where the state has a monopoly on all natural resources, timber, oil and gas and so on are virtually free. This has encouraged profligate consumption by industry. Krotov estimates that for every unit of national income, the Soviet Union uses two to three times as many resources as other countries. It should be possible to bring consumption into line with other countries, without reducing productivity.

Protocolo LIA:

Under the current economic system .. where the state has a monopoly on all natural resources .. timber oil and gas and so on a.. are virtually free This has encouraged .. hum .. profligate consumo .. excessivo? (...) Krotov estimates vamo deixá lá pra frente isso vamo vê se depois eu consigo voltar isso aí e vê .. qual o significado Krotov estimates that for every unit of national income .. the Soviet Union uses two to three times as many resources as other countries Ah ah .. eles desperdiçam recursos naturais usam demais Tá certo It should be possible to bring consumption into line with other countries without reducing productivity Tá certo Basicamente isso aqui esse parágrafo fala que na União Soviética o consumo é exagerado né .. o consumo dos recursos naturais né óleo

gás natural .. tudo o mais E .. esse consumo exagerado ao invés de gerá uma produtividade maior não é desperdiçado Então ... aqui ele fala que segundo o texto né seria possível é .. diminuí o consumo das indústrias soviéticas sem afetá a produtividade das mesmas E diminuí o consumo .. é .. atingindo então um .. um consumo parecido com o das indústrias de outros países né .. OK

Para a inferência de *timber*, o LIA parece utilizar-se de um recurso baseado nas relações da palavra com outras de seu contexto imediato, ou da oração onde a palavra está inserida. A palavra *timber* é acompanhada por *oil* e *gas*, sendo que *natural resources* aparece como seu termo superordenado. A inferência de um significado aproximado, nesse caso, *recurso natural*, perfeitamente satisfatória, parece ter sido feita através dessa associação semântica. Um significado mais preciso da palavra nesse caso teria sido impossível a partir desse contexto, que não oferece outros elementos de acesso, podendo, portanto, ser caracterizado como um *contexto pobre*. O significado *recurso natural* é explicitado pelo LIA, entretanto, somente na Tarefa II, ou seja, na tarefa de grifar palavras. Durante a leitura e o resumo, o leitor simplesmente ignora a palavra, reconhecendo seu significado como essencial ao sentido do texto.

Protocolo LIA (Tarefa II):

Timber .. *timber* é algum tipo de recurso natural eu não sei bem qual é também.

Os protocolos dos LIO e LIE abaixo mostram que esses leitores também chegam a sentidos aproximados para *timber*, suficientes para não causar rupturas na construção do sentido:

Protocolo LIO:

Bom deparei com outra palavra aqui é *timber* (...) Tem então *profligate* outra palavra que eu não sei (... ..) Eu não sei o que quer dizer *bring* também esqueci Eu conheço essa palavra mas eu esqueci o significado (... ..) Bom ele o texto aqui eu entendi que *timber* deve ser algum .. algum combustível fóssil certo? .. Bom o quarto parágrafo diz o seguinte Bom dentro do sistema econômico corrente onde o estado tem o monopólio de todas as reservas naturais *timber* óleo gás então são .. então o sistema econômico é virtualmente livre certo? Isto foi encorajado tem encorajado tá .. a .. concepção *profligate* que eu não sei o que que é pela indústria certo? Então Krotov ele estima que é .. para cada unidade .. nacional certo? *Income* .. outra palavra que eu esqueci de relatar que é *income* que eu não sei o que que é *In* é dentro come é vir certo? *Income* vir para dentro certo? Deve ser indústria dentro do país alguma coisa assim certo? Então a União Soviética usa de duas a três vezes as .. as .. muitas fontes tá .. e .. as mesmas .. as .. as muitas fontes tá? em outros países também certo? Bom ele diz que isso pode ser possível tá? *to bring* eu não sei o que que é .. a concepção dentro de uma linha com outros países é .. na qual podem vai reduzindo a produtividade alguma coisa assim (... ..)

Protocolo LIE:

..o sistema econômico lá vigente né na Rússia o sistema econômico deles lá onde o estado é tem o monopólio nos recursos naturais .. *timber* também não não é nos recursos naturais .. *timber* *timber* apareceu noutra texto e eu não sei que que é isso *timber* .. e.. óleo *oil* e gás e óleo e *timber* deve ser *tinier* alguma coisa química *tinier* gás e também .. hum .. são hum .. virtualmente livres.

Para o LIO, *timber* pode ser algum combustível fóssil, uma vez que se baseou em *oil* e *gas*, e não em *natural resources*. Não conseguiu estabelecer a ligação semântica entre *timber* e *natural resources*, ou identificar *natural resources* como o termo superordenado. A informação *are virtually free*, que é o predicativo de *timber*, *oil and gas*, aparentemente, também não foi usada pelo LIO, uma vez que ao processar esse trecho, identificou incorretamente o sujeito: o sistema econômico, ao invés de *timber*, *oil and gas*. O fato de o verbo estar no plural (*are*) não parece ter incomodado o leitor. Dessa forma, poder-se-ia dizer que deixou de verificar o nível de estrutura da oração ou das restrições sintáticas na oração, da mesma forma que deixou de usar conhecimento de mundo, especificamente no que consiste um monopólio por não possuir esse conhecimento ou por não ter conseguido ativá-lo para a verificação da informação.

Para o LIE, por sua vez, *timber* é entendido como *tinier*, alguma coisa química. Embora a relação entre *timber* e *gas* e *oil* possa ter sido o recurso utilizado para a criação de contexto, a inferência, nesse caso, parece ter sido baseada na percepção da forma ou som da palavra, ou um recurso fonológico. Os protocolos dos dois leitores, apresentados a seguir, ilustram melhor esses aspectos que acabei de salientar:

Protocolo LIO:

Bom deparei com outra palavra aqui é *timber* (...) Tem então *profligate* outra palavra que eu não sei (... ..) Eu não sei o que quer dizer *bring* também esqueci Eu conheço essa palavra mas eu esqueci o significado (... ..) Bom ele o texto aqui eu entendi que *timber* deve ser algum .. algum combustível fóssil certo? .. Bom o quarto parágrafo diz o seguinte Bom dentro do sistema econômico corrente onde o estado tem o monopólio de todas as reservas naturais *timber* óleo gás então são .. então o sistema econômico é virtualmente livre certo? Isto foi encorajado tem encorajado tá .. a .. concepção *profligate* que eu não sei o que que é pela indústria certo? Então Krotov ele estima que é .. para cada unidade .. nacional certo? *Income* .. outra palavra que eu esqueci de relatar que é *income* que eu não sei o que que é *In* é dentro come é vir certo? *Income* vir para dentro certo? Deve ser indústria dentro do país alguma coisa assim certo? Então a União Soviética usa de duas a três vezes as .. as .. muitas fontes tá .. e .. as mesmas .. as .. as muitas fontes tá? em outros países também certo? Bom ele diz que isso pode ser possível tá? *to bring* eu não sei o que que é .. a concepção dentro de uma linha com outros países é .. na qual podem vai reduzindo a produtividade alguma coisa assim (... ..)

Protocolo LIE:

..o sistema econômico lá vigente né na Rússia o sistema econômico deles lá onde o estado é tem o monopólio nos recursos naturais .. *timber* também não não é nos recursos naturais .. *timber timber* apareceu noutro texto e eu não sei que que é isso *timber* .. e.. óleo *oil* e gás e óleo e *timber* deve ser *tin*er alguma coisa química *tin*er gás e também .. hum .. são hum .. virtualmente livres.

Retornando ao protocolo do LIA acima, é interessante observar como esse leitor resolve o próximo problema, relacionado ao desconhecimento da palavra *profligate*. A continuação da leitura, mesclada com a releitura do trecho problemático ou do trecho anterior ao problemático parecem ter sido os procedimentos utilizados por esse leitor. No caso em questão, o contexto criado anteriormente, imediatamente anterior à palavra, já havia sido suficiente para levá-lo a uma hipótese sobre seu significado, por sinal adequada, isto é, *excessivo*. Esse contexto é *como consequência do monopólio em todos os recursos naturais ... oil and gas and so on are virtually free*. A continuação da leitura lhe serviu para confirmar sua hipótese. Entretanto, o leitor ainda se mostra insatisfeito, e pretende coletar mais elementos ou mais pistas para confirmar seu significado "provisório". Com os elementos da sentença seguinte ou *for every unit of national income, the Soviet Union uses two to three times as many resources as other countries*, ele confirma esse significado provisório. Esses elementos são muito elucidativos, uma vez que criam redundância para o leitor. Entretanto, para que pudesse fazer uso dessa redundância, o LIA teve que processar o trecho todo, incluindo as palavras *income* e a expressão *two to three times as many*. A inferência e construção do sentido do parágrafo foram dependentes, nesse caso, do processamento de trechos maiores de discurso, assim como do conhecimento acumulado do texto até então.

Os LIO e LIE, entretanto, não conseguiram criar essa redundância, uma vez que não entenderam, dentre outras coisas, a comparação marcada pela expressão *two to three times as many resources as other countries*, *profligate* nem *income*. Esse trecho foi muito problemático para esses leitores, que fizeram várias confusões sem conseguir processá-lo como um todo, uma vez que além das palavras acima, ainda desconheciam *under*, *consumption*, *income*, *bring*, *should*, *be possible*, *into line*, dentre outras. Também não entenderam a relação semântica entre *os recursos naturais eram virtualmente livres*, que é a causa do *profligate consumption* (efeito). Além disso, desconheciam também a expressão *as many as*, ou seja, que a União Soviética *usa duas a três vezes mais recursos que outros países*, portanto, um exemplo de consumo excessivo. Restrições, portanto, de nível sintático impediram o estabelecimento de uma relação de exemplificação, ou seja, de nível semântico. Também não lhes ficou clara a sentença seguinte, que contém a proposta de solução. Dessa maneira, a compreensão de todo o trecho ficou prejudicada.

O sentido que o LIO inferiu para *income* (indústria dentro do país) é distorcido para caber dentro de um contexto também distorcido. Usou, como mostra seu protocolo, um recurso

morfológico, ou de análise da palavra em prefixo e radical. Entretanto, foi malsucedido. A falta de conhecimento acumulado do texto assim como a não percepção da classe gramatical da palavra parecem tê-lo impedido de rejeitar o sentido inferido. Outros exemplos de utilização de recursos morfológicos, por esse mesmo leitor, podem ser encontrados no parágrafo 5, mais especificamente na tentativa de inferência das palavras *profitable* e *clean up*. Ainda nesse mesmo parágrafo, o leitor achou que desconhecia *profligate*, mas na realidade, também desconhecia *consumption*, incorretamente traduzida por *concepção* nas duas ocorrências no texto. No primeiro caso, o desconhecimento do adjetivo *profligate* e sua relação com *consumption* no sintagma nominal, parecem ter colaborado para que não descartasse o sentido equivocado (*concepção*). No segundo, o desconhecimento da palavra *bring* e da expressão *into line*, além de muitas outras anteriores a ela, podem ter tornado esse contexto muito denso para o leitor. Além disso, faz um erro de segmentação, que impede a correta identificação do sujeito de *are virtually free*. Ao invés de *timber, óleo e gas* são virtualmente livres o leitor afirma que o sistema econômico é virtualmente livre, exatamente o oposto do que acontece (acontecia) na União Soviética. Esse sentido, entretanto, não é confirmado através de outras fontes, talvez por uma falta de conhecimento acumulado, um maior envolvimento com o texto, falta ou dificuldade de ativação de conhecimento prévio do assunto, ou por todos esses fatores conjuntamente. Nesse caso, portanto, mais uma vez, a hipótese foi prematura, ou seja, não verificada através de outro nível de conhecimento, que aqui, no caso, poderia ter sido conhecimento de mundo. Esse exemplo parece mostrar, de forma clara, que a ativação ou uso de conhecimento prévio em LE não parece tão simples como tem sido enfatizado.

São várias as tentativas de inferência da palavra *consumption* pelo LIE. Depois de levantar a hipótese de uma palavra inexistente em português "*consumição*", também chega a *concepção*, que não faz sentido nenhum no trecho em questão. Para esse leitor, a inferência de *profligate* também passa por tentativas, tais como "*proliferação* ou "*profileração*". Nesse caso, uma competência lexical menos desenvolvida, também em português, pode ter impedido a rejeição desses sentidos. Esse processo, caracterizado por etapas de aproximação do significado baseadas em recursos lexicais, nesse caso, relacionados à forma da palavra, sem a utilização de recursos em outros níveis de conhecimento para confirmação dessa hipótese, é, mais uma vez, condizente com a hipótese que levantei acima, ou seja, de que para leitores menos proficientes, a inferência e a construção do sentido ocorrem em etapas separadas. A dificuldade de incorporação desses significados ao sentido em construção levam o leitor a retomar o processamento do texto várias vezes, pouco acrescentando em cada uma das tentativas.

Retornando ao texto base, pode-se dizer que, enquanto os quatro primeiros parágrafos caracterizam o problema, no caso, a crise ecológica, o quinto traz uma proposta de solução. Esse parágrafo apresentou grande número de pausas, até mesmo para o LIA, que, entretanto, as solucionou com bastante facilidade e eficiência, conforme demonstram o texto-base e o protocolo

a seguir:

Texto-base: parágrafo 5

The best way to cure industry of its wastefulness is to charge the going rate for natural resources, maintains Krotov. He also proposes the addition of an "ecological supplement" to the price, the surcharge reflecting the scarcity of the resource or the damage done to the environment. Under the old system, research into the science and technology needed to clean up industry also stagnated. "It was just not profitable for factory managers to introduce clean, safe technology," said Krotov.

Protocolo LIA:

The best way to cure industry of its wastefulness desperdício is to charge the going rate for natural resources maintains Krotov Is to charge? .. charge .. charge .. charge .. cobrar? E agora? A melhor maneira de curá a indústria de seu desperdício é .. cobrá .. the going rate for natural resources. He also proposes the addition of an ecological supplement .. to the price, the surcharge reflecting the scarcity of the resource or the damage done to the environment .. Under the old system research into the science and technology needed to clean up industry also stagnated .. It was not just profitable for factory managers to introduce clean safe technology .. uhm (...) OK A melhor maneira de .. curá a indústria né desse desperdício .. é cobrá as taxas .. cobrá taxas por recursos naturais para os recursos naturais .. ele ainda propõe tá uma .. um .. ele propõe uma .. um .. preço adicional nessa taxa que deve ser cobrada das indústrias dependendo da escassez né do recurso natural .. que vai ser utilizado ou .. é .. também .. é .. dependendo do dano é .. causado ao ambiente.

No processamento desse trecho do texto, o LIA explicita o significado de *wastefulness* quase que automaticamente, sem pausa. Entretanto, não há evidências, em seu protocolo, do tipo de recurso utilizado. A inferência, nesse caso, poderia ter sido possível através da percepção da relação semântica que a palavra estabelece com a informação do parágrafo anterior, ou do conhecimento acumulado desse parágrafo, que menciona o desperdício ou *profligate consumption*, causado pelo sistema de monopólio na União Soviética. Os dados da Tarefa II mostram, entretanto, que esse leitor chegou até o significado através de um recurso morfológico, mais especificamente a análise da palavra:

Protocolo LIA (Tarefa II):

wastefulness .. não sabia inferi analisando os sufixos aí é .. ness é um sufixo formador de substantivo isso aí é substantivo Então waste to waste é gastar desperdiçar Wasteful é um adjetivo relacionado né a isso wastefulness ..desperdício ..

Depois de processar toda a oração que contém a palavra, o leitor retoma o texto, na

tentativa de incorporar o significado de *charge*, inferida de maneira automática. Entretanto, esse significado é acessado rapidamente, de maneira semi-controlada, após pronunciar algumas vezes a palavra em voz alta, como se o som da palavra, mais uma vez, buscasse ou ativasse seu significado. Depois de inferir *charge*, retoma a sentença anterior, para inferir *rate*. O significado *taxas* também é inferido rapidamente, provavelmente através da identificação da função de substantivo ou núcleo no sintagma nominal *going rate*, assim como da relação sintática (objeto) de *rate* em relação ao verbo *charge*, e, especialmente, através da percepção da relação semântica sintagmática ou de colocação que a palavra estabelece com *charge*. Uma outra possibilidade pode ter sido o uso de todos esses recursos para a ativação ou confirmação de um significado já conhecido. Na Tarefa II, a palavra *rate* não é mencionada como desconhecida, o que pode ser usado como evidência para a segunda possibilidade, caracterizando, portanto, um processamento que chamaria de semi-controlado. A palavra *going* é simplesmente ignorada, confirmando que a segmentação havia sido feita de maneira apropriada, ou seja, considerando o núcleo do sintagma nominal, *rate*, e ignorando o adjetivo. Finalmente, o significado *taxa adicional (surcharge)* é acessado rapidamente, através da criação de um contexto, que se baseia nos elementos imediatos à palavra, ou mais especificamente, com recursos semânticos, relacionando a palavra com *ecological supplement to the price*. Evidências de uso desse recurso foram, mais uma vez, obtidas através da Tarefa II. Um outro recurso lexical, baseado na palavra *surcharge*, mais especificamente em sua morfologia, ou a combinação do prefixo *sur* (em excesso) com *charge* também poderia ter sido usada pelo leitor, de forma inconsciente. Essa redundância caracterizaria esse contexto como *potencialmente rico*.

Protocolo LIA (Tarefa II):

.. *the surcharge the surcharge* quer dizer ele tá .. voltando aí um pouco a .. *surcharge* seria isso esse *ecological supplement to the price* né quer dizer então uma taxa adicional aí alguma coisa assim quando eu tava lendo nem cheguei a pará na palavra foi automático foi uma coisa automática não sei se certa mas foi.

O trecho acima foi particularmente difícil para os LIO e LIE, como mostram seus protocolos a seguir:

Protocolo LIO:

Bom no quinto parágrafo eu já deparei com outra palavra que eu não sei é (...) /wɛstɪfʊlnɪs/ *wastefulness* certo? Eu não sei o que que é /ʃɑʁʒ/ *charge* também não /reɪt/ *rate* também não *Maintains* também não sei o que que é /skɑːsɪti/ *scarcity* também não sei o que que é (... ..) Bom eu sei o que é .. *clean clean* se não me engano é apagar Agora *clean up up* é pra cima apagar pra cima eu não entendi (...) Aparece aqui a .. o .. novamente a palavra .. não não é novamente não é diferente da anterior lá no .. parágrafo anterior que é /prɒfɪtəbəl/ *profitable*. Não sei o que que é *profit able* é capaz mas *profit* eu não sei o que

que é (... ..). Bom lendo aqui o quinto parágrafo certo? diz o seguinte .. que o melhor caminho tá? pra (...) /KəR/ *cure* deve ser cura não sei tô chutando certo? que o melhor caminho pra curar a indústria certo? .. deve ser alguma alguma .. coisa /Hɛt/ *rate* deve ser ratear não não é ratear acho que usar de maneira mais racional os fontes naturais certo? Ele também propõe tá? a adição de um suplemento ecológico no preço certo? (...) A sentença onde ele fala *the surcharge reflecting the /ʊskɑːsɪti/ scarcity of the resource or the damage done to the environment* eu só entendi a última parte que é a fonte ou o gerenciamento feito no meio ambiente certo? E a primeira parte eu não entendi o que é certo? Bom, dentro do velho sistema é.. pesquisas dentro da ciência e tecnologia precisou tá? é .. sei lá deve ser apagar a indústria também apagar uma indústria também estagnada alguma coisa assim certo? Isso foi justamente não sei lá aproveitável para .. o gerenciamento de fábricas .. para introduzir o *clean* que eu não sei o que que é O *clean* é apagar mas aqui não tem esse sentido de apagar certo? Não deve ser acabar né? Não acabar também não encaixa (...) Bom enfim eu não consegui inferir a palavra .. quer dizer um sentido .. pro português pro sentido da frase aqui pra *clean* certo? (...)

Depois de identificar, mais uma vez, todas as palavras desconhecidas durante a leitura, o LIO tenta inferi-las no resumo. A inferência das palavras parece seu foco principal, e não a construção do sentido. Como a densidade de palavras desconhecidas é muito grande, torna-se difícil, senão impossível para o leitor, construir um sentido. Para inferir *rate*, tenta usar, em primeiro lugar, um recurso lexical baseado na própria palavra ou item detonador, juntamente com conhecimento em LM, como se fosse cognata, isto é, *ratear*. Entretanto, verifica que esse significado não faz sentido no trecho em questão, e logo o abandona. Convém salientar que inferência de *rate* é dificultada pelo desconhecimento do verbo do qual *rate* é o objeto, ou *charge*, e, conseqüentemente, da relação sintagmática ou de colocação entre *charge* e *rate*, ou seja, *cobrar taxas*. Acaba acessando o significado *usar de maneira racional os recursos naturais* por pura adivinhação. Muitas das hipóteses prematuras desse leitor, assim como do LIE, são baseadas na noção de cognatos, ou em padrões de reconhecimento da palavra baseados em sua forma ou som e, na falta de outra fonte de conhecimento em LE, também na LM. Embora o LIA também use esse recurso, o procedimento restringe-se, nesse caso, ao levantamento de hipóteses, a serem confirmadas, posteriormente, através de outros recursos.

Ainda no protocolo do LIO, podem-se notar pausas referentes ao desconhecimento das palavras *maintains*, *scarsity* (consideradas cognatas) e também *surcharge*, que não são inferidas, constituindo lacunas no sentido da oração onde essas palavras estão inseridas. A palavra *damage*, na segunda parte dessa oração, é confundida com *manage*, sendo traduzida por *gerenciamento*. Esse trecho é muito denso para esse leitor, e o grande número de palavras desconhecidas parece impedir o levantamento de hipóteses que seriam possíveis através do contexto imediato e do conhecimento acumulado até o momento, como fez o LIA acima. Isso também o impede de fazer a ligação com *an ecological supplement to the price*, como também fez o LIA.

Dentre outras coisas, o LIE confunde a palavra *charge* com *change* (trocar). Embora esse significado pareça difícil de ser incorporado ao sentido do parágrafo, o leitor, ao invés de refazer sua hipótese, distorce o sentido construído para tentar incluí-lo. Esse tipo de comportamento, bastante comum no protocolo do LIE, também ocorre com frequência no protocolo do LIO, como pode-se perceber nos protocolos acima, como também em outros contextos já mencionados.

A análise da palavra em suas partes, ou em prefixo, sufixo e radical, ou a utilização de um recurso morfológico, é comum no protocolo do LIA e também no do LIO. Entretanto, enquanto o primeiro é bem sucedido, o segundo não é. O protocolo do LIO acima, especificamente quando se refere ao verbo frasal *clean up*, ilustra esse aspecto. Esse verbo aparece duas vezes no parágrafo 5. A primeira ocorrência é em *clean up* e, para tentar inferi-la, o LIO recorre da análise da palavra em suas partes, identificando corretamente a partícula *up*. Entretanto, parte de uma hipótese errônea sobre o significado da palavra *clean*. Colocando as duas partes juntas, não consegue um significado que faça sentido nesse contexto. A palavra *clean* ocorre de novo, agora como um adjetivo para *tecnologia*. Ainda há uma terceira ocorrência da mesma palavra no parágrafo 10, em que ela aparece como um adjetivo no comparativo (*cleaner*). O leitor, entretanto, não relaciona as três ocorrências e o sentido do trecho em questão fica não apenas incompleto ou vago mas especialmente distorcido. Além do desconhecimento da palavra *clean*, também parece contribuir, nesse caso, a segmentação errada e conseqüentemente, a compreensão distorcida do trecho *research into the science and technology needed to clean up industry also stagnated*. Ao invés de *a pesquisa em ciência e tecnologia necessária para limpar a indústria também estagnou*, em que o verbo principal está no passado *estagnou*, o leitor conclui que *a pesquisa dentro da ciência e tecnologia precisou apagar uma indústria também estagnada*, em que o verbo *estagnar* é visto como particípio passado com função de adjetivo (para *indústria*) e a forma *needed* como verbo auxiliar de *clean up*. Mesmo percebendo que esse significado é improvável, não permitindo a construção de um sentido para o trecho, o sujeito não consegue abandonar a hipótese, mantendo-a até o fim do trecho em questão. A dificuldade na reformulação de hipóteses observada no protocolo deste leitor, assim como no do LIE poderia ser vista como um comportamento característico de leitores menos proficientes em uma situação de LE, não se limitando, portanto, a esses dois casos em particular. Ela não parece ser tão simples como parece, uma vez que a percepção de inconsistências ou a verificação da informação do texto com o contexto mais geral do leitor ou uma interpretação a nível pragmático não ocorre com facilidade em LE, principalmente quando a competência lexical do leitor é insuficiente. Nesse caso, a percepção de inconsistências não ocorre, pois o leitor está preso às malhas das relações de nível micro-estrutural ou de decodificação.

Para a inferência da palavra *profitable*, o LIO tenta a mesma estratégia acima, ou seja, a análise da palavra ou um recurso morfológico. Apesar de identificar corretamente o sufixo *able*, explicita desconhecer o radical *profit*, não conseguindo, portanto, chegar a um significado. Tenta

levantar uma hipótese sobre o significado desse radical, com base em *pura adivinhação*, ou de várias etapas de aproximação do significado, sem, contudo, explicitar a hipótese que o orienta. Distorce, mais uma vez, o sentido que está construindo para "encaixar" esse significado. Para que esse tipo de recurso possa ser usado de maneira eficiente, entretanto, parecem necessários não apenas o conhecimento das duas partes da palavra, mas também de outros recursos adicionais. Muitas vezes, como no caso de verbos frasais, esse recurso não é eficiente, uma vez que o significado do todo não é equivalente ao significado das partes.

Retomando o texto base, observa-se que o sexto parágrafo complementa o anterior trazendo mais um aspecto da proposta de solução. São observadas, especialmente nesse parágrafo, diferenças significativas entre a competência lexical dos três sujeitos, particularmente entre os LIO e LIE, por um lado, e o LIA de outro, assim como seus efeitos na busca da compreensão. O texto-base, assim como os protocolos dos três leitores, abaixo reproduzidos, trazem evidências para esse aspecto:

Texto base: parágrafo 6

Krotov's cure for the Soviet Union's environmental ills is to privatise at least half of the state factories to introduce competition among producers. This would involve existing companies setting up new enterprises and getting involved in joint ventures with foreign firms. But the Soviet Union must move away from heavy industry and step up production of consumer goods, he said. Most importantly, the country must join the international monetary system so that the rouble becomes convertible.

Protocolo LIA:

((LS)) [do parágrafo todo] Bom .. o parágrafo aqui .. diz que a sugestão né de Krotov pra solucioná o problema da União Soviética né com relação ao meio ambiente é privatizá pelo menos metade das indústrias existentes na na União Soviética pra introduzi uma competição entre os produtores né Isso faria com que eles reduzissem cada vez mais os gastos pra se tornarem competitivos né creio eu Isso já é .. uma intuição minha Bom é .. ele também diz que essa mudança né quer dizer envolveria né .. ah .. faria com que as companhias existentes se unissem formassem joint ventures formassem grupos acordos né com firmas estrangeiras É .. é .. além disso tudo que foi falado há .. haveria necessidade de do país da União Soviética é .. juntasse ao sistema monetário internacional pra que a moeda corrente lá que é o rubro se tornasse uma moeda que pudesse ser convertida né em outras tipo dólar yen talvez qualquer outra moeda assim ..

Protocolo LIO:

Bom no sexto em .. eu deparei com a palavra /ʒoɪnt/ *joint* /ʒoɪnvenʃuR/ *joint venture* que eu não entendi muito bem o que significa /ʒoɪnvenʃuR/ *joint venture* tá? e /foʀaɪn/ *foreign* que eu também não sei o que que é (... ..) Eu não entendi também o que é /Houbel/ *rouble* num .. num sei o significado dessa palavra Bom nessa .. nesse sexto parágrafo a cura que o Krotov tá? propôs para

a União Soviética sobre o meio ambiente tá? acho que fará a privatização de pelo menos metade tá? provocará a privatização de pelo menos metade das fábricas do estado pra introduzir uma competição entre os produtores certo? Isso isso envolverá tá? companhias existentes é .. isso envolverá novas companhias existente certo? e também envolveu /ʒoʊnvenʃuə/ *joint venture* em .. deve ser com outras firmas certo? ou seja vai entrar novas firmas no mercado certo? juntamente fazendo uma ligação com outras firmas certo? Mais ou menos nesse sentido certo? Mas .. a União Soviética ela precisa ir em frente certo? Ir em frente tá? pra ter uma boa um bom parque industrial alguma coisa assim certo? e ter um consumo de produção bom diz o Krotov o Krotov né certo? (...) Bom também o mais importante é que o país ele precisa de um sistema monetário internacional que .. sustente acho que isso essa estrutura de mercado certo? É .. praticamente isso que eu entendi do texto Do texto não do parágrafo seis (... ..)

Protocolo LIE:

Krotov's curou curou o que é /Kʊr/ *cure the* /Kɪʊr/ *cure* os curadores não .. Krotov's curou para a União Soviética meio ambiente curou cure não é curou não trouxe disse ou melhor .. Krotov's /Kʊr/ *cure cure* /Kʊr/ *cure* para a União Soviética meio ambiente *ills* então Krotov disse para o meio ambiente da União Sovi .. meio ambiente da União Soviética .. *ills* não sei já falei né para privatizar /lɛst/ *least* /hɒlf/ *half* o passado .. /lɛst/ *least* semana /lɛst/ *least* dos fatores menos /hɒlf/ *half* num sei do estado dos fatores estaduais para introdução da competição among entre os produtores Ah certo! então Krotov disse que a União soviética .. então Krotov cura curou curou pra cura né pra União Soviética no ambiente Soviético .. curar no ambiente soviético Krotov disse privatizar menos nos .. que os fatores est .. privatizar menos todos .. de os estados os fatores de introdução na competição entre produções Suas .. seus envolvimento .. isto envolveria envolveria *would* envolveria existência existindo companhias *set* por pondo novas /ɪntɜːpraɪzɪs/ *enterprises* novas /ɪntɜːpraɪzɪs/ *enterprises* e *get get got got* pegar não .. *get got gotten* (...) *get* dar? não .. pegar acho que é pegar .. pegar isso envolveria a existência de existindo companhias pondo novos interesses e .. pegando envolvendo pegando um /ʒoʊnvenʃuə/ *joint ventures* com .. /fɔːrɛnʃ/ *foreign* firmas com várias firmas /ʒoʊnvenʃuə/ *joint ventures* /ʒoʊnvenʃuə/ *joint venture* é fazer .. unir firmas .. unir firmas uma faz trabalho pra um outra faz o trabalho pra outra /ʒoʊnvenʃuə/ *joint venture* isso aí eu sei .. tá vendo? *joint ventures* sistema japonês como no Japão /ʒoʊnvenʃuə/ *joint venture* mas a União Soviética mostrou mais .. mostrou .. mos .. mostrou .. /mʊv/ *move away* não sei para *heavy* para sua indústria é /stɪp/ *step up* a produção de *consumers* os bons consumos ele disse .. ele disse. Quem? O Krotov Mas a União Soviética mostra mais um .. *heavy metal* *heavy* indústria e a produção de /kɒnsʊmərˈɡʊd/ *consumer goods* ele disse Mais importante importante no país *must* /ʒoʊn/ *join international monetary* sistema também que o /həʊl/ *rouble becomes* /kɒnvɜːtəbəl/ *convertable* .. convertido *convertable become became becom become became* saber conhecer é .. A União mas a União Soviética o maior serão a indústria *heavy* pesada indústria pesada e uma produção de consum.. de bom de consumos de bons consumidores consumos ele disse. Mais importantemente o país mostrar *intern* .. *moneta* .. não sei mostrar sistema monetário internacional do que saber *convert* /hʌbl/ *ruble* não sei. Então o quinto parágrafo Krotov fala o que? Como a União Soviética /Kʊr/ *cure* Krotov's /Kʊr/ *cure* da União Soviética meio ambiente privatizar menos porque .. que .. de .. os .. estado os fatores estaduais para introduzir a competição entre os produtores. Então .. privatizar .. menos .. os menos os

fatores .. do meio ambiente porque pra 'ver maior competição entre o que? Entre os produtores, certo? E isto envolveria a existência de companhias novos interesses com novos interesses e pegando envolvendo /30unvenTures/ *joint ventures* com novas firmas. Então /30unvenTures/ *joint ventures* sistema japonês certo? a união de novas .. união de firmas sistema japonês Um faz o trabalho pra outro uma troca de favores /30unvenTures/ *joint ventures* A União .. mas a União mas ele fala mas a União Soviética tem uma indústria muito pesada de consumo bom de bom consumo indústria pesada certo? De bom consumo ele disse o coisa aí Mas mais importantemente que o país mostrar a *inter monetary* .. o sistema monetário internacional do que saber conhecer convertido né fazer isso daí Mais importante o país tá bem com o sistema monetário internacional né porque fazer .. isso nas /30unvenTures/ *joint ventures* essas coisas novas certo? Então esse é o quinto parágrafo Quinto não sexto Nesse parágrafo não sei o que é /KuR/ *cure* /HaLf/ *half* (... ..) /MoVue/ *move away from* .. *move away* .. /STIp/ *step* .. *consumer goods* sei mais ou menos vão vê né /30Nt/ *joint* .. *must* /30Nt/ *joint* então *venture* é .. sistema japonês então é /30unvenTures/ *joint venture* .. *the international monetary system* rubro é /Hobol/ *rouble* isso daí .. certo ..

A diferença de extensão entre os protocolos dos três sujeitos é significativa e não pode ser ignorada, na medida em que reflete as idas e vindas do LIO e especialmente do LIE ao texto, e suas dificuldades na construção da compreensão. No processamento deste parágrafo observa-se, pela primeira vez, uma tentativa, por parte do LIA, de tornar explícita uma inferência (não lexical) com base no conhecimento acumulado do texto até o momento, que a meu ver, pode ser vista como um sinal de compreensão. Essa inferência é sinalizada pela observação do leitor "*... isso faria com que eles reduzissem cada vez mais os gastos para se tornarem competitivos né creio eu Isso já é .. uma intuição minha*".

Não se observam pausas neste parágrafo, embora os dados da Tarefa II mostrem que o leitor apenas dispunha de significados aproximados para *joint ventures* e *enterprise*, e que a palavra *rouble* era desconhecida, tendo sido inferida a partir da informação criada através do contexto. Nesse sentido, devo salientar a importância dos dados da Tarefa II, esclarecendo muitos aspectos que teriam ficado sem explicação através dos protocolos.

Protocolo LIA (Tarefa II):

Eu não sei o que quer dizer realmente assim tipo *enterprise joint ventures* mas são termos de economia mesmo né .. *joint ventures*? *cartel trust* deve ser coisa assim tipo parecida né acordos tá falando que com firmas estrangeiras algum tipo de acordo união aliança ...((LB)) Bom aqui eu não sabia o que que é mas deve ser moeda acho que não precisa nem .. /Robol/ *rouble*.

Retornando à construção do sentido, entretanto, nota-se, contudo, que a palavra *enterprise* foi simplesmente apagada. O significado aproximado de *joint venture* foi, aparentemente, resultante de um processamento semi-controlado, marcado por pausa, que permitiu ao leitor recuperar um

conceito previamente conhecido, dado que o termo em inglês é usado em português, sem tradução. A construção de um contexto em torno da palavra, mais especificamente, através de uma relação com companhias existentes e firmas, pode ter sido usado para confirmar esse significado.

Os LIO e LIE, também como o LIA, embora tenham ignorado a palavra *enterprise*, constroem um sentido para o parágrafo sem ela. O LIO, e também o LIE, conseguem, aparentemente, ativar o mesmo significado para *joint venture*. Para explicar o significado da palavra, o LIO faz referência ao sistema japonês, provavelmente o contexto em que aprendeu o conceito. Entretanto, apesar de inferir a expressão, e das várias tentativas de incorporação desse significado ao parágrafo, não consegue construir um sentido.

Embora o LIO mencione *ligação com outras firmas*, a palavra *joint* parece continuar desconhecida, uma vez que o leitor não reconhece *join* mais abaixo. O sentido do trecho é prejudicado, pois também desconhece a palavra *foreign*. *Step up* e *join* são "traduzidas" por *ter*. Nesse caso, *step up* não tinha sido identificada como problema e o leitor a ignorou na primeira leitura. Entretanto, ela parece ter sido necessária para a construção do sentido, e a lacuna foi preenchida por *ter*. A falta de compreensão dos trechos *move away from heavy industry* e *production of consumer goods* dificultou o reconhecimento de *step up*. A relação entre *bom* e *good*, causada pela frequência de *good* enquanto adjetivo, juntamente com o desconhecimento do significado da palavra como substantivo, parecem ter impedido o levantamento de outras hipóteses, mesmo quando o significado *bom* não se mostrou viável. Nem mesmo a flexão de plural (*goods*), inconsistente com o comportamento de adjetivos em inglês, não parece ter chamado a atenção do leitor. Além disso, o leitor também desconhecia *rouble*; o acesso a esse significado, por sua vez, foi dificultado pelo desconhecimento da palavra *convertible*. Assim, a construção de um sentido para todo o trecho, incluindo a sentença seguinte, acabou sendo baseada em pura adivinhação.

O uso explícito do conhecimento prévio do assunto do texto, tanto para a inferência de palavras como para a construção do sentido, apesar de identificado por Haarstrup (1989a), como uma das fontes mais usadas em seu estudo, não é fácil de se identificar nos protocolos deste estudo. Além do contexto acima, um outro exemplo de tentativa de utilização desse conhecimento, entretanto, mal sucedida, é encontrado no protocolo do LIE. Esse leitor estava tentando construir um sentido para o parágrafo 6. Seu protocolo é bastante longo e complexo, uma vez que esse trecho apresenta uma densidade muito grande de palavras desconhecidas. O único sentido que consegue construir é apresentado no final da primeira parte do parágrafo:

Protocolo LIE:

Mais importante o país tá bem com o sistema monetário internacional né porque fazer .. isso nas /30unvenTures/ *joint ventures* essas coisas novas certo?

Esse sentido, *estar bem com o sistema monetário internacional*, pode ter sido motivado por conhecimento de mundo equivocado, que confunde *sistema monetário internacional* com *fundo monetário internacional*. Uma explicação para essa associação pode ter sido o fato de que a atuação desse órgão foi muito discutida no Brasil, alguns anos atrás.

Retomando o texto base, observa-se que o parágrafo 7, também como o 6, complementa o 5, trazendo outro aspecto da proposta de solução de Krotov. A relação de continuidade ou de adição de mais um aspecto da proposta, aos outros já salientados, é sinalizada por *A second component of Krotov's scheme ...* O protocolo do LIA abaixo, juntamente com o texto base, mostram claramente que o leitor percebeu essa relação, uma vez que retoma a informação anterior antes de incluir a nova. Nesse sentido, não apenas *second component* foi fundamental, mas principalmente *scheme*. A percepção desse tipo de relação pode ser indicativa da tentativa de usar a macro-estrutura do texto como uma fonte de conhecimento, tanto para a inferência de palavras, mas principalmente, como um elemento facilitador para a compreensão. Entretanto, pode-se dizer que esse leitor apenas tenta estabelecer algumas relações entre alguns dos parágrafos e porções maiores do texto, sem, contudo, usar sua organização como um todo. Esse comportamento parece estar restrito ao LIA, uma vez que os outros dois leitores não consideram a possibilidade de uso da estrutura mais geral do texto, pois estão presos a um nível de processamento da micro-estrutura do texto.

Texto base: parágrafo 7

A second vital component of Krotov's scheme is to give local governments control over the resources in their region, by handing over what is now state property. This would give local governments a direct interest in preventing pollution and protecting their local environment. he says. "Ecological standards and rules can only be applied effectively when the local authorities are the main agent for monitoring and enforcement.

Protocolo LIA:

A second vital component of Krotov's scheme is to give local governments control over the resources in their region by handing over what is now state property. This .. would give local governments a direct interest in preventing pollution and protecting their local environment. People with the greatest stake stake (...) stake .. interesse? in the environment are the local people Quer dizer as pessoas com o maior interesse no ambiente são as pessoas da região Bom stake .. talvez seja interesse .. Vou passando pra frente Ecological standards and rules can only be applied effectively when the local authorities are the main agent for /monITORIN/ monitoring and enforcement Aqui ele diz que né além toda essa mudança no setor econômico né na privatização das indústrias na .. na instituição de vários vários acordos várias joint ventures .. entre empresas existentes na União Soviética e interesses estrangeiros além de é .. além de se direcioná né as a indústria o caráter da indústria pra indústria de bens de consumo além de torná de fazer com que o país é .. fizesse parte do sistema

monetário internacional Esse segundo parágrafo diz que .. um componente .. um segundo componente quer dizer uma outra uma outra peça vital do do esquema desse Krotov desse economista né de Leningrado do Instituto de Leningrado é .. dá mais poder dá mais controle dá um maior poder aos governos locais .. né poder esse sobre os recursos sobre a exploração sobre o aproveitamento desses recursos na região né .. Ele diz que isso daria né aos governos locais um maior interesse na .. em em se prevenir a poluição em se protegê o seu próprio ambiente local .. É .. aqui ele diz né que as regras e os padrões ecológicos que .. só podem ser aplicados efetivamente quando as autoridades locais são o principal agente né monitorando quer dizer controlando esse esforço um esforço.

No processamento desse parágrafo, o LIA faz duas pausas. A primeira ocorre logo após o referente *this*. Embora a razão não seja explicitada, parece causada pela necessidade de recuperação da informação substituída ou retomada por *this*.

Uma outra pausa refere-se à palavra *stake*. Pelo protocolo, e também pela Tarefa II, pode-se depreender que a palavra era desconhecida, tendo sido inferida a partir do contexto do parágrafo, ou através da informação *People with the greatest stake in the environment are the local people*, que é *potencialmente redundante*, pois é uma citação de Krotov que retoma a informação contida na oração anterior.

Embora não se observem mais pausas nesse protocolo, em que o sentido é construído bastante facilmente, praticamente de forma automática, nota-se, entretanto, através de uma comparação com o texto-base, que o final do parágrafo (*for monitoring and enforcement*) foi entendido como *monitorando esse esforço*, possivelmente pelo fato de o leitor não ter percebido a função de substantivo do gerúndio *monitoring*. A palavra *enforcement*, "traduzida" por *esforço*, por sua vez, foi inferida, aparentemente, a partir da semelhança da palavra com o português, ou de uma pista lexical que tem por base a forma da palavra em LE e o conhecimento da LM ou noção de cognato. Nesse caso, o comportamento do LIA assemelha-se ao dos outros dois leitores, caracterizado pela falta de observação de outros aspectos, como, por exemplo, a relação entre *enforcement* e *monitoring*, que nesse caso, não é de sujeito e objeto, mas de dois substantivos, ligados pela conjunção *and*. A consideração de um outro nível, neste caso, sintático ou de estrutura da oração, teria sido importante. Apesar de equivocado, o significado inferido não causou rupturas ao sentido do trecho, e o leitor não percebeu inconsistências; o conhecimento acumulado não foi, nesse caso, suficiente para excluir essa possibilidade.

A palavra *enforcement* foi inferida incorretamente (*força*) pelo LIO, possivelmente através da percepção da semelhança com o português, e também a partir de informações da própria oração, quando fala em *regras* e *monitoramento*. Esse leitor, entretanto, parece reconhecer *monitoramento* como um substantivo, contrariamente ao LIA. A palavra *enforcement* foi inferida como *eficientes* pelo LIE, praticamente a partir da pista lexical ou da forma da palavra juntamente com pura adivinhação, uma vez que o conhecimento acumulado desse trecho é completamente distorcido. O

sentido da última sentença (que contém a palavra *enforcement*) fica prejudicado, uma vez que outras palavras também são desconhecidas. *Ecological standards* é simplificado para *ecologia*. *Can only be applied effectively* é entendido como algo sem sentido ou o seu próprio efetivamente e *local authorities* por *local das autoridades* e finalmente *are the main agent for monitoring and enforcement* é traduzido por *são agentes monitorando eficientes*. Em todos esses observam-se problemas relacionados principalmente à estrutura sintática dessas frases. Assim, a primeira tentativa de construção de um sentido para essa sentença é o trecho completamente sem sentido a seguir:

Protocolo LIE:

.. *ecological standards* tá próprio ... então o seu próprio efetivamente quando o local das autoridades são agentes monitorando eficientes Então vai ficar mais eficiente o controle sobre a população porque o governador vai tomar conta de sua própria .. região acho que é isso tá ..

Juntando essa informação truncada com informação da sentença anterior, o leitor faz uma segunda tentativa de construção do sentido:

Protocolo LIE:

Ecologia .. é própria tem mais .. eficiência quando as autoridades .. agentes são agentes de seu próprio ambiente certo? Certo ..

A relação da informação deste parágrafo com os anteriores, sinalizada principalmente por *um segundo componente do esquema de Krotov* acima referida, não foi percebida pelo LIO, que deixou de criar redundância através da citação de Krotov, que retoma a informação da oração anterior, principalmente por desconhecer a palavra *scheme*, assim como outras no parágrafo, tais como *greatest*, *stake*, e também pela falta de conhecimento acumulado mais completo. Embora mencionasse em seu protocolo que iria descobrir, provavelmente através da continuação da leitura, verifica-se que não foi bem sucedido. O protocolo abaixo é ilustrativo desses aspectos.

Protocolo LIO:

Não sei o que é *scheme* [soletra em português s-c-h-e-m-e]. Não sei o que que é (... ..) Eu não entendi a a .. expressão *handing over* certo? *Over* é sobre *handing hand* é mão botando a mão sobre? .. não sei (... ..) Bom eu encontrei outro problema aqui .. que é .. quando ele fala assim .. *people with the greatest stake* (...) não sei o que é *stake* certo? Aliás eu não sei o que é a expressão *greatest stake* certo? Não sei mesmo mas vou ver se eu descubro (... ..) Bom o que ele fala aqui no .. no .. sexto parágrafo sext .. não primeiro segundo terceiro quarto quinto sexto sétimo no oitavo certo? no oitavo parágrafo é o seguinte que o segundo componente vital na proposta dever ser proposta de Krotov certo? E um controle gover *governamental local* tá? sobre as fontes e .. naquela região

certo? .. É agora eu não entendi como fica *handing over* né (...) tá mas deve ser algum sobre colocando a mão colocando esse controle governamental sobre propriedades as (...) as agora as propriedades do estado certo? Bom isto dará ao governo local o direito .. um direito .. no .. um interesse no controle tá? da poluição e proteção dos seus locais de meio ambiente certo? Então depois ele fala aqui que .. pessoas com greatest stake que eu não sei o que é do meio ambiente, são as pessoas locais ele diz né Bom os padrões ecológicos e as regras podem ser bem aplicadas efetivamente quando as autoridades locais são o principal agente de monitoramento e .. é.. de .. né (...) Bom (...) ele fala assim *monitoring and enforcement* certo? é a força e monitoramento naquela região certo? na região local Tá bem vou passar para o próximo parágrafo o nono (... ..)

Protocolo LIE:

Vital .. segundo .. vital o segundo componente vital de Krotov's /SK&M/ *scheme* não sei é trazer é dar né é eleger acho que é é botar governadores .. então um segundo componente vital para ele ele Krotov é botar governadores .. dar gover .. controle aos governadores certo? pro local .. pros locais de suas próprias regiões certo? Então o governador vai controlar o meio ambiente de sua própria região certo? de sua própria região Assim fazendo agora sua própria proprietário por *handing handing* não sei que .. o que agora estado proprietário Isto .. isto daria .. o que .. os governadores locais uma direta faria com que os governadores locais tivessem um direto interesse em proteger o ambiente porque é da sua região certo? Porque é o local de seu povo ele disse tá porque .. direta proteção prevenindo o meio ambiente e protegendo o local meio ambiente Certo? então isso daria uma direta .. prevenção da poluição da população de poluição né e um e a proteção do meio ambiente .. tá então os .. então os .. os com grandes ... gostam de seus próprios meio ambiente né não vai querer que sujam que .. seu local e a ecologia .. *ecological standards* tá próprio ... então o seu próprio efetivamente quando o local das autoridades são agentes monitorando eficientes Então vai ficar mais eficiente o controle sobre a população porque o governador vai tomar conta de sua própria .. região acho que é isso tá Esse é o sétimo parágrafo. Então, o segundo .. o segundo componente vital pra Krotov é que os governadores controlem os recursos de suas regiões .. e novas propriedades novos estados propriedades certo de suas próprias regiões porque isso dá o que? Isso dá aos governadores os recursos na .. um direto interesse na prevenção da população da poluição e proteção do meio ambiente e para os grandes não querem que o meio ambiente seus .. né vantagens pro povo ele disse Ecologia .. é própria tem mais .. eficiência quando as autoridades .. agentes são agentes de seu próprio ambiente certo? Certo. Esse é o sét .. parágrafo sétimo parágrafo Agora o oitavo .. parágrafo .. oitavo parágrafo (...)

A semelhança de *scheme* com o português (esquema) não foi percebida pelo LIO nem pelo LIE. Nos dois casos, a pronúncia foi incorreta: o LIO apenas a soletra no protocolo e a pronuncia incorretamente na Tarefa II (/ʃɛM/). O LIE a pronuncia ora /SK&M/ ora /ʃɛM/. A palavra, entretanto, é inferida pelo LIO como *proposta*, não como *esquema*, possivelmente a partir do contexto *segundo componente vital ...*, que é um contexto potencialmente rico. Apesar de ter chegado a um significado para *scheme*, não retoma o texto para tentar reprocessar o parágrafo e incorporar esse significado.

O LIE não conseguiu inferir *stake*, ignorando a sentença que contém essa palavra, uma vez

que já lhe havia sido difícil entender a sentença anterior. Apesar dos problemas encontrados, conseguiu um sentido para a sentença anterior que foi bastante razoável:

Protocolo LIE:

Então o governador vai controlar o meio ambiente de sua própria região certo? de sua própria região ou numa segunda tentativa isto daria .. o que .. os governadores locais uma direta faria com que os governadores locais tivessem um direto interesse em proteger o ambiente porque é da sua região certo? Porque é o local de seu povo ...

Retomando a construção do sentido do texto pelo LIA, observa-se no parágrafo 8, apenas uma pausa, causada mais uma vez, pelo desconhecimento, ou mais especificamente, pelo não reconhecimento automático da palavra *boundaries*, como mostram o texto base e o protocolo abaixo:

Texto base: parágrafo 8

The Russian republic contains most of the Soviet Union's natural resources – and has the greatest environmental problems. But the government of the republic, led by Boris Yeltsin, has already begun to take steps to remedy its own problems. It recently announced its intention to take control of all the natural resources within its boundaries. Yeltsin has introduced a 400-day programme to revitalise the Russian economy, with backing from President Gorbachov. And from 1 January, an innovative "polluter pays" policy comes into force.

Protocolo LIA:

Bom .. *The Russian republic contains most of the Soviet Union's natural resources — and has the greatest environmental problems But the government of the republic led by Boris Yeltsin .. has already begun to take steps to remedy .. its own problems. It /reSɛntɪ/ recently announced its intention to take control of all the natural resources within its /baʊndə-rɪs/ boundaries .. /baʊndə-rɪs/ boundaries .. fronteiras? O que é /baʊndə-rɪs/ boundaries? Deve ser fronteiras .. /baʊndə-rɪs/ boundaries .. Yeltsin has introduced a 400-day programme to revitalise the Russian economy with backing from President Gorbachov .. And from first January an innovative "polluter pays" policy comes into force* Aqui nesse parágrafo diz que a Rússia a república da Rússia contém a maior quantidade de recursos naturais e também tem os maiores problemas com relação ao ambiente Mais esse governo .. ou seja o governo dessa república, desse .. da .. da Rússia já tomou alguns já vem tomando né alguns passos em direção ao .. à solução desses problemas né .. Esses passos .. podem ser os seguintes recentemente o governo anunciou sua intenção de controlá né é .. de tê um controle de todos os recursos naturais dentro de suas fronteiras O governo já anunciou um programa de revitalização da economia russa com uma política é .. de .. de multas né aos às pessoas que poluem né quer dizer eles vão passar a aplicar multas às pessoas que provocam a poluição que poluem o ambiente Quer dizer em resumo isso deve ter ficado meio confuso A Rússia que é uma república da União Soviética contém né o .. a maior quantidade de recursos naturais da União Soviética e também contém os maiores problemas com relação ao meio ambiente Mais a Rússia já vem tentando .. é solucionar seus problemas

através de várias medidas né descritas aí.

A palavra *boundaries*, assim como ocorreu outras vezes, é repetida algumas vezes em voz alta. Não fica claro, entretanto, pelos dados do protocolo, quais os recursos utilizados pelo LIA. São várias as possibilidades: já ter encontrado a palavra antes e ter-se esquecido dela, e, nesse caso, a construção do contexto seria usada para ativar esse conhecimento; relação semântica com *within*, que aparece no contexto imediato da própria palavra, juntamente com o conhecimento acumulado do texto até esse ponto; conhecimento acumulado do parágrafo que contém a palavra; conhecimento prévio da situação na Rússia; ou ainda todas essas fontes de conhecimento juntas. O protocolo da Tarefa II, nesse sentido, não é muito esclarecedor:

Protocolo LIA (Tarefa II):

Não sei se é isso mesmo /baunDə-rs/ *boundaries* .. creio que isso aqui seja fronteira limite ..

Embora a construção do sentido do parágrafo tenha procedido sem maiores problemas, observa-se, contudo, comparando o protocolo com o texto-base, que *400-day* e *backing* foram apagadas, não causando rupturas, uma vez que eram dispensáveis a esse sentido.

O LIO, por sua vez, consegue inferir nesse parágrafo a palavra *greatest*, em sua segunda ocorrência, como mostra seu protocolo abaixo:

Protocolo LIO:

Bom acho que agora eu descobri /GRɛtɛst/ *greatest* Deve ser alguma coisa muito (...) o maior problema (...) Bom eu deparei com uma palavra que eu não conheço é /aNaunNsɛd/ *announced* (...) não sei o que que é (...) Bom tem uma palavra aqui chamada assim /baunDə-rs/ *boundaries* certo? Bom eu sei que /baunDə-r/ *boundary* deve ser contorno certo? /baunDə-rs/ *boundaries* eu não sei o que que é Pode ser um falso cognato certo? (...) Bom o que ele diz aqui no nono parágrafo [referindo-se ao oitavo] é o seguinte que a república russa é .. contém a maioria dos fontes naturais da União Soviética e devido a isso é .. tem os maiores problemas do meio ambiente ambientais né? Mas o governo da república certo? Não sei o que é que é /Lɛd/ *led* Sei que /Lɛd/ *led* é a luzinha que acende no *drive* so isso Tá tem é .. /alRɛd/ *already* Sei que /Rɛd/ *ready* é pronto tá? pronto certo? /alRɛd/ *already* não sei o que que é Começou com pequenos com passos pra remediar esse problema certo? Quer dizer são os problemas ambientais tá? Ele recentemente é .. anunciou sua intenção de fazer um controle de todas as fontes naturais certo? com /baunDə-rs/ *boundaries* que eu não sei o que que é certo? Esse Yeltsin também introduziu é .. um programa diário para revitalizar a economia russa con .. conjuntamente com o presidente Gorbachev e .. em primeiro de janeiro um .. pagamento .. dever ser um imposto novo né certo? E ele fala *innovative polluter pays* .. eu não sei o que é "*polluter pays*" /poLaɪs/ *policy comes into force* .. certo? Eu não entendi bem o final disso aí o final do parágrafo (... ..)

Bastante interessante é a tentativa de inferência de *boundaries* pelo LIO. Embora explicita conhecer a palavra *boundary* como *contorno*, não consegue, através do contexto da oração e do parágrafo, fazer uma adaptação para *fronteira* ou *limite*, mais adequado para a construção do significado da sentença. Essa falta de flexibilidade, que poderia ser explicada pela associação da palavra a apenas um significado, também é observada em outra pausa neste parágrafo, mais especificamente, naquela causada pela palavra *led*, tendo sido também freqüente em outros parágrafos e também no estudo de produto. Mais uma vez, os problemas, nesse caso, não se restringem ao contexto da palavra, mas também incluem a oração anterior àquela que contém a palavra, ou, mais especificamente, o parágrafo todo. O conhecimento acumulado não parece ter sido, nesse caso, suficiente.

As pausas causadas por *400-day* e *backing* foram solucionadas como *diário* (confundindo com *daily*) e *conjuntamente*. Este último significado parece ter sido possível a partir do contexto *with backing from President Gorbachov*. O leitor deixou de observar a preposição *from*, que teria eliminado a possibilidade *conjuntamente*. Apesar de inadequados, não interferiram na construção do sentido. Embora o LIO não tenha, aparentemente, entendido o sentido da última sentença, consegue inferir que se trata de um imposto novo.

O LIE infere *400-day* corretamente, usando *backing* em inglês. Esse procedimento é usado com freqüência por esse leitor, dado que o número de palavras desconhecidas é bastante grande; o uso da palavra em inglês, em alguns trechos, lhe permite, focalizar o trecho em questão por etapas, focalizando uma palavra desconhecida de cada vez.

Como já havia mencionado anteriormente, um outro exemplo de um comportamento que é característico do processo de construção do sentido pelo LIE é a tentativa de inferir um significado através da ativação de conhecimento sintático de verbos (nesse caso, *take* e *begin*), através de uma explicitação de suas partes principais, que são repetidas em voz alta, seguindo por um outro, em que o verbo é inserido em um outro contexto mais simples e conhecido, provavelmente o contexto em que a palavra foi ensinada. Este tipo de procedimento é usado inúmeras vezes no protocolo desse leitor, sendo, algumas delas, nesse parágrafo (vide, por exemplo, *get* e *become*, no parágrafo 6). Embora esse recurso tenha sido eficiente para acessar o significado de *begin* e *take*, não solucionou o problema do leitor, uma vez que *take* fazia parte de dois verbos frasais *take steps* e *take control*, cujo sentido não poderia ser inferido a partir da soma de suas partes. A integração dos vários níveis (lexical, fonológico e sintático) de uma competência lexicial *rica* fica aparente, nesse caso.

O LIE interrompe a leitura no final deste parágrafo, pois se mostrava cansado, uma vez que sua tarefa de leitura já durava horas, como pode-se observar pela extensão de seus protocolos.

O nono parágrafo do texto-base, a seguir, apresenta as possíveis implicações das mudanças propostas por Krotov:

Texto base: parágrafo 9

The new law means that about 100 of the 500 major industrial concerns in the Leningrad region will have to shut down because they cannot meet the new standards set by the Russian authorities. An optimistic Krotov believes that the extra money coming into the republic from the export of its resources both abroad and to other parts of the Soviet Union will help to reduce unemployment by providing funds to set up new, cleaner industries. "Economic reform in the Russian republic will make it possible to harmonise entrepreneurial activity and the rational use of resources with the protection of the environment," said Krotov.

Nesse trecho, o protocolo do LIA abaixo indica duas pausas. A primeira está relacionada a *concerns*, empregada no texto como substantivo. A palavra, aparentemente, parece reconhecida pelo leitor apenas em sua função mais corrente ou verbo, causando-lhe estranheza e ansiedade como substantivo. O procedimento empregado pelo leitor é a continuação da leitura em busca de informações adicionais. Entretanto, no final no parágrafo, o desconhecimento da palavra *entrepreneurial* detona outra pausa, fazendo com que sua atenção volte para a solução desse novo problema. Para isso, relê a oração que contém essa palavra, sem, contudo, chegar a nenhuma conclusão. Pela primeira vez, seu processo de construção do sentido sofre uma ruptura, uma vez que, na concepção do leitor, essa palavra é imprescindível para o sentido do texto e não pode ser apagada. O protocolo, a seguir, ilustra esse aspecto:

Protocolo LIA:

Bom .. *The new law means that about 100 of the 500 major industrial concerns in the Leningrad region will have to shut down because they cannot meet the new standards set by the Russian authorities. The new law means that about 100 .. of the 500 major industrial concerns .. will have to shut down .. concerns? ..* Que que que é *concerns*? .. 'Tá estranha essa colocação a colocação dessa palavra aí .. Bom vamos lá *An optimistic Krotov believes that the extra money coming into the republic from the export of its resources both abroad and to other parts of the Soviet Union will help to reduce unemployment by providing funds to set up new cleaner industries Economic reform in the Russian republic will make it possible to harmonise .. /entreprẽNeunal/ entrepreneurial ..* Essa palavra eu não sei o que que é não /entreprẽNeunal/ *entrepreneurial .. Economic reform in the Russian republic will make it possible to harmonise /entreprẽneunaL/ entrepreneurial activity and the rational use of resources with the protection of the environment* Esse último parágrafo .. harmonizá (...) ah .. tá difícil essa palavra (...)

Como a pesquisadora estivesse presente nesse momento da coleta, resolve interferir para tentar ajudar o leitor. A interação é apresentada no protocolo a seguir:

PE: Alguma idéia alguma hipótese?

LIA: Não .. a princípio não .. A reforma econômica na Rússia tornará possível

harmonizá .. /*entrepreNeunal/ entrepreneurial* .. deixa eu voltá lá pra cima, vamo vê se eu consigo .. vê .. esclarecê alguma coisa aqui. Pera aí *The new law means that about 100 of the 500 major industrial concerns in the Leningrad region will have to shut down because they cannot meet the new standards set by the Russian authorities* An optimistic Krotov believes that the extra money coming into the republic from the export of its resources both abroad and to other parts of the Soviet Union will help to reduce unemployment by providing funds to set up new, cleaner industries Economic reform in the Russian republic will make it possible to harmonise /*entrepreNeunal/ entrepreneurial* activity .. ahn ahn não sei não sei mesmo O parágrafo diz que essa nova lei .. que a Rússia que o governo russo implantou instituiu provavelmente fará com que uma grande quantidade de fábricas né sejam obrigadas a fechar por não conseguirem .. atingir os novos padrões estabelecidos pelas autoridades russas. Mas ele acredita mas o Krotov não .. o economista é .. acha que .. acredita que o dinheiro extra que vai entrá na república né através da exportação dos recursos dessa república pro exterior e pra outras partes da União Soviética vai ajudá a reduzir o desemprego né é .. assim é .. através é .. da .. é.. de de através da .. criação de fundos que ajudem a estabelecê construí novas indústrias né novas indústrias que estejam de acordo com a .. a .. os padrões dessa nova lei instituída pelo pelas autoridades russas Agora o que tá falando nesse último parágrafo que é um comentário dele eu num essa palavra tá é meio importante aí eu não sei ..

PE: Você tem alguma idéia da classe gramatical dessa palavra? Pode ajudá?

LIA: É um adjetivo

PE: É Pode deixá de lado?

LIA: Não não dá acho que não dá pra se deixá de lado. 'Tá 'tá 'tá .. dando uma característica dessa atividade. Essa atividade 'tá sendo comp .. 'tá sendo .. comparada aí com o uso racional dos recursos .. quer dizer ..

Embora o LIA identifique a classe gramatical da palavra *entrepreneurial*, não consegue inferir seu significado nem construir um sentido sem ela, pois a considera essencial, embora o sentido da última oração do parágrafo não fique prejudicado. A pausa detonada pela palavra *concerns*, núcleo do sintagma nominal *industrial concerns*, salientada no início do parágrafo, acaba sendo solucionada a partir do momento que o leitor percebe que pode construir o sentido sem ela, mais especificamente invertendo o foco para o adjetivo ao invés do substantivo; *industrial concerns* é, então, simplificado por *indústrias*.

O LIO, nesse caso, chegou ao mesmo significado para *concerns*, provavelmente usando o mesmo recurso, embora ela não tenha sido um problema consciente, pois não havia sido identificada por uma pausa; na construção do sentido, foi apagada. Seu protocolo, entretanto, apresenta muitas outras pausas, também detonadas por palavras desconhecidas, tais como *unemployment*, *step up*, *join*, dentre outras, que acabam dificultando sua compreensão, como mostra seu protocolo a seguir:

Protocolo LIO:

Bom o próximo parágrafo que eu 'tou lendo aqui eu não entendi aqui o que é a

palavra /ʃut/ *shut* é .. /ʃutDauN/ *shut down* certo? (... ..) Uma palavra aqui também eu bom .. também eu (...) Bom eu já vi essa palavra mas eu não lembro o significado dela que é /Bout/ *both* certo? E tem a palavra seguinte que é também /aBRoud/ *abroad* que eu também não sei o que que é eu nunca vi na vida (... ..) Tem outra palavra aqui também que eu não não entendi nunca vi na vida /entrepreNeunAl/ *entrepreneurial* certo? Não sei o que que é (...) Bom esse décimo [referindo-se ao nono] diz o seguinte que as novas as novas leis tá? que .. foram seriam impostas de 100 das 500 maiores indústrias de na região de Leningrado terão é .. devido a isso certo? eles não podem colocar colocar novos padrões .. pelas autoridades russas Não ficou muito legal certo? (...) Acho que é o seguinte Acho que de 100 dessas indústrias maiores indústrias da região de Leningrado eles não conseguiram eles não obedeceram as autoridades russas colocando esses novos padrões de .. de .. de .. lei contra o proteger o meio ambiente certo? Bom o otimista Krotov acredita que o dinheiro o dinheiro extra certo? vindo tá pra dentro da república devido a exportação de seus .. de suas fontes, certo? precisa ser acho que aplicado em outras partes da União Soviética e .. e que ajudarão a reduzir *unemployment* certo? a reduzir deve ser a diminuição de .. fundos a diminuição de fundos pelo .. não não é reduzindo .. reduzindo alguma coisa certo? Depois ele fala pelo fornecimento de fundos pra .. *set up new cleaner industries* as .. eu não sei o que é *cleaner* certo? Depois ele diz aqui que uma reforma econômica da república russa tá? fará .. é .. será possível harmonizar o /entrepreNeunAl/ *entrepreneurial* que eu não sei o que que é, atividade atividade e o uso racional dos das fontes com proteção do meio ambiente diz Krotov certo?

Devido ao grande número de palavras desconhecidas, a inferência, assim como a construção de contexto para *unemployment* torna-se impossível. O leitor chega ao significado *diminuição de fundos* sem, contudo, explicitar os recursos mobilizados. Como esse significado não se encaixa no sentido em construção, este é abandonado.

Nesse trecho, também desconhece os verbos *step up* e *join*. O que parece ocorrer, nesse caso, é que durante a leitura, o leitor não havia identificado *step up* como problema. Entretanto, na construção do sentido, a lacuna causada pela falta do significado foi preenchida por *ter*. A falta de compreensão de *move away from heavy industry*, assim como do trecho seguinte *production of consumer goods*, impossibilitou o acesso a *step up*. Assim, o sentido da oração toda acabou sendo baseada em pura adivinhação, incluindo o da oração seguinte.

Retomando o texto-base, observa-se no último parágrafo, o resultado ou avaliação final sobre a proposta de solução para a crise ecológica proposta por Krotov:

Texto base: parágrafo 10

While these reforms will bring benefits to the Russian federation, they will cause problems in other parts of the country and in the countries of Eastern Europe, which, until now, have had almost unlimited access to raw materials for virtually nothing

A única pausa identificada no protocolo do LIA refere-se a *raw materials*, como mostra o protocolo abaixo:

Protocolo LIA:

Bom último parágrafo ((LB)) [do parágrafo todo] (...) Hum .. Enquanto essas reformas trarão benefícios .. à Rússia .. elas trarão elas causarão problemas em outras partes do país e em outros países da Europa do Leste que até agora .. tiveram quase tiveram um acesso quase ilimitado to *raw materials* .. *raw materials* seria matéria prima .. for virtually nothing por quase nada .. Bom (...) essas reformas elas 'tão trazendo benefício pra Rússia mas 'tão causando problemas em outras partes do país .. porque o estado compra quer dizer ele vai comprar da Rússia e .. a Rússia tá exportando a .. essa matéria prima esses recursos naturais tanto pros outros países como pras outras partes da União Soviética né Quer dizer esse dinheiro que 'tá entrando na Rússia vai sê benéfico para ela Mais enquanto isso .. os outros países os outros estados e os outros países vão se prejudicá com isso .. Por que?

Apesar da pausa, o leitor infere *matérias primas (raw materials)* rapidamente, dando a impressão de já conhecê-la, da mesma forma que ocorreu com *joint-ventures*, num processamento semi-controlado. Nesse caso, as evidências da Tarefa II são mais uma vez esclarecedoras, mostrando que a palavra *raw* realmente era desconhecida e ainda continua; o significado do sintagma *raw materials* é que foi inferido, aparentemente, a partir de um recursos sintático, ou da relação que *raw* estabelece com *materials*, juntamente com conhecimento acumulado do texto até o momento. O leitor também pode ter-se utilizado de um outro recurso, isto é, o uso do conhecimento acumulado para ativar um significado conhecido, embora não completamente estabelecido. O fato de ainda mostrar desconhecimento da palavra *raw* na Tarefa II não exclui essa possibilidade, uma vez que *raw materials* e *raw* são conceitos diferentes. O mesmo caso pode ser observado com as palavras *joint-ventures* e *joint*, que aparecem no mesmo parágrafo.

Protocolo LIA (Tarefa II):

Raw materials aqui matérias primas né Inferi também *Raw* .. eu já ouvi essa palavra antes mas não guardei qual o significado dela ..

O processamento do parágrafo 10 pelo LIO também é interrompido pelo desconhecimento da palavra *raw*, como mostra seu protocolo a seguir:

Protocolo LIO:

Tou lendo agora o último parágrafo (... ..) Eu não sei o que é a palavra /Haw/ *raw* /tuHaw/ to *raw* Bom o último parágrafo diz o seguinte enquanto essas reformas tá? são benéficas para a federação russa elas causam elas causarão problemas em outras partes do país do país certo? E nos países da Europa

externa deve ser no oeste né certo? Europa Bom países da Europa aí grudados ali na União Soviética .. na qual até agora eles eles não tem .. tido um acesso limitado não limitado né ilimitado de materiais para (...) /tuHau/raw /tuHau/ to raw eu não sei o que que é Depois ele fala materiais para virtualmente nada certo? É praticamente .. eu não entendi o finalzinho .. devido o que eu não sei a palavra o que significa a palavra /tuHau/to raw ..

Entretanto, o leitor é malsucedido, pois faz a segmentação de forma incorreta da primeira oração, fazendo de *raw* um verbo (*to raw*). Essa segmentação acaba por eclipsar a percepção da classe gramatical dessa palavra (adjetivo) e sua relação com o substantivo *materials* no sintagma nominal, impedindo a ligação com a última parte da oração ou *for virtually nothing* e, conseqüentemente, a construção de um significado para a última parte do parágrafo.

Retornando ao protocolo do LIA, observa-se, entretanto, em sua parte final, que o leitor não parece satisfeito com sua compreensão. A pesquisadora, presente nesse momento da coleta de dados, interage com o leitor, na tentativa de ajudá-lo a explicitar seu problema. O que se pode depreender, nesse caso, é que o LIA não tinha uma noção de onde estava o problema, dado que não havia sido detonado, desta vez, por uma palavra.

PE: Alguma coisa não 'tá fazendo sentido pra você?

LIA: É aí é que tá eu não entendi porque essas reformas .. trarão benefícios pra Rússia mas causarão problemas pras outras partes .. do país (...) Tá essas reformas algumas dessas reformas poderiam .. aumentá ainda mais o acesso a essas matérias primas .. às outras partes (...) porque .. bom .. seria talvez .. bom .. a Rússia exportando pra pras outras partes facilitaria então o acesso dessas partes às matérias primas Quer dizer é .. voltando aqui a um problema discutido logo no começo o problema na União Soviética é que .. como tudo é estatal tudo é monopolizado o acesso a essas matérias primas é muito fácil quer dizer então eles eles consomem muito eles desperdiçam matéria prima E isso .. é prejudicial pro ambiente quer dizer .. eles poderiam diminuir o consumo da da matéria prima sem afetar a produtividade Né agora então deve ser mais ou menos isso quer dizer essas reformas de uma .. de alguma forma facilitariam ainda mais esse .. o o acesso .. às matérias primas e fariam com que .. então .. é .. eles .. num .. meçam direitinho a quantidade certa pra pra consumo talvez consumam em em .. excesso continuem consumindo em excesso ..

O LIA não parecia entender porque a proposta de Krotov, embora trazendo benefícios para a Rússia, pudesse ocasionar problemas em outras partes da União Soviética. Sua interpretação é de que esses problemas seriam de consumo excessivo, o que é inconsistente com a informação do texto contida na proposta. O leitor não parece ter levado em conta, nessa interpretação, a informação da última oração, ou seja, que até aquele momento, esses países tiveram acesso quase ilimitado às matérias primas, mas que, com a introdução das medidas propostas por Krotov, terão que pagar por esses recursos. Portanto, os problemas estariam relacionados, mais especificamente, aos gastos

futuros com essas matérias primas.

A análise, que acabei de apresentar, não parece deixar dúvidas com relação à compreensão alcançada pelo LIA. Os resumos dos parágrafos confirmam de maneira inequívoca esse aspecto, reafirmada, mais uma vez, pelo resumo conciso que o leitor faz nessa parte final do texto, em que retoma pontos relevantes e explicita questões fundamentais. Apesar de alguns problemas menores, o sentido que constrói mostra coerência, o que pode ser caracterizando como compreensão (Koch & Travaglia 1989). Entretanto, seu protocolo também revela que esse sentido foi construído de forma linear, sem questionamentos, mantendo-se fiel aos dados do texto, num processo em que os significados não são exatamente negociados mas aceitos, e que poderia ser mais bem definido como uma "tradução", mais próximo, portanto, de uma interpretação semântica. A interpretação pragmática, nesse caso, não parece ter ocorrido.

Uma confrontação dos itens lexicais chaves do texto base (responsáveis pela saliência-textual ou relevância-autor) (vide início deste capítulo), com os itens contextualmente relevantes que o leitor escolhe focalizar (responsáveis pela relevância-leitor), assim como uma confrontação das cadeias coesivas estabelecidas a partir dos itens lexicais chaves, com os sentidos construídos pelo leitor, seguindo o arcabouço teórico proposto por Cavalcanti (1983, 1989) confirmam, de maneira mais objetiva, os resultados da análise. Nesse caso, os itens focalizados pelo leitor em sua interação com o texto parecem os mesmos itens que o autor salienta, considerados, portanto, itens chaves no texto. A identificação desses itens pelo leitor e, principalmente, das cadeias coesivas por eles encabeçadas, permitem a identificação do foco do texto e de sua estrutura problema-solução, levando-o à antecipação de seu conteúdo. Para Cavalcanti, entretanto, a interpretação pragmática estaria no "... equilíbrio entre a saliência textual e a relevância-leitor". Esse equilíbrio, entretanto, não parece ter ocorrido, dado que não se nota negociação de sentidos, mas apenas a aceitação da saliência-leitor ou seja, uma interpretação semântica.

Com relação aos outros dois leitores, os dados, entretanto, não mostram a construção de um sentido geral para o texto, ou uma continuidade de sentidos que caracterize a coerência e, conseqüentemente, a compreensão, embora alguns focos esporádicos possam ser identificados.

A confrontação dos itens lexicais do texto-base, com os itens contextualmente relevantes do protocolo do LIO (o LIE estaria representado nessa análise, uma vez que seu padrão de leitura se assemelha ao do LIO), assim como uma confrontação dos sentidos construído pelos LIO com as cadeias coesivas do texto-base, também confirmam a conclusão baseada na análise dos protocolos. Esse processo, que se inicia a partir do título *A capital cure for the Soviet Union's ecological ills* deixa claro que desde o início da Tarefa I, o leitor já mostra desconhecimento das palavras *capital cure* (apesar da semelhança com o português) e *ills*, portanto, dois dos três itens lexicais chaves (*Soviet Union* é o terceiro), que recorrem em várias partes do plano principal do discurso, mantendo a unidade temática e a continuidade do sentido, como mostrei no início deste capítulo. A análise dos

resumos indica que somente no quinto parágrafo o LIO consegue inferir *cure*. Entretanto, a inferência desse significado nessa parte do texto, não parece útil para o leitor, que não retoma o texto para tentar reprocessar as outras ocorrências da palavra e incorporar esse sentido aos trechos em que ocorrem. Dessa forma, *ecological* (ou apenas *ecologia*) e *Soviet Union*, passam a ser os itens contextualmente relevantes para o leitor, usados para a identificação do foco e da estrutura problema-solução do texto e, conseqüentemente, para a antecipação de seu conteúdo e ativação de conhecimento de mundo. Em nenhum momento, o leitor menciona *capital* (cognata), sugerindo que desconhece essa palavra, ou não a considera relevante (o LIE, apesar de identificar *capital*, não o relaciona ao resto do texto). Esse conhecimento limitado impossibilitou a identificação das cadeias coesivas encabeçadas pelos itens lexicais chaves (principalmente por *capital cure*) e a função dessas cadeias na estruturação do texto. Uma análise dos itens que fazem parte dessas cadeias mostra que os desconhecidos são muito numerosos, e as tentativas de inferência, em sua maioria, mal-sucedidas. Dessa forma, a escolha dos itens contextualmente relevantes pode ter sido causada por uma competência lexical inadequada. Como conseqüência das limitações acima, o conhecimento acumulado do leitor também é prejudicado, uma vez que seus sistemas de valores não conseguem ser efetivamente ativados, levando o sujeito a uma atitude passiva diante do texto. Dessa forma, nem a interação semântica e muito menos a pragmática, parecem ocorrer.

5.3 – Discussão da análise dos protocolos

A discussão dos resultados da análise, conduzida na seção anterior, tem por objetivo responder à pergunta geral de pesquisa colocada no início deste capítulo, que busca caracterizar a relação entre competência lexical e compreensão enquanto processo da leitura em LE. Essa relação, já evidenciada através do estudo de produto, apresentado no capítulo anterior, não é apenas confirmada, mas complementada através deste estudo. Os resultados, nesse caso, parecem mostrar que o número de problemas encontrados pelos leitores no texto, a frequência e a natureza dos recursos mobilizados para solucioná-los, as modalidades de processamento da informação, o desempenho desses sujeitos enquanto processo ou seus padrões de leitura, assim como o resultado dessa compreensão, podem ser explicados, em grande parte, pelo seu nível de competência lexical. Não uma competência lexical definida de forma estreita, que considera apenas os significados das palavras, ou o léxico separado da gramática e de outros níveis, mas uma competência lexical definida de forma ampla e abrangente, incorporando os vários níveis (pragmático-discursivo, lexical, semântico, sintático, morfológico, fonológico), além de uma dimensão cognitiva ou de uso.

Todas as pausas (com exceção de uma) identificadas nos processos de construção de sentido dos leitores deste estudo, assim como outros problemas, não marcados por pausas, são

detonados por dificuldades que, na superfície podem ser considerados lexicais, ou atribuídas ao desconhecimento de palavras, mas que uma análise mais detalhada mostra insuficiências ou dificuldades também em outros níveis (fonológico, morfológico, sintático, semântico e pragmático-discursivo), exigindo do leitor, para sua solução, vista na perspectiva da construção do sentido, recursos, portanto, que vão além do significado da palavra. Em outras palavras, na perspectiva das soluções, poder-se-ia dizer que a efetiva atualização de significados potenciais em uma situação de uso (nível discursivo) parece exigir adaptações e relações em níveis que não se restringem ao lexical, ou apenas ao significado de palavras isoladas.

A saliência do léxico, portanto, com relação aos outros níveis, poderia explicar porque o vocabulário tem um *valor de face* (2) tão grande na aprendizagem da leitura e de línguas em geral, e porque, ao mesmo tempo, muitos estudos que se propõem a investigar a força dessa relação, apresentando aos sujeitos significados de palavras desconhecidas após uma primeira leitura do texto, não conseguem mostrar melhoras sensíveis na compreensão.

Três modalidades de processamento ocorrem nos protocolos dos três leitores:

- Processamento automático, não marcado por pausas, caracterizado pelo reconhecimento de palavras e construção do sentido do texto, que ocorre através de uma interação, poder-se-ia dizer perfeita e automática, entre pistas e fontes de conhecimento. Há, nesse caso, fluência de acesso aos significados. Esse tipo de processamento, portanto, não depende apenas de competência lingüística bem desenvolvida, mas também de conhecimento prévio, assim como da percepção de pistas para a ativação das várias fontes de conhecimento e construção de um contexto. Em outras palavras, está relacionado ao uso automático e eficiente de recursos, contrariando, o que é muitas vezes afirmado na literatura, de que esse tipo de processamento é dependente apenas de competência lingüística ou seja, é "independente de contexto". A função do contexto, contudo, nesse caso, é distinta dos outros processamentos, uma vez que os significados das palavras são já estabelecidos;
- Processamento semi-controlado, marcado por pausas, caracterizado pelo reconhecimento de palavras e construção de sentido, que se utiliza das várias pistas e fontes de conhecimento, ou seja, do contexto para a ativação de significados já conhecidos (pelo menos a nível receptivo). Nesse caso, o contexto é utilizado para orientar decisões, por exemplo, com relação ao significado mais adequado àquela situação, quando a palavra apresenta vários significados; e, finalmente,
- Processamento controlado, marcado por pausas, em que as pistas do texto ativam as várias fontes de conhecimento para o acesso aos significados das palavras que, nesse caso, contrariamente ao que ocorre no processamento anterior, são significados desconhecidos. Este processamento é o que tem sido identificado na literatura como "através do contexto", envolvendo testagem de hipótese, adivinhações, dentre outros.

Apesar de a função do contexto diferir nos três processamentos, ela está presente em todos

eles, no momento de efetiva atualização de significados potenciais, e na construção de um sentido para o texto. Embora essas três modalidades de processamento possam ser observadas nos protocolos dos três leitores, devo salientar que a frequência em que ocorrem são diferentes, como mostro mais adiante. Além disso, apesar de as três modalidades terem sido identificadas nos dados, ressalto que a divisão não é rígida, podendo ser mais bem caracterizada como um contínuo, que tem nos seus polos o processamento automático e controlado.

Respondendo à primeira pergunta de pesquisa deste estudo, ou seja,

1. Tomando por base a categorização de recursos bilaterais do leitor proposta no Capítulo II (onde recurso é definido como a interação entre fontes/níveis de conhecimento (incluindo o prévio ou de mundo) e as restrições do texto, resultando no estabelecimento do contexto do texto e do contexto geral), quais são os recursos efetivamente utilizados pelos leitores/sujeitos para o reconhecimento de palavras e construção de um sentido para o texto ?

posso afirmar que a categorização, elaborada no Capítulo II, é adequada para caracterizar os recursos utilizados pelos leitores durante sua interação com o texto, embora nem todos os recursos dessa categorização tenham sido efetivamente utilizados. Não poderia, entretanto, discutir esses recursos sem fazer referência aos leitores, pois, embora os mesmos recursos possam ser identificados nos protocolos dos três, eles diferem em relação à frequência com que são utilizados, sendo alguns deles mais frequentes em um comportamento mais proficiente, que parece caracterizar o LIO, outras mais frequentes no comportamento menos proficiente dos outros dois leitores. Conseqüentemente, a eficiência no uso desses recursos também difere, uma vez que mesmos recursos são usados com resultados positivos e negativos pelos três leitores. Além disso, problemas iguais são solucionados pelos três leitores através de recursos diferentes.

Os recursos lexicais foram frequentes nos protocolos dos três leitores. Observa-se um grande número de tentativas de inferência/inferências de palavras, ou de acesso aos significados, e conseqüente construção do sentido do texto com esses significados, através da forma e do som do item detonador.

O uso do som, ou de um recurso fonético para ativação ou acesso ao significado, é utilizado inúmeras vezes pelo LIA e também pelo LIE; a palavra é, nesse caso, pronunciada várias vezes em voz alta, como se o som acessasse um significado que, nesse caso, não parece totalmente estabelecido, num processamento que denominei acima de semi-controlado. Enquanto as tentativas do LIA são em geral bem-sucedidas, as do LIE são mal-sucedidas, uma vez que o significado da palavra em LE é, na maior parte das vezes, desconhecido.

Um recurso idiossincrático do LIE é a tentativa de acessar significados de verbos através da pronúncia em voz alta de suas três partes (presente, passado e particípio passado). Nesse caso, entretanto, o conhecimento sintático não parece apenas o foco, mas também o som das três partes

do verbo em uma repetição mecânica, que poderia refletir a forma como essa palavra foi ensinada ou à abordagem de ensino a que o sujeito/leitor foi exposto. As tentativas de inferência através desse recurso nem sempre foram bem-sucedidas, uma vez que a competência lexical do leitor é particularmente vaga e imprecisa.

Outras vezes, a busca do significado é realizada através da forma da palavra. Nessas situações, o conhecimento vago de formação lexical, além de conhecimento semântico insuficiente levam, principalmente os LIO e LIE, a confusões entre palavras que têm formas semelhantes ou que Laufer (1990) denomina *sinformes*, como por exemplo entre *dairy* and *daily*, *without* e *which*, *damage* e *manage*, *run* e *rain*, *health* e *heat*, dentre muitas outras.

Todos os recursos lexicais acima mencionados procuram acessar o significado através da ativação de conhecimento na LE, e podem ser, portanto, denominados de recursos lexicais intralingüísticos. Entretanto, também podem ser observados recursos lexicais que são intermediados pela LM, no caso, o português, ou um recurso interlingüístico. A tendência de utilização de competência lingüística ou lexical da LM, mais especificamente, do som ou da forma da palavra em LE para estabelecer relações com palavras que se assemelham em forma ou som às palavras da LM, é muito comum nos protocolos principalmente dos LIO e LIE. Esse tipo de associação, ou utilização da "noção de cognaticidade" é bem sucedida em alguns casos, como por exemplo *ecological*, *Soviet*, *economy*, *according*, *result*, *problems*, *industry*, *technological*, *progress*, *market*, *mechanisms*, *use*, *rational*, *natural*, mas falha em outros. Nos casos malsucedidos, as associações são feitas com palavras na LM que não têm nenhuma relação semântica, apenas semelhanças formais com as palavras da LE, como por exemplo *reasons* por *reações*, *direct* por *direito*, *consumption* por *concepção*. Falsos cognatos entrariam nessa categoria.

Contudo, também podem-se observar exemplos, nos protocolos dos três leitores, em que a associação com a LM é feita através do som. Uma associação entre pronúncia incorreta da palavra e resultado malsucedido indicam que a percepção de uma relação de "cognaticidade" pode ser dependente não apenas da forma mas também do som. As relações estabelecidas através do aspecto fonético, tanto intralingüísticas como interlingüísticas, portanto, poderiam sugerir um efeito facilitador para a subvocalização em LE, contrariamente ao que ocorre em LM e, conseqüentemente, a necessidade de uma atenção maior à pronúncia das palavras, aspecto praticamente ignorado nas abordagens instrumentais ou naquelas que consideram a leitura como uma habilidade independente.

Da mesma forma que a noção de cognaticidade foi, muitas vezes, superutilizada ou extrapolada para outras palavras, não foi outras vezes percebida. Como o reconhecimento da semelhança com o português é variável, ou dependente de outros fatores, talvez pudesse sugerir uma escala de "cognaticidade" em que se possa distinguir entre duas ou mais classes de cognatas; as "mais cognatas" seriam aquelas que, mesmo pronunciadas de forma inadequada, ou de acordo com a fonética do português, ainda seriam cognatas, enquanto as "menos cognatas", seriam aquelas que

dependeriam da pronúncia correta para serem identificadas.

Outro fator importante para o reconhecimento eficiente de cognatos seria uma competência lexical ou lingüística em LM mais desenvolvida. Insuficiências. Insuficiências lexicais em LM poderiam explicar significados como "profligação" (*profligate*), "consumição" (*consumption*), "esforçamento", "enfortemento", "enforçamento" (*enforcement*), encontrados nos protocolos do LIE.

Ainda outra possibilidade para explicar a relativa dificuldade/facilidade de cognatos não estaria apenas relacionada à forma da palavra ou a LM, mas, especificamente, às pistas percebidas pelo leitor e às restrições do texto, com base em sua competência lexical em LE, e, conseqüentemente, aos recursos mobilizados para inferi-las; em outras palavras, dependeriam do tipo de contexto em que ocorrem (*potencialmente rico*).

Todos esses aspectos que acabei de mencionar, relacionados à eficiência na utilização da noção de cognaticidade poderiam sugerir que a "facilidade" atribuída às palavras cognatas pode ser ilusória, e que seu uso, enquanto elemento facilitador na leitura, poderia ser restrito ao levantamento de hipóteses, confirmadas ou rejeitadas através de uma combinação com os outros recursos, tais como por exemplo, relações semânticas ou sintáticas ou de conhecimento prévio. Enquanto o primeiro tipo de comportamento ou a identificação de cognatas para acesso direto ao significado, através de hipóteses prematuras é uma característica dos leitores LIO e LIE, o segundo, ou a confirmação da hipótese inicial através de outros recursos é característico do LIA. Em outras palavras, os bons leitores mobilizam vários recursos ao mesmo tempo, fazendo uso de suas potencialidades de forma mais efetiva.

A utilização de recursos lexicais, baseados no som e na forma ou grafia das palavras, intermediados pela LM também pode ser observado em outros estudos, sempre caracterizando leitores menos proficientes (vide Haynes 1983; Laufer 1990, dentre outros), salientando a influência maior da LM nos estágios iniciais da aprendizagem de uma língua. Também podem sugerir a maneira como as palavras são memorizadas ou como o léxico é armazenado. Estudos sobre aquisição do vocabulário trazem alguma evidência nesse sentido. Nos estágios iniciais da aprendizagem, salienta Henning (1973), as memorização de palavras parecem se basear em parâmetros acústicos e formais, enquanto, nos estágios mais avançados, os parâmetros semânticos ou ligados aos significados são mais salientes. A tendência de associação com forma e som, poderia estar sendo incentivada pelas abordagens instrumentais, que enfatizam a necessidade de se reconhecer semelhanças com a LM e usá-las como elemento facilitador da compreensão em LM. Embora concorde que, em princípio, a noção seja útil sob o ponto de vista do ensino, principalmente da motivação do leitor nos primeiros estágios de aprendizagem, considero-as um tanto simplificadora, na medida em que é vista como a panacéia para todos os males da leitura, criando expectativas para o aluno/leitor que nem sempre se confirmam na prática, levando esses alunos a generalizações, que trazem confusões e conseqüentemente, um sentimento, muitas vezes, de frustração. Uma explicitação de suas restrições,

que vão além da apresentação de uma lista de falsos-cognatos, ou de suas possibilidades para o levantamento de hipóteses, a serem confirmadas através de outros níveis de conhecimento, parecem colocar essa noção numa perspectiva mais realista, com resultados, a meu ver, mais eficientes.

Um outro recurso utilizado pelos três leitores deste estudo diz respeito ao uso da morfologia, ou mais precisamente, da noção de prefixos, radical e sufixos através da análise da palavra. Como sufixos, em geral, marcam flexões e classes de palavras, enquanto prefixos introduzem mudanças nos significados, poder-se-ia dizer que esse recurso é lexical na medida em que apenas considera a palavra testada, e principalmente seu prefixo e radical como elementos de acesso ao significado, deixando de lado possíveis relações com outras palavras na oração. Entretanto, também poderia ser considerado um recurso sintático ou gramatical, quando for intermediado por uma relação gramatical ou de classe gramatical, acessada pelo sufixo, e, que nesse caso, envolveria a relação com outras palavras. Embora esse tipo de recurso tenha sido usado pelos três leitores, apenas o LIA é bem-sucedido. A diferença entre o sucesso/insucesso poderia ser explicada, uma vez que no caso do LIO, por exemplo, o conhecimento das partes da palavra era incompleto, pois conhecia o sufixo mas desconhecia o radical. Além disso, o leitor avançado parecia ter uma noção clara de quando usar esse recurso, ou em que situações ele poderia ser eficiente. O mesmo não acontece com o leitor LIE e LIO. Este último tenta se utilizar desse recurso para acessar o significado de verbos frasais, como por exemplo *handing over*, dentre outros. Nesse caso, a soma das partes não é igual ao todo, havendo, portanto, a necessidade de aprendizagem de um novo conceito. O desconhecimento das outras palavras ou a densidade de palavras desconhecidas em torno do verbo impediu o leitor de confirmar, através de outros recursos, a hipótese levantada a partir da análise da palavra ou vice-versa.

Enquanto recursos lexicais caracterizam o comportamento dos leitores LIO e LIE, recursos que envolvem outras palavras, ou relações da palavra dentro da oração, são mais frequentes no protocolo do LIA. Essas relações podem ser semânticas e sintáticas.

Exemplos de recursos sintáticos, ou baseados no comportamento sintático ou gramatical das palavras, não são muito frequentes nos protocolos dos LIO e LIE. O que se observa, nesse caso, são problemas causados pela falta de percepção das relações que as palavras estabelecem umas com as outras na oração, ou ainda, de percepção das restrições gramaticais ou das funções que certos sufixos, como por exemplo, *-ing* podem desempenhar além de gerúndio, ou seja, como adjetivo ou substantivo. Essas insuficiências aparecem refletidas em segmentações incorretas ou em confusões na identificação de sujeito e de sintagmas nominais, dentre outros. O LIA, por outro lado, embora muitas vezes de forma implícita, parece utilizar-se desses recursos, tais como classes e funções de palavras para a identificação de sintagmas nominais e de seus núcleos, como critério para mascaramentos ou apagamentos conscientes de palavras. A falta desses recursos, por outro lado, levam os LIO e LIE, a apagamentos ou inferências incorretas que introduzem lacunas e distorções

ao sentido. Em todos casos, os recursos se limitam a relações dentro da oração.

Recursos semânticos, que consideram porções menores do texto, como, por exemplo, relações próximas à palavra, de procedência ou de origem sinalizada pela preposição *from*, relações de causa-efeito ou ainda de exemplificação podem ser observadas no protocolo dos três leitores, porém, com mais frequência, no protocolo do LIA. Também relações semânticas que se sobrepõem às relações sintáticas acima identificadas, como por exemplo, entre verbo e objeto em *charge rate* (*cobrar taxa*). Essas relações dizem respeito à colocação ou associação sintagmática da palavra *cobrar* com relação ao seu objeto (*taxa*); ainda podem ser observadas associações entre as palavras de um sintagma, entre dois ou mais sintagmas ou entre porções maiores da oração. Outras vezes, entretanto, essas associações não são percebidas, como por exemplo entre *pulp* e *paper*, ou associações de *timber* com o termo superordenado *natural resources*. Alguns problemas, identificados nos protocolos dos três leitores, podem ser atribuídos à associação da palavra com apenas um significado. Por exemplo, a palavra *concerns*, como substantivo, no sintagma *industrial concerns* causa estranheza ao LIA, por desconhecer especificamente esse significado da palavra. A solução dada a esse problema por esse leitor, nesse caso, é o mascaramento da palavra, passando a focalizar o seu adjetivo *industrial*. Um outro exemplo, ainda no protocolo desse mesmo leitor, é um foco no adjetivo do sintagma nominal, ao invés do seu núcleo. O conhecimento de apenas um significado da palavra, como por exemplo, *led*, confundiu o LIO. É interessante notar que, nessa situação, o significado explicitado pelo leitor não foi o do verbo *lead*, mais básico ou *core*, como no caso da palavra *concerns* acima, mas um termo mais especializado, ou seja, o substantivo *led*, que é associado à "luzinha do *drive* do computador". Um outro exemplo, ainda relacionado ao mesmo problema, ou seja, à associação das palavras a apenas um significado, poderia explicar a falta de flexibilidade de ajuste do significado ao texto em construção. O LIO, por exemplo, conhece a palavra *boundary* como *contorno*, mas não consegue "ajustá-la" para *fronteiras* ou *limites*, que seria mais consistente com o sentido no texto. O conhecimento de um significado mais remoto ou mais técnico, sem o conhecimento do mais comum ou generalizável não parece incomum em situações de LE, dado o tipo de exposição que esses leitores têm da língua. A exposição prematura e formal à leitura, principalmente de textos especializados ou técnicos, a necessidade de uso do computador (deve-se lembrar que os programas eram, até pouco tempo, todos em inglês), distanciados de situações de vivência da língua, poderiam explicar esse tipo de inversão; nesse caso, o leitor aprende os conceitos em inglês, mas utiliza-se da fonética do português para pronunciá-las. Juntamente com essa situação, observa-se uma visão estreita de competência lexical, em que a palavra é vista como portadora de um único significado, não incentivando o leitor a buscar outras possibilidades para a confirmação de suas hipóteses. Esse tipo de visão da aprendizagem do léxico (associação entre duas palavras em duas línguas) é comum em abordagens de ensino de línguas em geral. O desconhecimento de outras palavras em torno do item detonador, ou a densidade de palavras desconhecidas, e,

conseqüentemente, a falta de conhecimento acumulado, dificultam a construção de contexto local ou mais geral.

Recursos semânticos, que se utilizam de trechos mais longos de discurso, e, até mesmo o texto como um todo, ou seja, recursos de conhecimento acumulado, não são encontrados nos protocolos dos LIO e LIE. Suas tentativas são, nesse caso, frustradas, uma vez que os trechos processados contêm muitas lacunas ou são muitas vezes distorcidos por inferências mal sucedidas. Por exemplo, a falta de conhecimento acumulado pode explicar a inferência malsucedida da palavra *clean* pelo leitor intermediário, mesmo depois de ter aparecido três vezes no texto. O conhecimento acumulado distorcido leva, conseqüentemente, a inferências distorcidas. O LIA, por sua vez, consegue mobilizar esse recurso com bastante freqüência. Suas inferências e construções de sentido locais parecem sempre guiadas ou confirmadas por um recurso de conhecimento acumulado, o que lhe permite estabelecer relações entre trechos maiores do texto, parágrafos ou subtópicos.

Recursos de conhecimento de mundo ou do assunto do texto, explicitados pelos leitores, são raros até mesmo no protocolo do LIA, apesar de alguns estudos, como por exemplo o de Haarstrup (op. cit.) e Scott (1990) colocarem esse conhecimento dentre os mais utilizados pelos seus sujeitos. Algumas hipóteses poderiam ser levantadas para explicar esses resultados. A primeira delas seria a falta de conhecimento específico do assunto; embora o tema geral *ecologia* fosse familiar e motivante para os alunos, deve-se salientar que está sendo abordado no contexto da União Soviética, além de estar relacionado a questões econômicas. Uma outra hipótese poderia ser, no caso dos LIO e LIE, a dificuldade de ativação desse conhecimento prévio, causada pelas inúmeras lacunas de sua competência lexical. Como já salientei no Capítulo II, a ativação desse conhecimento não parece tão fácil em LE como tem sido sugerida, uma vez que o conhecimento do assunto não implica em conhecimento das palavras, como em LM. A análise que considera os itens chaves do texto/itens relevantes do LIO mostra claramente que sua interação com o texto foi prejudicada, impedindo a compreensão e também uma interpretação pragmática ou uma leitura mais crítica. O LIA poderia estar usando esse conhecimento como parte de seus vários recursos, sem, entretanto, explicitá-lo, adequando o sentido do texto à realidade.

Recursos textuais, que se utilizam da organização ou macro-estrutura do texto ou de sua estrutura retórica, superestruturais, ou referentes à estrutura canônica de cada tipo de discurso, assim como gráficos e tipográficos, tais como letras maiúsculas, grifos, aspas (com exceção de uma letra maiúscula, salientada pelo LIE na identificação do autor do texto), não foram utilizados, de forma explícita, por nenhum dos leitores. Esses resultados são coincidentes com os de Olshavsky (op. cit.), em que a identificação da estrutura do texto não foi usada por nenhum dos leitores do seu estudo. Contudo, a continuidade de sentidos apresentada no protocolo do LIA, evidenciada através da recuperação de sentidos que construiu em parágrafos anteriores, poderia sugerir a consideração da organização do texto como um recurso para a construção do sentido.

Recursos pragmático-discursivos também não foram identificados. A ausência desses recursos, que, de certa forma, aparece refletida no produto da compreensão ou seja, no tipo de interpretação feita pelos leitores, contudo, poderia ser resultante de fatores diferentes ou ter explicações distintas para cada uma dos leitores, pelo menos para o LIA, de um lado, e LIO e LIE de outro. Esse aspecto será abordado na discussão a seguir, que visa responder às segunda e terceira perguntas de pesquisa:

2. Com base nesses vários recursos, como se caracteriza o padrão de leitura apresentado por cada um dos sujeitos?
3. Qual a relação entre o padrão de leitura de cada sujeito, sua competência lexical e sua compreensão enquanto produto?

Algumas das diferenças sensíveis entre o desempenho de cada um dos três sujeitos-leitores, em sua interação com o texto, mais marcantes, principalmente, entre o LIA, por um lado, e os outros dois leitores, de outro, já foram parcialmente demonstradas a partir da apresentação dos recursos utilizados. Enquanto, até agora, o foco foram os recursos, na discussão que inicio a seguir saliento os padrões de leitura de cada um dos sujeitos, delineados a partir dos recursos mobilizados para resolver problemas causados, em sua maioria, pelo nível de competência lexical apresentado, embora a influência de fatores interacionais ou pragmáticos do contexto de leitura, tais como atitudes e motivações do leitor, tipo de texto, propósito da leitura, dentre outros, não possa ser ignorada.

Antes de buscar relacionar, entretanto, competência lexical, recursos e produto da compreensão, considero importante caracterizar melhor a competência de cada um dos sujeitos, trazendo, para isso, quando relevante, alguns dados do seu Perfil I e II. Essa caracterização não só é importante para um melhor entendimento da relação entre esses aspectos, mas, também, para suprir uma lacuna que se identifica na área de leitura em LE. Embora os estudos em geral salientem que o conhecimento lingüístico ou mais especificamente lexical de leitores em LE seja limitado, não se tem dados na literatura mostrando em que sentido ele é limitado.

A competência lexical dos três sujeitos difere bastante, confirmando os dados obtidos através dos resultados dos testes de produto apresentados no Perfil II. Ela pode ser colocada em um contínuo, que têm os LIA e LIE em seus polos. A competência mais desenvolvida do LIA pode ser atribuída à exposição maior à língua, freqüentando inclusive escolas de línguas. Os dados do perfil II, ou resultados de testes, mostram escores altos de competência lexical e inferência de palavras. Essa competência é confirmada através pelos recursos utilizados e, como mostro a seguir, pelo padrão de leitura que apresenta, lhe permitindo inferir grande parte das palavras desconhecidas e usar esses significados na construção de um sentido para o texto em questão. A identificação, na auto-avaliação, do vocabulário como dificuldade na leitura se confirma, na medida em que, realmente, o

leitor desconhece palavras de frequência mais baixa. Entretanto, o padrão de leitura apresentado por esse leitor mostra que as dificuldades são facilmente superadas, refletindo, portanto, um leitor exigente com seu desempenho.

Os dados de Perfil II do LIO, por sua vez, revelam que, apesar de apresentar um escore de conhecimento de vocabulário acima da média, ele é baixo, ou seja, é pelo menos metade daquele apresentado pelo LIO; no teste de inferência, seu escore ficou abaixo da média. Esses resultados são confirmados pelo padrão de leitura exibido, mostrando um grande número de palavras desconhecidas, e conseqüente dificuldade de inferência e de construção do sentido do texto. Esses resultados confirma, mais uma vez, sua auto-avaliação, em que reconhece apresentar problemas com o vocabulário.

Finalmente, os escores bastante baixos ou abaixo da média em conhecimento de vocabulário e inferência do LIE confirmam-se, mais uma vez no padrão de leitura exibido. As dificuldades com relação ao vocabulário salientadas em sua auto-avaliação são também confirmadas.

Um dos traços salientes da competência vaga e imprecisa apresentada pelos LIO e LIE, e principalmente LIE, é a dificuldade com verbos, principalmente aqueles que não aparecem no texto em sua forma infinitiva. O LIE tenta, constantemente, acessar esses significados, mas suas tentativas são, na maioria das vezes, malsucedidas. O desconhecimento dos verbos torna difícil a segmentação, necessária para a leitura em blocos maiores de significados, contribuindo para a leitura palavra por palavra. O LIO e LIE mostram desconhecer verbos tais como *reach, lead, ban, charge, maintain, damage, become*, dentre outros. Para o LIE, a lista ainda inclui *introduce, get, give, monitor, contain, begin*. Verbos frasais, tais como *clean up, set up, move away, step up, hand over, take steps, shut down*, são particularmente difíceis para esses dois sujeitos. As tentativas de solução malsucedidas mostram que a combinação das duas partes não foram suficientes para o reconhecimento de seus significados. Em todos esses exemplos, a presença de outras palavras desconhecidas na mesma oração impediram a rejeição dos sentidos errôneos.

Uma análise da natureza das palavras desconhecidas mostra que, enquanto o LIA desconhece palavras que poderiam ser classificadas como de média e baixa frequência (*ills, timber, progligate*), dentre outras, pouco numerosas no texto-base, os leitores intermediário e iniciante, além dessas, demonstram imprecisão também no reconhecimento de palavras de alta frequência, básicas, tais como *bring, talk, reach, clean without, every, unit, times, should*, além de muitas outras, como *set, take*, que entrariam na formação de verbos frasais. Além disso, enquanto para o LIA as palavras desconhecidas são substantivos, verbos e adjetivos, para os LIO e LIE incluem também palavras de função, como preposições, advérbios, quantificadores, dentre outros. A falta de palavras básicas, ou de um vocabulário que pode ser denominado *procedimental (procedural)* dificulta enormemente a tarefa de inferência e de construção do sentido desses leitores, uma vez que essas palavras básicas são difíceis de serem inferidas, pois não estão relacionadas a um assunto ou conteúdo específico.

É interessante notar, pelos comentários em seu protocolo, que os leitores, principalmente o LIO, têm uma noção clara das palavras desconhecidas/conhecidas. Uma escala de familiaridade poderia ser deduzida a partir de alguns trechos do protocolo do LIO a seguir:

"*reach* eu não sei o que que é..."; "*unscheduled, Swansea* também não sei"; "Bom eu sei o que é ... *clean*. *Clean* se não me engano é *apagar* Agora *clean up* é *pra cima apagar pra cima* eu não entendi"; "Eu não entendi muito bem o que significa *joint venture* tá?"; Bom eu deparei com uma palavra que eu não conheço é *announced*(...) não sei o que que é"; "Bom no próximo parágrafo que eu tô lendo aqui eu não entendi qui o que é a palavra /*jut*/ *shut* .. é *shut down* certo? (... ...)"; "Uma palavra aqui também eu bom ... também eu (...) Bom, eu já vi essa palavra mas eu não lembro o significado dela, que é /*Bout*/ *both* certo? E tem a palavra seguinte que é também /*aBRoud*/ *abroad* que eu também não sei o que que é, nunca vi na vida (... ...)"; "Tem outra palavra aqui também que eu não entendi nunca vi na vida: *entrepreneurial* certo? Não sei o que que é (...); "Eu não sei o que quer dizer *bring* também esqueci Eu conheço essa palavra mas esqueci o significado."

Esses diversos graus de familiaridade das palavras sugerem que o conhecimento de significados de palavras pode ser representado por uma escala de familiaridade, confirmando, através dos dados de processo, a validade de uma escala ou lista de verificação (*checklist*) como aquela usada no capítulo anterior, enquanto instrumento para avaliação do vocabulário.

A todas essas distinções qualitativas ou referentes à natureza das palavras desconhecidas desses leitores, podem-se acrescentar diferenciações quantitativas ou de número de palavras como um elemento definidor dos padrões de leitura dos sujeitos/leitores. Embora a ênfase do trabalho apresentado neste capítulo seja qualitativa, considero importante destacar, mesmo que rapidamente, alguns aspectos quantitativos, uma vez que, nos dados, os dois aspectos aparecem por vezes entrelaçados (vide Olshavsky 1976-7). Dois parâmetros, portanto, são importantes para definir ou caracterizar a competência lexical dos sujeitos: parâmetros que levam em conta sua extensão ou quantidade (número de palavras desconhecidas) e sua qualidade ou profundidade.

A diferença no número de pausas identificadas por cada um dos leitores é bastante significativa. As Tabelas 5.3 e 5.4 abaixo apresenta as pausas e soluções de cada sujeito, assim como aquelas comuns aos três sujeitos. Convém salientar que são consideradas soluções, nesse caso, inferências bem-sucedidas, o que não necessariamente implica em construção do sentido.

Tabela 5.3 – Pausas/problemas e soluções dos sujeitos

Leitores	Pausas	Soluções
LIA	23	14 (61%)
LIO	81	11 (13%)
LIE	96	15 (16%)

Tabela 5.4 – Pausas/problemas comuns aos sujeitos e soluções

Leitores	Pausas	Soluções
LIA	23	14 (61%)
LIO	23	5 (22%)
LIE	23	4 (17%)

Tendo caracterizado a competência lexical dos leitores como determinante na escolha dos recursos utilizados, passo a seguir a delinear e comparar seus padrões de leitura.

Três padrões de leitura distintos podem ser identificados, estando o do LIO mais próximo ao do LIE. Embora os padrões desses leitores não sejam iguais, apresentam semelhanças em muitos aspectos, sendo, a mais saliente, a dificuldade de inferir palavras e de processar trechos longos, de forma automatizada, uma vez que a construção do sentido é sempre truncada por uma palavra desconhecida.

O LIA lê praticamente o texto todo em voz alta, processando a maioria das palavras de forma automática. Sua pronúncia é excelente, com exceção de algumas poucas palavras desconhecidas, em que apresenta, na maioria, erros de acentuação. Ao encontrar palavras essas palavras, que algumas vezes repete em voz alta, destacando-as do resto do texto, desautomatiza o processo de leitura. Lê e relê trechos problemáticos (que em geral contém as palavras desconhecidas), segmentando-os corretamente e destacando-os dos demais, repetindo-os também em voz alta. Combina muitas vezes esse procedimento com a leitura e releitura de trechos anteriores ao trecho problemático, processando, para isso, trechos mais longos. Predições ou hipóteses levantadas na primeira leitura são confirmadas ou rejeitadas na segunda ou subseqüentes. Para facilitar a compreensão de alguns desses trechos mais difíceis, o leitor os divide em trechos menores, mantendo em inglês algumas partes enquanto processa outras. Parece monitorar sua compreensão de maneira eficiente e consistente e sua atitude, de forma geral, é analítica e de cooperação.

Além disso, esse leitor intermediário consegue mascarar palavras desconhecidas, ao percebê-las como não essenciais ao sentido do texto. Entretanto, isso só acontece depois de processar todas as palavras, em geral de forma automática, o que permite construção de um contexto local ou do texto, trazendo mais elementos para a inferência de outras palavras. Os exemplos de processamento não automático ou de compensação através da construção de um contexto de forma controlada ou semi-controlada são poucos ou apenas 23, como mostrei no quadro acima, considerando-se a extensão do texto. Nesses casos, embora apenas os recursos mais salientes sejam por vezes explicitados na análise, o leitor parece sempre combinar mais de um recurso, constituindo um traço distintivo do padrão de leitura desse sujeito. Embora o recurso decisivo ou explicitado tenha sido, por exemplo, a construção de um contexto imediato à palavra, observa-se que o conhecimento acumulado do texto até então também está sendo utilizado. A mobilização de mais de um recurso se confirma, no caso dos recursos lexicais, que têm a função de levantamento de hipóteses. Isso

significa dizer que esse leitor leva em consideração porções maiores do texto e não apenas a palavra. Esses resultados, ou seja, o uso de duas ou mais fontes de conhecimento paralelamente coincide com aqueles apresentados pelos sujeitos mais avançados do estudo de Hastrup (op. cit.).

O protocolo do LIA ainda parece mostrar que o reconhecimento automático de grande parte das palavras permite o estabelecimento de contexto, que é utilizado não apenas para a decodificação de palavras, mas principalmente para a construção do sentido, ou seja, para processamentos de nível mais alto. Em outras palavras, a economia de memória obtida através do reconhecimento automático aumenta a capacidade de processamento do leitor, que é usada para explorar outras pistas do texto. Além disso, depois de inferir uma palavra, seu significado é incorporado facilmente ao sentido em construção. Isso não acontece com o LIE e LIO, que, embora algumas vezes consigam inferir palavras, têm dificuldades para incorporar esses sentidos ao trecho em questão. As evidências parecem sugerir que, para esses dois leitores, há necessidade de uma fase de decodificação que precede a construção do sentido, enquanto para o LIA, a decodificação, mesmo a não-automática, é possível juntamente com a construção do sentido. O que os processamentos dos leitores LIO e LIE, por um lado e LIA de outro parecem mostrar, portanto, são relações diferentes entre inferência e compreensão. Em outras palavras, o que quero dizer é que enquanto para o LIA, num comportamento que poderia ser típico de leitores mais proficientes, a construção do sentido é o foco da leitura, e a inferência apenas vista enquanto um processo mediador, para o LIO, o acesso a significados ou a inferência de palavras de forma desautomatizada, pela falta de uma competência lexical mais desenvolvida, torna-se o foco de sua atenção imediata e seu propósito efetivo na leitura. Enquanto para o LIA a inferência e construção do sentido são parte de um mesmo processo, realizados ao mesmo tempo, para o LIO e LIE são processos que se realizam em etapas diferentes, que se sucedem. A inferência, nesse caso, antecede à construção do sentido. O leitor só conseguiria iniciar esse processo depois que tivesse chegado aos significados das palavras. A compreensão resultaria, portanto, do conjunto desses significados. Daí a busca incessante de significados que caracteriza seu padrão de leitura. Entretanto, o simples acesso aos significados não é suficiente para a construção do sentido; há necessidade, nesse caso, também de outros tipos de informação que a palavra traz consigo ou dos vários tipos de relações que podem estabelecer entre si. Dessa forma, apesar de os dois aspectos, inferência e compreensão, estarem relacionados, não significa que a habilidade de inferência possa garantir a compreensão.

A leitura do LIO, por sua vez, é silenciosa, mas pronuncia em voz alta, em inglês, trechos problemáticos, bastante numerosos nesse protocolo. Nota-se, nessas situações, que a pronúncia das nem sempre é fácil, dificultando o uso de pistas lexicais que têm como base a LM. Embora o LIO, ao encontrar trechos problemáticos, continue a leitura como um procedimento consciente para inferir palavras e construir o sentido desses trechos, não há procedimentos de volta e releitura do trecho problemático, como nos protocolos dos LIA e LIE. Embora o sujeito afirme desconhecer o

significado de muitas palavras, não apresenta hipóteses a respeito. A simples identificação das palavras-problemas, sem tentativa de inferência, é um procedimento usado com muita frequência por esse leitor (pelo menos 24 vezes, em todo o texto). Essas palavras tornam-se lacunas ao sentido do texto. O leitor parece, nesse caso, ter consciência de seus recursos ou de suas insuficiências na construção de um contexto para a decodificação e principalmente para a construção do sentido. A atitude do LIO perante ao texto é bastante objetiva, concentrando-se estritamente na "busca" de informações, como se estivesse desempenhando uma tarefa.

Nesse sentido, o padrão de leitura do LIO é diferente do LIE, pois, apesar de suas dificuldades e insucesso, tenta reler os trechos problemáticos, embora acabe se perdendo ao longo da leitura, sempre retomando o início do parágrafo toda vez que surge um problema. O leitor, nesse caso, não parece ter consciência de que seus recursos são muito limitados para a tarefa e insiste na decodificação palavra por palavra. A diferença entre os padrões de leitura desses dois sujeitos se reflete claramente na extensão de seus protocolos.

O padrão de leitura apresentado pelos LIO e LIE, portanto, mostra leitores dependentes da construção de um contexto para a decodificação de muitas palavras. Entretanto, o contexto para esses leitores muitas vezes não é muitas vezes estabelecido, ou construído de forma distorcida. A distorção dos contextos é resultado de um ciclo vicioso causado pela necessidade de inferir significados, resultando em hipóteses prematuras geralmente baseadas em pistas lexicais, usando em geral a forma da palavra e conhecimento da LM. A premência de significados forçam a formação de hipóteses e a pista mais saliente e disponível parece ser a da palavra e do conhecimento lingüístico em LM. Sem conhecimento lingüístico da LE que permita estabelecer comprovar ou rejeitar essas hipóteses, esses significados são mantidos na construção do sentido. Ao perceber que o significado não se encaixa no contexto, e apesar da dificuldade em se construir um significado com a palavra, o sujeito, entretanto, não consegue abandonar a hipótese. O comportamento é curioso. Ao invés de reformular sua hipótese, o leitor acaba distorcendo o sentido do texto em construção para encaixar a palavra ou apoiar a inferência incorreta. Em outras situações, a palavra não é inferida, introduzindo lacunas na construção do sentido. Essa dependência na forma da palavra também é identificada em Laufer & Sim (1985), Bensoussan & Laufer (1984) e Haynes (1984). Esse último autor mostra que, aprendizes de L2, em geral, deixam a forma da palavra desconhecida prevalecer sobre pistas sintáticas. Para Laufer & Sim (op.cit.) esse comportamento não apenas ocorre com aprendizes fracos, mas também com aprendizes melhores.

Como consequência desse comportamento, o conhecimento acumulado do texto resultante é vago e impreciso e muitas vezes distorcido. A inferência de outras palavras, por sua vez, vai se basear em um contexto também distorcido. Muitas vezes, as lacunas resultantes de inferências mal sucedidas ou simplesmente ignoradas são preenchidas por um significado baseado em pura adivinhação (*wild guessing*) do sentido da palavra.

As grandes diferenças, portanto, no padrão de leitura dos LIA, por um lado, e LIO e LIE de outro parecem ser o processamento automático, e, conseqüentemente, a liberação de capacidade de memória e processamento para o estabelecimento do contexto de forma controlada ou semi-controlado, e para processamentos de nível mais alto, que não se restringem à decodificação. Para o LIA, a construção do contexto é conseguida através de vários recursos diferentes (sintáticos, semânticos, fonológicos, lexicais, morfológicos, que envolvem fontes de conhecimento e as restrições do texto, e que variam em extensão, podendo abranger desde as palavras em torno do item detonador, como porções maiores da oração ou do parágrafo, ou ainda do texto como um todo, ou seu conhecimento acumulado.

Assim, enquanto o padrão de leitura do LIA se caracteriza, principalmente, pela construção do contexto através de recursos de natureza diversa, o dos LIO e LIE se caracteriza por lacunas e apagamentos, e, principalmente, pelo uso de recursos lexicais, e dificuldade de processamento de trechos longos, de forma automatizada, limitando a leitura a um processo de decodificação.

A análise e discussão conduzida parece mostrar bastante claramente a relação entre os padrões de leitura dos sujeitos, seu conhecimento lexical, e a construção do sentido. Um aspecto, entretanto, que chama a atenção em todos os protocolos, inclusive o do LIA, são pausas apenas para sinalizar problemas, não havendo, portanto, manifestações explícitas de julgamentos de valores e de posicionamentos, ou tentativas de negociação de significados. Em outras palavras, não se observa uma leitura mais crítica, essencial para o estudo acadêmico. Embora esse comportamento possa ser explicado nos protocolos dos LIO e LIE, uma vez que toda sua atenção estava voltada para a decodificação não-automática, não se justifica no protocolo do LIA.

A análise do produto da compreensão, através dos resumos no final de cada parágrafo, confirmam esses aspectos. Os resumos dos três leitores podem ser caracterizados como "traduções". Entretanto, há uma grande diferença entre eles. Enquanto as traduções do LIA demonstram compreensão, como já afirmei antes, aquelas apresentadas pelos outros dois leitores, apesar de algum sentido em trechos isolados, não revelam uma continuidade de sentidos que possa caracterizar a compreensão. As traduções do LIA podem ser consideradas, portanto, "traduções" bem-feitas enquanto os dos outros dois leitores apenas "tentativas de tradução", com muitas lacunas e distorções, sem dúvida alguma, refletindo o seu processo de leitura.

A interpretação do LIA é, entretanto, bastante próxima à do autor. Pode-se dizer que houve, nesse caso, uma interpretação semântica. No caso dos três leitores, entretanto, não há evidências do que Cavalcanti denomina interpretação pragmática, ou do que Moita Lopes denomina uma interpretação de nível pragmático.

A explicação para esse resultado, entretanto, não parece a mesma para os três leitores. No caso dos LIO e LIE, a falta de uma leitura mais crítica poderia ser explicada, pelo menos num primeiro momento, por uma competência lexical menos desenvolvida, que os impediu de identificar

os itens lexicais chaves e as cadeias coesivas. Para o LIA, a ausência de uma interpretação pragmática, entretanto, poderia ser atribuída aos fatores pragmáticos ou interacionais (Koch & Travaglia 1989) relacionados às condições de produção da leitura, que no arcabouço teórico proposto no Capítulo II, referem-se às restrições do texto (organização, ponto de vista, intenções, assunto) e restrições do contexto de leitura, ou o que Cavalcanti denomina variáveis de desempenho (fatores psicossociais do leitor, seu foco de atenção imediata, expectativas, atitudes e motivação) assim como o propósito da leitura. Esses elementos, que funcionam como molduras interpretativas, trazem para a leitura uma característica típica da interação humana, permitindo ou proporcionando condições para que haja uma negociação do significado entre os interlocutores (Moita Lopes 1986) através dos procedimentos de interpretação.

Como havia salientado no início deste capítulo, os leitores deste estudo tiveram como molduras para sua leitura *um artigo de assunto geral mas não desconhecido, leitura em inglês como LE, leitura no nível de idéia principal, leitura para coleta de dados de pesquisa*. As condições de leitura, portanto, determinadas por esses elementos, não podem deixar de ser consideradas, de certa forma, artificiais. Apesar da tentativa de trazer, para a situação de leitura como coleta de dados, as condições de uma situação normal, através de uma familiarização dos leitores com a tarefa de introspecção e de proporcionar-lhes condições de interação com o texto, uma certa artificialidade não pode ser evitada. Além disso, embora o texto fosse sobre um tema conhecido para os sujeitos, ou seja, a ecologia, a maneira como ele é abordada ou o tópico do texto, relacionado a problemas econômicos na (ex) União Soviética, pode não ter sido familiar e suficientemente motivante para o leitor, dificultando seu engajamento. Nesse sentido, a ausência de uma leitura mais crítica poderia ser explicada pela falta de motivação pelo tópico.

Uma outra explicação poderia ser a falta de um propósito mais específico de leitura, além da elaboração de um resumo. Não posso deixar de salientar que as interpretações de um resumo adequado podem variar (vide Scaramucci 199). Embora, para o professor, o nível de compreensão de um resumo seja a idéia principal ou central do texto, essa idéia principal pode variar em termos do que deve incluir. Para muitos, e pode ser o caso do leitor, essa visão de idéia principal em princípio pode não significar posicionamentos, mas uma interpretação o mais próximo possível do texto. Em outras palavras, pode refletir o tipo de concepção que esse aluno tem do que é um resumo. Ainda uma outra possibilidade de explicação poderia estar relacionada às próprias condições ou à natureza da tarefa. Elaborar um resumo no final de cada parágrafo pode, sem dúvida alguma, ter levado o leitor a uma leitura e compreensão lineares.

As atitudes do leitor com relação à leitura em geral e, especificamente, com respeito à leitura em LE, poderia oferecer mais uma explicação para uma interpretação que se restringe ao nível semântico. A predominância, na escola, de uma visão de leitura como decodificação, e a atitude do professor, sempre impondo sua interpretação como a única correta, pode ter levado os alunos a uma

postura passiva perante o texto, que os impede de se posicionar, e a uma concepção de leitura como extração do significado. Em LE, esse tipo de atitude parece ainda mais forte, pois se observa que uma leitura mais crítica ou mais profunda não é normalmente esperada em situações de LE. Quando menciono a visão de leitura do leitor como uma possibilidade de explicação para seu comportamento passivo, não estou me referindo às concepções *local* e *global* usada neste trabalho, inferida a partir do questionário metacognitivo, embora não deixem de estar, de certa forma, relacionados. Nesse caso, as concepções de leitura referem-se a processos internalizados da leitura, que são utilizados para explicar comportamentos como, por exemplo, o dos leitores LIO e LIE deste estudo, mais centrados nas palavras e no som, ou na decodificação, considerados, portanto, mais *locais*, ou o do LIA, mais voltado para o significado e, portanto, mais *global*, como mostram os dados do Perfil II, apresentados no início deste capítulo. Entretanto, não posso deixar de salientar que essa concepção de leitura, definida como *local* ou *global*, pode estar, na realidade, refletindo as dificuldades desses leitores e, não estarem sendo usadas como um dos fatores orientadores ou determinantes desse comportamento. Essa hipótese, entretanto, estaria sujeita a confirmação, através de estudos posteriores.

Finalmente alguns dados do Perfil I, apresentados no início deste capítulo, também podem ser esclarecedores nesse sentido. A falta de posicionamento crítico do leitor pode ser uma questão já percebida de forma consciente, uma vez que, na tarefa de auto-avaliação, afirma não se considerar um bom leitor, mesmo em LM. Por outro lado, esses dados de perfil também poderiam indicar um leitor exigente, que subestima seu desempenho, e, portanto, também inseguro em seu comportamento enquanto leitor. Essa insegurança poderia impedi-lo de ser mais assertivo, incentivando sua atitude passiva perante o texto.

A resposta afirmativa à quarta e última pergunta de pesquisa

4. Há evidências de um possível nível limiar de lexical para a inferência de palavras e para o processo de compreensão?

já parece previsível a partir da discussão conduzida até agora. Todas as evidências parecem mostrar que, a competência lexical mais desenvolvida, tanto qualitativamente, ou com relação à sua profundidade, como quantitativamente, ou relativa ao número de palavras conhecidas e natureza dessas palavras (principalmente palavras básicas ou de alta frequência), como aquela apresentada pelo LIA deste estudo, permite uma maior sensibilidade às restrições semânticas, sintáticas, fonológicas e morfológicas do texto, possibilitando a construção de um contexto de forma automática de grande parte das palavras, e também a construção de um contexto de forma controlada ou semi-controlada, quando necessário. A importância de palavras básicas está, principalmente, em servir como conhecimento *procedimental*, permitindo o acesso a palavras de baixa frequência. Esses dois aspectos de uma competência lexical mais desenvolvida, portanto, oferecem ao leitor mais proficiente

deste estudo a possibilidade de usar sua capacidade de processamento não apenas para a decodificação, mas principalmente para a construção do sentido ou processamentos de nível mais alto.

Leitores com uma competência lexical menos desenvolvida, por outro lado, apresentam sérias dificuldades de processamento automático, tendo que construir um contexto de forma controlada para a inferência de cada palavra. Há grandes dificuldades também na construção desse contexto, uma vez que sua competência lexical insuficiente não oferece outros recursos, além de lexicais, ou que levam em conta apenas as restrições de forma ou som das palavras. Além disso, o grande número de palavras desconhecidas torna necessário, com muita frequência, o estabelecimento de contexto de forma controlada para a inferência dessas palavras, tomando toda sua capacidade de processamento. As limitações na atenção humana e capacidade de processamento da memória, ou as demandas cognitivas exigidas do leitor com uma competência lexical menos desenvolvida poderiam explicar porque bons leitores em LM não conseguem, freqüentemente, aplicar suas habilidades de leitura a textos em L2 ou LE, ou porque uma compensação com uma competência em LM nem sempre é possível.

Embora o presente estudo não permita uma caracterização mais precisa desse limiar (o que, a meu ver, seria um objetivo difícil de ser alcançado senão impossível de ser alcançado, uma vez que esse limiar parece oscilar), oferece algumas sugestões a respeito de suas características. Poder-se-ia dizer que esse limiar não é absoluto, mas relativo ou dinâmico, uma vez que algumas inferências e algum sentido foram possíveis, embora com prejuízos para a compreensão. Em princípio, esse limiar poderia ser caracterizado em termos quantitativos, consistindo num conjunto de palavras básicas ou ainda palavras de alta frequência; e em termos qualitativos, relativos à qualidade desse conhecimento.

Da mesma forma que se pode identificar um limiar de competência lexical para a construção do sentido ou para a compreensão, também poder-se-ia inferir um limiar para a inferência, embora não haja elementos suficientes para se afirmar que o limiar seja o mesmo. Esses resultados não confirmam os de Scott (1990), uma vez que o autor não encontra evidências para um nível limiar de conhecimento de vocabulário que impeça a inferência. As evidências apresentadas, portanto, sugerem que mesmo sujeitos com vocabulário bastante limitado em L2, são razoavelmente bem sucedidos na inferência de palavras desconhecidas. Isso implicaria, para o autor, na necessidade, para a maioria das inferências, de muito menos contexto do que se poderia imaginar, isto é, um contexto restrito à algumas palavras de cada lado do item alvo. Não se pode esquecer, no entanto, que no estudo de Scott (op. cit.), a inferência é o foco, não a compreensão. Embora porções menores de texto sejam suficientes para a inferência, podem não o ser para a compreensão ou para a construção do sentido, o que poderia sugerir, portanto, limiares diferentes para os dois processos.

Este capítulo apresentou e discutiu a análise dos dados resultantes da investigação que

focalizou o vocabulário e a compreensão enquanto processo da leitura. O próximo capítulo traz uma discussão abrangente dos resultados dos estudos de produto e de processo, antes de apresentar as conclusões, limitações e contribuições teóricas, metodológicas e práticas desta tese.

Notas

(1) As transcrições fonéticas foram incluídas nos protocolos apenas para salientar pronúncias incorretas, em algumas situações relevantes para a análise e discussão. No entanto, muitas outras palavras, além das transcritas, foram pronunciadas incorretamente.

(2) O termo *valor de face* está sendo usado para *face validity*, na falta de um termo melhor em português. Outras possibilidades seriam *validade de face*, *validade de fachada* ou *de aparência*. Esse conceito, geralmente usado na área de avaliação, como um parâmetro de julgamento de testes, refere-se aqui ao vocabulário, ou à característica que tem de ser percebido pelos alunos, professores e pessoas em geral, como um aspecto importante da aprendizagem de uma língua.

(3) Emprego o termo "tradução" da forma como é vista por um leigo, apesar de estar ciente de que não implica numa interpretação apenas semântica.

CAPÍTULO VI

CONCLUSÕES

Este capítulo está subdividido em quatro seções. Na primeira delas (6.1), apresento um resumo das conclusões, que serve de introdução para a (6.2), a discussão comparativa dos resultados dos estudos apresentados nos Capítulos IV e V. A terceira seção (6.3) aponta as limitações do estudo, enquanto a última (6.4) salienta suas contribuições teóricas, metodológicas e práticas.

6.1 – Resumo das conclusões

Este estudo teve como objetivo principal investigar a relação entre a compreensão em leitura e a competência lexical ou o papel do vocabulário na compreensão em leitura em LE. Dada a natureza do problema, o estudo é investigado através de dois estudos diferentes. O primeiro tem como foco a compreensão enquanto produto da leitura a partir de uma análise predominantemente quantitativa. Foram sujeitos desse estudo 49 leitores de inglês como LE de níveis diferentes de proficiência em leitura. O segundo tem como foco a compreensão enquanto processo da leitura a partir de uma análise de natureza predominantemente qualitativa. Foram sujeitos desse estudo 3 leitores de níveis intermediário avançado, intermediário e iniciante em proficiência em leitura, selecionados dentre os sujeitos do primeiro estudo.

Os resultados dos dois estudos parecem oferecer evidências da importância de uma competência lexical bem desenvolvida como um fator determinante na leitura. Embora outros fatores tais como conhecimento conceitual, proficiência em leitura em LM e diferentes concepções em leitura possam ser considerados parte dessa relação, ou desempenhar um papel na proficiência em leitura, seu uso eficiente parece estar condicionado à competência lexical. Em outras palavras, para que a interação desses fatores possa ocorrer de forma eficiente, ou para que uma compensação possa ocorrer através deles, é necessário um nível limiar de competência lexical que envolva aspectos qualitativos e quantitativos. Os aspectos qualitativos vão além da consideração de significado da palavra, isto é, extrapolam o nível lexical, envolvendo a capacidade de uso desse conhecimento em outros níveis tais como semântico, sintático, fonológico, morfológico e pragmático-discursivo, permitindo não apenas a inferência de palavras mas, principalmente, a construção de um sentido para o texto.

Os aspectos quantitativos dessa competência, por sua vez, dizem respeito à sua extensão,

ou ao número de palavras desconhecidas, assim como à natureza dessas palavras. Um conjunto de palavras básicas de alta frequência ou de uso generalizável parece ser um elemento distintivo dessa competência, fundamental para que a busca de significados de outras palavras de baixa frequência possa ocorrer.

Uma competência lexical assim definida, entretanto, embora necessária, não é suficiente para que uma interpretação pragmática ou uma leitura crítica possa ocorrer. Nesse caso, os fatores pragmáticos ou seja, relacionados ao contexto de leitura, principalmente as atitudes e motivações do leitor, o tipo de texto e o propósito da leitura têm um papel fundamental, promovendo o engajamento do leitor com o texto e permitindo não apenas a interação, mas a negociação de significados.

6.2 – Discussão comparativa

A discussão dos resultados gerais deste estudo ou, mais especificamente, das análises conduzidas nos Capítulos I e V será orientada pelas perguntas de pesquisa levantadas no Capítulo I.

Em relação à primeira delas,

1. Como se caracteriza a relação entre a competência lexical e a compreensão enquanto produto e processo da leitura? Ou seja, qual o papel do vocabulário na construção do sentido do texto e, conseqüentemente, na compreensão?

poder-se-ia afirmar que os resultados dos dois estudos apontam para um papel determinante do vocabulário na compreensão vista enquanto produto e processo da leitura. Enquanto o estudo de produto, considerando o grupo de 49 sujeitos como um todo, focalizando os resultados de testes, oferece evidências que mostram a força dessa relação, o estudo de processo, considerando apenas três sujeitos, apresenta resultados que permitem caracterizar melhor a natureza dessa relação.

As evidências do estudo de produto foram obtidas principalmente através das correlações entre os escores obtidos nos testes de vocabulário/inferência e os escores dos testes de leitura em LE, confirmadas pelo resultado da análise fatorial, que mostra a existência de um componente lingüístico/lexical forte. Embora as correlações apresentadas sejam diferentes nas três medidas de leitura usadas neste estudo, isto é, teste do tipo cloze, escore do vestibular e teste de leitura de perguntas abertas/resumo, elas são consistentes nos três casos. Dentre essas três medidas, as correlações mais baixas foram apresentadas pelo teste de leitura com perguntas de respostas abertas/resumo. Essas correlações, entretanto, podem ser explicadas pelo fato de ter havido uma superestimativa do conhecimento lexical dos leitores na seleção de palavras a serem incluídas nos

testes de vocabulário.

O estudo com foco no produto oferece, ainda, dados que permitem caracterizar a competência lexical dos sujeitos, definida a partir de um conceito *rico* de vocabulário. Apesar de a análise ter demonstrado uma grande associação entre os aspectos ou subtestes que compõem esse conceito *rico*, tais como entre flexão/derivação, significado, colocações, dentre outros, observam-se diferenças de desempenho dos vários sujeitos nos vários subtestes, mostrando que alguns deles são mais difíceis que os outros, sugerindo um contínuo de dificuldade. Uma explicação para a relativa facilidade/dificuldade de alguns testes poderia estar no tipo de abordagem de ensino de LE a que esses sujeitos/alunos têm sido expostos. A facilidade relativa do teste de classes gramaticais, e a dificuldade do teste de significados adicionais, que vão além do significado mais básico ou generalizável, assim como dos testes de associação e colocação, já eram de certa forma previsíveis, uma vez que as abordagens de ensino de leitura e de vocabulário, em geral, enfatizam apenas o significado da palavra como o único aspecto importante, além de sugerir ao aluno que esse significado é único, focalizando, muitas vezes, as palavras isoladamente, sem considerar suas redes semânticas e possibilidades de colocações.

Um outro resultado importante do estudo de produto foi mostrar que a competência desses leitores não é apenas insuficiente sob o ponto de vista qualitativo ou de profundidade, envolvendo os vários aspectos do conceito *rico* de vocabulário, mas também no número de palavras e natureza das palavras desconhecidas. Esse último aspecto parece distinguir os dois subgrupos que podem ser identificados nesse estudo: um grupo, cujo conhecimento de vocabulário envolve as palavras avaliadas pelos testes de vocabulário; e um segundo, cujas palavras também inclui outras no texto. Isso significa dizer que, enquanto o primeiro grupo se caracteriza pelo conhecimento de palavras de frequência alta e desconhecimento de palavras de frequência média e baixa, o segundo se caracteriza também pelo desconhecimento de palavras de frequência alta ou seja, palavras básicas (*core words*).

O estudo de processo oferece evidências, através dos dados dos protocolos dos três leitores, que confirmam essa subdivisão, além de fornecer elementos adicionais para uma caracterização mais completa da competência lexical desses leitores. O leitor intermediário avançado poderia ser um exemplo do primeiro grupo, enquanto os leitores intermediário e iniciante pertenceriam ao segundo grupo. Além disso, o estudo de processo ainda complementa esses resultados através de um foco na dimensão de uso, mostrando os efeitos desses níveis diferentes de competência no desempenho (enquanto processo dos três leitores ou seu padrão de leitura), e, conseqüentemente, na compreensão. As várias confusões com formas de palavras "parecidas" nos protocolos dos leitores intermediário e iniciante, sugerem que, para esses leitores, muitas palavras são conhecidas apenas superficialmente, indicando que esses sujeitos têm uma idéia vaga sobre morfologia, por exemplo. Essa falta de uma competência mais desenvolvida, que, em geral, parece se limitar apenas à busca de significados de palavras, leva os leitores a dificuldades de inferência e, sobretudo, de

processamento, necessário para a construção do sentido do texto. Nesse caso, o processamento automático não ocorre e, como compensação, os leitores se vêem obrigados a tentar a decodificação de forma controlada, necessitando, para isso, do estabelecimento de um contexto, buscando, muitas vezes, nesse processo, não apenas seus recursos limitados em LE mas, sobretudo, recursos que envolvem a LM. Entretanto, a tentativa de compensação não ocorre de forma eficiente. Como o número de palavras desconhecidas é muito grande, essa decodificação mobiliza todos os recursos dos leitores, causando uma sobrecarga em sua capacidade de processamento, o que os impede de usar os recursos para a construção de um sentido para o texto ou para processamentos de nível mais alto.

A predominância de um processamento automático, característica marcante do leitor intermediário avançado, não significa, no entanto, que a função do contexto seja eliminada. Nesse caso, o que se pode observar é que não apenas os significados das palavras já são estabelecidos, mas os outros níveis de conhecimento também são desenvolvidos, a ponto de permitirem uma interação perfeita e automática entre as restrições do texto e as fontes de conhecimento do leitor. Os recursos resultantes dessa interação são usados, nesse caso, de maneira automática. Dessa forma, o contexto é criado de forma eficiente, permitindo não apenas o reconhecimento ou o acesso automático aos significados, mas, principalmente, a disponibilidade de espaço de memória ou de processamento para uso em outros níveis mais altos ou de construção do sentido propriamente dito, ou ainda quando houver necessidade de uso de contexto de forma controlada. Isso significa que o leitor mais proficiente deste estudo consegue usar o contexto de forma controlada, porque não precisa dele com muita frequência. Por outro lado, os outros dois leitores, embora tenham necessidade de se utilizar do contexto de forma controlada ou desautomatizada, não conseguem fazer uso dele com eficiência. Além disso, a competência lexical insuficiente dos leitores menos proficientes leva-os a usarem predominantemente recursos lexicais, baseando-se apenas na palavra ou item detonador, que, na maior parte das vezes, são provenientes de conhecimentos na LM ou recursos interlingüísticos, levando-os a hipóteses prematuras sobre os significados das palavras ou mesmo incorretos, sem que possam confirmar essas hipóteses através de outros recursos. O grande número de significados incorretos e ainda lacunas (uma vez que muitas palavras não são inferidas) leva-os a um conhecimento acumulado do texto que é, na maior parte das vezes, distorcido, e, por isso, não usado com eficiência na construção de um sentido para o texto. Dessa forma, a decodificação e a busca da compreensão acabam por se utilizarem exclusivamente de porções menores do texto, geralmente próximas às palavras desconhecidas ou itens detonadores.

Assim, enquanto o processamento controlado e o uso freqüente de recursos lexicais caracterizam leitores menos proficientes, o uso do processamento automático e de recursos que se utilizam de vários outros níveis de conhecimento (sintático, fonológico, semântico, morfológico) caracterizam leitores mais proficientes. Em outras palavras, isso significa dizer que, enquanto para

o estabelecimento de contexto, o leitor mais proficiente se utiliza de porções maiores do texto, que envolvem relações com outras palavras, o estabelecimento de contexto pelos leitores menos proficientes se limita à palavra ou item detonador. Não estou afirmando, entretanto, que não possam ser encontrados os mesmos tipos de recursos nos protocolos dos três leitores; embora isso ocorra, os resultados ou a eficiência nesse uso diferem bastante.

Dessa forma, não apenas a construção do sentido mas também o sucesso na inferência de palavras ou a decodificação semi-controlada ou controlada parecem dependentes de uma competência lexical mais desenvolvida, como mostram as correlações entre esses aspectos identificadas no estudo de produto e processo. A inferência poderia ser vista, nessa situação, como um elemento mediador, mas fortemente determinada pela competência lexical. Embora a inferência, como fator mediador, possa ser observada no estudo de processo, o que se observa é que nem sempre a inferência bem sucedida leva à construção de um sentido para o texto. Há a necessidade, nesse caso, de relações entre outros níveis dessa competência que vão além da inferência de palavras isoladas. Embora no protocolo do leitor avançado a inferência e a construção do sentido pareçam fazer parte de um mesmo processo, o mesmo não se pode dizer dos outros dois leitores, que parecem mostrar dois processos distintos.

Quanto à segunda pergunta desta pesquisa, isto é,

2. Qual a contribuição, para essa relação, sob o ponto de vista de processo e de produto, de fatores como proficiência em LE, conhecimento conceitual, leitura em LM assim como diferentes concepções de leitura em LE?

é importante lembrar que os dois estudos usaram procedimentos diferentes. No estudo de produto, a investigação consistiu na operacionalização desses fatores através de testes. As correlações entre esses componentes e as medidas de leitura em LE mostram, com exceção das concepções em leitura, associações relevantes, confirmando sua importância na leitura. Essas correlações, entretanto, não são tão altas como as de conhecimento de vocabulário/inferência, mostrando uma contribuição em menor escala desses fatores. Esses resultados são consistentes com aquele apresentado pela análise fatorial, que revela três fatores comuns às 21 variáveis, embora os Fatores 1 e 2 sejam mais explicativos. O primeiro deles, Fator 1, é o componente lingüístico, uma vez que apresenta o conhecimento de vocabulário e proficiência lingüística fortes. O Fator 2, representado pelos testes de leitura I e II e suas subpartes, é o componente de leitura. Ele envolve também as variáveis *concepções de leitura e conhecimento conceitual*. Apesar de baixa, a contribuição dessas variáveis ainda é maior para esse fator. O Fator 3, representado pelos testes de leitura em LM incluindo os escores em LM do vestibular, é o componente de LM. Esse fator é o mais fraco dentre eles, podendo ser considerado conjuntamente com o Fator 2, formando, praticamente, um único subgrupo: um componente de leitura, que engloba proficiência em leitura em LM, conhecimento prévio e diferentes

concepções de leitura. Com base nesses resultados, conclui-se que a compreensão em leitura é resultante de dois fatores, que contribuem com pesos diferentes: um fator lingüístico e um fator de leitura, que engloba proficiência em LM, concepção de leitura e conhecimento conceitual. Esses resultados confirmam o que já se pressupunha a respeito da natureza da leitura em LE, ou seja, que esta não envolve apenas fatores lingüísticos, mas também depende de outros componentes.

No estudo de processo, a relação entre o conhecimento lexical, a compreensão e os aspectos acima mencionados foi mostrada de duas maneiras. A primeira delas consistiu em utilizar os escores de conhecimento conceitual, proficiência em LM e concepção de leitura obtidos na primeira fase – ou o que denominei Perfil 2 – para uma caracterização mais precisa dos três leitores. Esses dados são confirmados a partir dos desempenhos ou padrões de leitura mostrados pelos três sujeitos. Uma segunda maneira foi investigar a interação dos fatores de conhecimento conceitual, de proficiência em LM e também de concepções diferentes de leitura sob o ponto de vista de processo propriamente dito, isto é, nos recursos utilizados pelos leitores na inferência de palavras e na busca da compreensão.

No estudo de produto, a relação entre o conhecimento conceitual e a compreensão é investigada, especialmente, através do cruzamento dessa variável com os testes de leitura I e II ou de perguntas abertas/resumo. De uma maneira geral, o grupo de sujeitos se mostrou bastante homogêneo com relação a esse conhecimento. Apesar de alguns sujeitos estarem nos extremos, ou seja, terem acertado todos os seis conceitos, ou um sujeito ter errado todos, esses casos são poucos em relação ao grupo como um todo, cujos acertos foram acima de 75%. Correlações positivas podem ser também observadas entre conhecimento conceitual e testes de leitura I ou com palavras inventadas, mostrando que a contribuição do conhecimento conceitual parece ter sido ligeiramente maior nesse caso. Entretanto, é necessário salientar que os sujeitos que apresentaram uma competência lexical falha ou insuficiente aparentemente não conseguiram compensá-la através do conhecimento conceitual, uma vez que a comparação entre os resumos e respostas às perguntas nos testes de leitura I e II, nesses casos, não mostram alterações.

No estudo de processo, os resultados não esclarecem melhor o papel do conhecimento prévio na leitura em LE. Há, nos dados dos protocolos, apenas uma evidência explícita de uso de conhecimento prévio do assunto ou do contexto geral, contrariando outras evidências na literatura, como por exemplo Scott (1990), que enfatiza a importância desse conhecimento na inferência de palavras. Diferentes hipóteses poderiam ser levantadas para explicar a falta de explicitação de uso desse conhecimento pelos sujeitos deste estudo. No caso dos leitores iniciante e intermediário, ele poderia estar associado à dificuldade de ativação dos esquemas de conhecimento prévio, causada pela competência lexical vaga e imprecisa. No caso do leitor intermediário avançado, entretanto, o fato poderia ser explicado pela não necessidade dessa ativação, uma vez que sua competência lexical é bastante boa e eficiente. Uma outra possibilidade, a me ver, mais plausível, seria o uso desse

conhecimento de maneira automática, não necessariamente explicitado enquanto parte de uma competência lexical, mas presente em todos os níveis dessa competência.

As correlações entre as medidas de proficiência em leitura em LM e LE, embora não tão fortes, mostram que essa variável pode ter uma influência na leitura em LE ou pode ser considerada como um fator de importância secundária da proficiência em leitura. Esses resultados também são confirmados pela análise fatorial. No estudo de processo a relação entre a proficiência em leitura em LM e LE aparece mais uma vez, principalmente quando se cruzam os perfis dos sujeitos e seu desempenho em leitura em LE. O leitor intermediário avançado, por exemplo, apresentou também um escore acima da média, em leitura em LM, enquanto os escores dos leitores intermediário e iniciante foram abaixo da média. Sob o ponto de vista do processo propriamente dito, ou dos recursos utilizados pelos leitores, pode-se dizer que a competência lingüística ou lexical em LM, é usada, várias vezes, como um recurso interlingüístico no acesso de significados às palavras na LE pelos três leitores/sujeitos desse estudo. Entretanto, enquanto para o leitor intermediário avançado, esses recursos são bem sucedidos, uma vez que os utiliza juntamente com outros recursos intralingüísticos, para os leitores intermediário e iniciante são mal sucedidos, levando-os à formação de hipóteses prematuras, não confirmadas através de outros recursos. Sua utilização, portanto, enquanto compensação para uma competência lexical menos desenvolvida não fica evidenciada. Em outras palavras, poder-se ia dizer que o uso eficiente de recursos interlingüísticos, no caso, de proficiência em LM, depende de uma competência lexical mais desenvolvida na língua alvo.

Considerando a última variável, ou seja, as diferentes concepções em leitura, inferidas a partir do questionário metacognitivo, pode-se observar, através do estudo de processo, que não mostram correlações relevantes com os escores de leitura. Várias hipóteses podem ser aventadas para explicar as correlações mais baixas das diferentes concepções em leitura. Em primeiro lugar, isso significa que tanto uma concepção *local* como *global* poderia estar associada ao sucesso na compreensão e não apenas uma concepção *global*, como a literatura parece reconhecer. Em segundo lugar, os escores que mostram a concepção de leitura, na realidade, não classificam o leitor como apresentando exclusivamente uma concepção ou outra, ou seja, totalmente *global* ou *local*, mas mostram um contínuo, isto é, os leitores são ora *globais* ora *locais*, combinando os dois tipos de estratégias. Uma terceira possibilidade de explicação estaria relacionada a uma limitação do instrumento de coleta, ou seja, do questionário metacognitivo, ou ainda da amostragem para a coleta. Com relação ao primeiro caso, observa-se que muitas das respostas, principalmente aquelas na categoria *neutro*, poderiam indicar estratégias de leitura não fixas, mas que podem variar dependendo do propósito da leitura e do tipo de texto, dentre outros. Nesse caso, o questionário não parece ser sensível a esse aspecto, uma vez que não considera essas diferentes possibilidades. Com relação à amostra, devo ressaltar que quando os dados foram coletados, todos os sujeitos (os de nível II de Inglês Instrumental mais, os de nível I menos), já haviam sido expostos às aulas de instrumental e a

uma abordagem que enfatizava o uso de estratégias, principalmente estratégias mais globais. Dessa forma, os dados já poderiam estar refletindo a influência dessas aulas e não mais apenas as concepções individuais de leitura e as estratégias realmente usadas por esses sujeitos. Uma outra explicação, levantada agora a partir do estudo de processo é a de que os sujeitos podem apresentar uma concepção *local* ou *global*, como consequência do tipo de dificuldades ou de necessidades apresentadas na leitura. Dessa forma, as dificuldades é que estariam determinando a concepção e não o contrário. Essa possibilidade, entretanto, teria que ser mais bem investigada.

Uma análise das estratégias metacognitivas denominadas facilitadoras, tais como *ler em blocos, sem focalizar cada uma das palavras, ou distinguir conscientemente palavras essenciais ao sentido do texto e palavras cujos sentido podem ser ignorados, ou ainda usar o contexto* mostram uma dependência muito maior em relação ao conhecimento lingüístico do que se quer admitir e muito maior do que alguns leitores em LE, como por exemplo, os de nível intermediário e fraco que apresento neste estudo, poderiam apresentar.

Devo salientar que os resultados discutidos até agora já permitem responder de forma positiva à terceira pergunta de pesquisa, ou seja,

3. Há evidências de um nível limiar de competência lexical para a compreensão enquanto processo e produto da leitura?

A investigação dessa questão em termos absolutos pressupõe uma avaliação do número de palavras conhecidas ou da extensão total do vocabulário do indivíduo, como salientei no Capítulo II, e foge do escopo deste trabalho. A possível existência de um nível mínimo de competência lexical é, neste estudo, investigada em termos relativos, isto é, considerando a competência do sujeito com relação às palavras do texto utilizado no teste de leitura, competência essa que é usada posteriormente para projetar generalizações sobre o seu léxico total.

No estudo de produto, as evidências para um nível limiar foram obtidas, principalmente, através do cruzamento dos escores de conhecimento de vocabulário, apresentados pelos sujeitos, e seus escores do teste de leitura II, ou seja, o teste com as palavras reais. Os resultados dessa análise mostram, de maneira satisfatória, a correspondência entre esses escores, ou seja, escores baixos em leitura correspondem a escores baixos em conhecimento de vocabulário. Entretanto, seis exceções, que se caracterizam pela falta de correspondência entre escores altos e escores baixos puderam ser observadas. Embora esses resultados possam ser explicados, em princípio, por algum fator externo à avaliação, podem também sugerir a interação de outros fatores e a possibilidade de compensação desse limiar através de outras fontes de conhecimento. Isso significa dizer, que embora se possa inferir a existência de um nível lexical de conhecimento de vocabulário para a compreensão, não se poderia afirmar que esse limiar é absoluto ou estático, mas sim relativo ou dinâmico, podendo oscilar dependendo do texto, propósito da leitura, ou ainda de diferentes concepções de leitura, níveis

diferentes de conhecimento conceitual e proficiência em LM, dentre outros. Uma evidência de que esse limiar é relativo, porque varia de acordo com a tarefa ou propósito da leitura, é obtida através de uma comparação do nível de competência lexical exigido na tarefa de resumo e na de pergunta de detalhe, incluídas no teste de leitura I e II, ou de perguntas de respostas abertas/resumo. Nesse caso, pode-se notar a necessidade de um nível mais alto de competência lexical para responder à pergunta de detalhes do que para elaborar o resumo. Entretanto, a possibilidade de compensação com outras fontes de conhecimento, tais como o conhecimento prévio do assunto ou conhecimento da LM, sem maiores prejuízos para a compreensão não parece se confirmar através dos resultados do estudo de processo. Apesar de algumas poucas inferências baseadas no conhecimento da LM e conhecimento prévio, os dados dos protocolos mostram que as tentativas de uso da competência lingüística em LM em geral, através de recursos interlingüísticos, não parece eficiente sem uma competência mais desenvolvida em LE, que permita verificar essas hipóteses, através de outros níveis ou recursos, como já salientei acima. A interação com os outros aspectos poderia ocorrer, efetivamente, depois que esse limiar fosse alcançado. Uma possível explicação para a diferença entre os dois estudos pode estar no texto, uma vez que no estudo de produto, ele é mais curto e aparentemente mais fácil, enquanto no de processo é mais longo e mais difícil, confirmando mais uma vez a existência de um limiar relativo.

Embora os resultados desta tese não permitam uma caracterização desse limiar em termos absolutos ou quantitativos, indicando *quais e quantas* seriam as palavras suficientes para sua composição, objetivo esse, a me ver, não apenas difícil de ser alcançado como também de utilidade relativa, uma vez que esse limiar pode oscilar dependendo do tipo de texto e propósito da leitura ou nível de compreensão almejado, oferecem algumas sugestões com relação às características desse limiar. Este parece incluir duas dimensões, uma quantitativa ou de número de palavras desconhecidas e, principalmente, uma dimensão qualitativa, isto é, de profundidade desse conhecimento. As palavras, nesse limiar, seriam, principalmente, palavras básicas de alta frequência, de uso generalizável. A segunda dimensão envolveria os vários aspectos de uma competência lexical definida em termos de um conhecimento *rico* de vocabulário, conforme a proposta deste estudo.

As correlações entre conhecimento de vocabulário e inferência apresentadas no estudo de produto, confirmando as associações encontradas através da AC, poderiam sugerir um nível limiar de competência lexical inclusive para a inferência de palavras. Algumas evidências para esse limiar também podem ser encontradas no estudo de processo, embora não possa se afirmar com segurança se o mesmo nível é necessário para os dois processos, ou se a construção do sentido exige um limiar mais alto. Em princípio, os resultados indicam que a inferência de palavras não necessariamente conduz à compreensão.

Além desses aspectos, o estudo de processo também mostra a inter-relação e a importância na leitura de fatores pragmáticos e interacionais, tais como a situação de leitura, seu propósito, o tipo

de texto assim como as atitudes do leitor. Embora o leitor intermediário avançado tenha conseguido construir um sentido para o texto, que pode ser interpretado como compreensão, observa-se que, nesse caso, o sentido foi construído de maneira linear, mantendo fidelidade aos elementos do texto. Esse resultado, portanto, poderia ser visto como uma interpretação apenas semântica, uma vez que todas as rupturas feitas em seu protocolo foram causadas por problemas e não para negociar os sentidos do texto, emitindo julgamentos de valor ou reafirmando posicionamentos a respeito do assunto. Em outras palavras, embora haja evidências de comunicação entre leitor/autor e de uma concordância intersubjetiva em sua leitura, não há, por outro lado, evidências explícitas de uma interpretação pragmática nos moldes propostos por Cavalcanti (1983, 1989) e Moita Lopes (1986). Os fatores interacionais ou da situação da leitura, com já salientei no Capítulo V, mais especificamente as atitudes do leitor, suas expectativas e interpretações com relação aos propósitos da leitura e o tipo de tarefa podem ter sido responsáveis por tal interpretação.

Os outros dois leitores, por sua vez, além de não terem chegado a uma interpretação pragmática, também não conseguiriam uma continuidade de sentidos que caracterizam a compreensão. Nesse caso a explicação estaria relacionada à sua competência lexical menos desenvolvida, que os impediria de perceber os itens lexicais-chaves e de identificar as cadeias coesivas encabeçadas por esses itens.

6.3 – Limitações

Para que as conclusões acima apresentadas assim como as contribuições a serem abordadas na Seção 6.3, possam ser interpretadas de maneira mais adequada, apresento, nesta seção, algumas limitações deste estudo, relacionadas, principalmente ao desenho de pesquisa e instrumentos usados.

A primeira delas diz respeito à generalização de resultados que transcenda à amostra estudada. Muitas hipóteses são levantadas neste trabalho, seguindo sua natureza exploratória, descritiva e observacional. Generalizações para outras situações teriam que ser feitas com reservas, ou apenas após sua confirmação através de outras investigações. Este estudo tem, portanto, validade externa limitada. Um elemento que poderia ser orientador dessas decisões, ou usado para adequar essas generalizações para outras situações de leitura em LE no Brasil, seriam os perfis detalhados dos leitores deste estudo que apresento no Capítulo V.

A segunda limitação se relaciona à natureza da relação entre vocabulário e a compreensão em leitura. Embora a importância da competência lexical para a compreensão tenha sido amplamente demonstrada, ela não pode ser caracterizada como uma relação de causa-efeito, uma vez que o estudo é exploratório, e não houve tentativa de se influenciar a compreensão com nenhum tipo de instrução ou tratamento dado ao vocabulário. Dessa forma, as implicações para uma situação de

ensino também devem ser consideradas com reserva. Entretanto, um aspecto deste estudo, que dá maior consistência e validação aos seus resultados, é o fato de ter incluído leitores com níveis diferentes de competência lexical e proficiência em leitura, permitindo a comparação das soluções dadas aos mesmos problemas pelos três leitores. Problemas não solucionados pelos leitores menos proficientes são efetivamente resolvidos pelo mais proficiente. As hipóteses levantadas através deste estudo, principalmente aquelas que se referem à competência lexical, baseadas em um conceito *rico* poderão vir a ser confirmadas, por exemplo, através da avaliação dos resultados de uma abordagem fundamentada nesse conceito, da forma como é sugerida nas *contribuições práticas* mais adiante.

Algumas limitações deste estudo ainda poderiam ser relacionadas aos instrumentos de coleta de dados, com foco na compreensão enquanto produto da leitura. Dada a natureza do estudo, a maioria dos instrumentos tiveram que ser elaborados. Apesar de muitos terem sido pré-testados, outros não o foram, e não se pode deixar de considerar a medida de erro que toda avaliação envolve. Além disso, uma validação "clássica" dos instrumentos também não foi realizada, uma vez que, dado o número de instrumentos necessários para o estudo, ela seria muito extensa, justificando um estudo à parte. Entretanto, esse aspecto foi minimizado pelo seu desenho abrangente; incluindo uma triangulação de métodos e diferentes técnicas estatísticas. Muitos dos testes, principalmente os de vocabulário, tiveram sua consistência medida através das correlações e associações da A, assim como pela análise fatorial. Os testes de leitura também tiveram sua consistência verificada através das correlações apresentadas entre si. Os dados de processo, além de terem sido validados pelos de produto e vice-versa, também foram confirmados através dos resumos ou dados de produto, eliciados no próprio estudo de processo. Além disso, para tentar minimizar a falta de familiaridade dos sujeitos com a tarefa, o que tem sido freqüentemente citado como uma limitação da técnica introspectiva, foi feito um aquecimento usando-se um outro texto.

Uma outra limitação está relacionada à complexidade da coleta de dados pelo fato de envolver várias sessões. Embora os testes tenham sido realizados em sessões diferentes, não se pode eliminar totalmente a hipótese de que os resultados possam ter sido influenciados por alguma variável externa à avaliação, tais como falta de motivação dos sujeitos ou pela ordem de apresentação dos testes. A aplicação do teste de leitura I, com as palavras inventadas, por exemplo, foi precedida pela aplicação do teste de vocabulário em contexto, com as palavras reais, uma vez que não se desejava que as perguntas do teste de leitura, em português, servissem de pistas para a inferência de certas palavras. Apesar de ter havido um intervalo de quase um mês entre as duas aplicações, não se pode deixar de considerar a possibilidade de que os sujeitos tenham se lembrado dos detalhes do texto. Essa hipótese é aventado, uma vez que os resultados, ou mais especificamente os escores de quatro sujeitos no teste de leitura I, não puderam ser totalmente explicados.

Dada a abrangência dos instrumentos, pode-se dizer que muitos dados efetivamente coletados não foram analisados em sua totalidade, podendo ser explorados mais detalhadamente em

pesquisas futuras. Os testes do tipo coze, tanto em LM como LE, analisados neste estudo sob o ponto de vista do produto da compreensão, constituem material bastante rico para a análise também do processo de leitura, mais especificamente do contexto, inferência, compreensão e principalmente avaliação da compreensão e da proficiência global enquanto processo. Análises futuras desses dados poderiam, portanto, mostrar evidências dos recursos utilizados pelos leitores no processamento da leitura, complementando os dados de processo deste estudo. Um outro exemplo diz respeito a algumas informações coletadas a partir dos questionários, principalmente com relação à concepção de leitura dos sujeitos e do que seria um *bom leitor*, que foram ignoradas. Os dados do questionário metacognitivo, embora tenham sido explorados apenas para a inferência das concepções de leitura dos sujeitos enquanto grupo, classificadas em um contínuo *local/global*, oferece ainda muitas possibilidades de investigação. Uma delas seria verificar até que ponto as estratégias percebidas como importantes pelos leitores são efetivamente usadas no seu processo de leitura.

Apesar de ser uma pesquisa de desenho complexo, envolvendo a triangulação de muitas informações, e três medidas de compreensão em leitura em LE, isto é, perguntas de respostas abertas/resumo, coze e escores do vestibular, além da auto-avaliação, ela ainda não esclarece de maneira satisfatória alguns aspectos da avaliação da compreensão, principalmente relacionadas ao método coze. Devo salientar, nesse caso, que os escores do vestibular foram utilizados apenas como medidas de consistência para outros testes; sua validade, entretanto, é limitada, uma vez que não se tinha informações sobre os critérios de correção do exame utilizados no exame que as gerou. O teste do tipo coze foi usado como uma medida adicional de proficiência em leitura. Nesse estudo, entretanto, não foi elaborado pelo critério da aleatoriedade, que consiste no apagamento das palavras de forma aleatória, dada a existência de uma possível associação entre testes elaborados por esse critério e habilidades de nível mais baixo restritas, em geral, ao nível de oração. O critério adotado nesse caso, foi o de processamento de palavras, de forma a eliminar seu viés e torná-lo também uma medida de processamento de nível mais alto ou de habilidades discursivas. Embora as três medidas se apresentem correlacionadas, o que poderia ser uma indicação da adequação do método coze também como medida de leitura, pode-se observar que as correlações mais altas entre conhecimento de vocabulário e compreensão ainda foram apresentadas pelo teste do tipo coze e as mais baixas pelos escores do teste de perguntas abertas/resumo. Dessa maneira, torna-se difícil concluir se esse critério de elaboração foi realmente mais efetivo enquanto medida de leitura. Estudos posteriores seriam necessários para confirmar esse aspecto.

6.4 – Contribuições

Para abordar de maneira mais objetiva as contribuições deste estudo, subdivido-as em três

modalidades: teóricas, metodológicas e práticas. Focalizo a seguir cada uma delas.

6.4.1 – Contribuições teóricas

Uma contribuição deste estudo estaria em mostrar a adequação de uma competência lexical definida a partir de um conceito *rico* de vocabulário, composta de vários níveis de conhecimento que extrapolam o lexical ou de significado. Além disso, uma competência que não envolve apenas conhecimento, mas também uso, ou seja, que inclui, além de aspectos lingüísticos também aspectos cognitivos, mostrando, a me ver, de maneira mais clara, a inter-relação entre esses aspectos que parece se perder quando segmentações são propostas. Uma competência lexical assim abrangente mostra-se, portanto, mais condizente com um modelo de leitura interativa, e poderia ser usada para fundamentar o ensino e a avaliação do vocabulário necessário para a leitura.

Uma segunda contribuição, específica do estudo de processo, é uma maior compreensão do processo da leitura em LE, e dos recursos utilizados pelos leitores na construção do sentido do texto. Em geral, os estudos de processo que consideram a relação entre leitura/vocabulário se restringem à inferência de palavras não investigando como esses significados são efetivamente incorporados ao sentido em construção, através do uso de recursos ou da interação entre fontes de conhecimento e percepção das restrições do texto.

6.4.2 – Contribuições metodológicas

O desenho de pesquisa inovador, na medida em que procura aliar dois paradigmas ou métodos reconhecidamente contraditórios ou conflitantes, isto é, o quantitativo e qualitativo, assim como a dicotomia produto e processo, colocando-os em uma relação de complementaridade, também pode ser considerado uma contribuição deste estudo para a metodologia de pesquisa. Os resultados obtidos vêm confirmar a eficiência dessa combinação, permitindo não só uma investigação mais abrangente, mas mais consistente com sua base teórica, mostrando contribuições que nenhuma investigação seria capaz de oferecer de modo exclusivo.

Uma outra contribuição está relacionada ao seu desenho de pesquisa abrangente, baseada numa triangulação metodológica que envolve diferentes métodos de coleta de dados, de técnicas de análise e de informações. Como todos os métodos e técnicas apresentam vieses e limitações que muitas vezes se misturam à realidade observada, poder-se-ia dizer que uma combinação oferece melhores condições de se avaliar o problema.

Finalmente, uma terceira contribuição ainda está relacionada à abrangência do desenho de

pesquisa, mas, mais especificamente, às variáveis investigadas. Dada a natureza exploratória deste trabalho, a natureza interativa da leitura e as limitações identificadas na literatura, essa abrangência era essencial, permitindo num mesmo estudo, por exemplo, levantar algumas hipóteses e eliminar outras que necessitariam de outros estudos. Em outras palavras, permitiu explorar caminhos que, no me ponto de vista, caracterizam estudos exploratórios e justificam sua importância.

6.4.3 – Contribuições práticas

As contribuições práticas deste estudo podem ser vistas enquanto um conjunto de subsídios para o ensino e avaliação da leitura e do vocabulário em uma situação de LE. Embora os resultados deste trabalho sejam de natureza exploratória, o que significa que nenhum tratamento específico tenha sido dado ao vocabulário, considero ser possível, a partir da consistência apresentada pelas evidências dos dois estudos, e também pelo fato de ter incluído leitores de níveis diferentes de proficiência, sugerir implicações para essa situação de ensino e avaliação.

Iniciando pela avaliação da leitura, embora essa contribuição também seja pertinente para o ensino, devo salientar, mais uma vez, a importância de um propósito claro e realista para a leitura de textos, sem o qual torna-se impossível uma leitura adequada; a utilização de *tarefas*, por exemplo, seriam uma possibilidade; além de um propósito, elas explicitariam para o leitor também um interlocutor, de forma que ele possa ajustar sua leitura e respostas ao nível de compreensão esperado.

Uma contribuição para a seleção de textos usados tanto na avaliação como no ensino pode ser sugerida a partir deste estudo. Embora a legibilidade, em um modelo de leitura interativo, não possa ser determinada, em princípio, apenas pelos elementos do texto, mas seja resultante da interação das restrições do texto e conhecimentos do leitor, poderia indicar algumas áreas de dificuldades em potencial. Enquanto a seleção de textos em geral é baseada no número de palavras de baixa frequência, segundo um princípio de dificuldade definida por falantes nativos, este estudo parece mostrar que para alunos numa situação de LE também as palavras de frequência média e alta são importantes. Um outro critério, desta vez qualitativo, seria o de contextos *potencialmente ricos* ou *opacos*. Textos com uma concentração grande de contextos *ricos* seriam mais fáceis, na medida em que oferecem ao leitor várias possibilidades de acesso ou poucas restrições de uso. Contextos *opacos*, por sua vez, seriam contextos em que as restrições do texto são muito numerosas, permitindo ao leitor poucas possibilidades de acesso. Sem dúvida alguma, a efetiva dificuldade ou facilidade desses textos iria depender dos conhecimentos do leitor. Nesse sentido, o que se observa, em geral, é que o critério de palavras inferíveis/não inferíveis ou a dificuldade do texto costuma ser visto sob a ótica do falante nativo ou ideal, com uma competência completamente desenvolvida, e não na perspectiva do leitor não nativo. Avaliações mais realistas, levando-se em conta as possibilidades reais

dos leitores não-nativos, ou seus níveis de conhecimento lexical, minimizariam as fontes de frustração e incapacidade frente ao texto que, em geral, resultam de tais situações.

Considerando-se a avaliação da leitura e da proficiência mais global, embora algumas potencialidades de uso do teste do tipo coze possam ser observadas a partir deste estudo, elas têm que ser vistas com uma certa cautela, uma vez que dependeriam de confirmações futuras. A sua utilização na avaliação não apenas do produto da leitura, como no caso deste estudo, mas também do processo, necessário em situações de ensino ou em testes de rendimento, é uma das possibilidades a serem exploradas. Entretanto, esses testes, assim como qualquer outro, precisam de pré-testagem e calibração para serem validados.

Os resultados deste estudo também mostram que as medidas de auto-avaliação se apresentam associadas às outras medidas que tiveram por base testes convencionais, mostrando que os alunos têm uma percepção razoável de suas dificuldades. Essas evidências poderiam, portanto, sugerir a auto-avaliação como uma medida complementar na avaliação da leitura.

Especificamente com relação ao vocabulário, devo salientar que embora sua avaliação de forma discreta ou isolada de outras habilidades não pareça uma atividade muito interessante, pelo menos em termos de ensino, não posso deixar de considerar sua importância para situações de pesquisa, como a deste estudo. Nesse sentido, a operacionalização de um conceito *rico* de vocabulário mostrou-se adequada para um contexto de LE. Comparações entre os desempenhos dos sujeitos nos vários subtestes de vocabulário, apesar de mostrarem uma gradação de dificuldade entre eles, sugere uma consistência bastante grande das medidas. Entretanto, devo lembrar que uma avaliação mais qualitativa do conhecimento de vocabulário implica uma complexidade grande de instrumentos e da própria avaliação que muitas vezes são dispensáveis, considerando-se os objetivos da avaliação. Nesse sentido, este estudo oferece uma contribuição, ao mostrar a possibilidade de se substituir um teste bastante complexo por uma simples *escala de avaliação*, ou mais precisamente de auto-avaliação, que consiste apenas em uma lista de palavras a serem avaliadas pelos alunos. Essa medida mostrou-se associada com a bateria abrangente de conhecimento de vocabulário, contrariando alguns estudos na literatura que a identificam apenas como medida de reconhecimento de vocabulário.

A necessidade de se repensar a ênfase a ser dada à competência lexical no ensino de leitura em inglês em um contexto de LE também pode ser considerada uma implicação deste estudo. Apesar da frequência com que os leitores em LE salientam sua importância e apontam sua dificuldade, essa competência não é geralmente reconhecida e muito menos enfatizada em abordagens que se baseiam em modelos de leitura em LM, como, por exemplo, a abordagem instrumental.

Uma exposição dos alunos às abordagens que focalizam estratégias de leitura sem os conhecimentos básicos para seu uso não parece eficiente, uma vez que este estudo mostrou que os recursos utilizados por cada um dos três leitores, seu padrão de leitura assim como o

sucesso/insucesso na inferência de palavras e construção do sentido do texto, não parecem ter sido determinados apenas pelo tipo de estratégia usado, mas especialmente por sua competência lingüística, que também envolve uma dimensão de uso. Uma análise das estratégias denominadas facilitadoras, tais como *ler em blocos, sem focalizar cada uma das palavras, ou distinguir conscientemente palavras essenciais ao sentido do texto e palavras cujos sentido podem ser ignorados*, ou ainda *usar o contexto* mostram uma dependência muito maior em relação à competência lingüística do que se quer admitir, e muito maior do que alguns leitores em LE, como por exemplo, os de nível intermediário e fraco deste estudo, poderiam apresentar. Além disso, as abordagens que focalizam estratégias de leitura incentivam a necessidade de se conviver com a noção de vagueza e ignorar significados de palavras ou a convivência com sentidos aproximados de palavras, restringindo a consulta a dicionários e as interrupções durante a leitura para busca de significados e definições. Embora tais procedimentos sejam, sem dúvida alguma, importantes sob o ponto de vista da leitura, acabam subestimando a importância do conhecimento de vocabulário e não permitindo um insumo adequado e significativo para o seu desenvolvimento. Não se pode esquecer que o único contato que os alunos têm como a língua em uma situação de LE é, geralmente, durante as poucas horas de aula de leitura.

A possibilidade de uma compensação dessa competência lexical inexistente através de outras fontes também parece estar implícita nessa abordagem. Por exemplo, o uso do conhecimento da LM através de uma ênfase na utilização da noção de cognatos ou do contexto para inferir palavras. Entretanto, o uso de recursos lexicais e o reconhecimento de cognatos não parecem tão eficientes como tem sido demonstrado. Fatores como pronúncia correta, outras palavras na oração, além de um reconhecimento de falsos-cognatos parecem ser fundamentais para que a noção possa ser usada eficientemente. Em outras palavras, seu uso requer uma base lingüística mais sólida. Sem essa base lingüística, algumas compensações podem ocorrer, mas não sem prejuízos para a compreensão. Em geral, a tentativa de uma compensação do conhecimento lexical através de outras fontes, como por exemplo, através de recursos interlingüísticos leva, muitas vezes, a hipóteses prematuras a respeito das palavras e a inferências incorretas e sentidos distorcidos, e, conseqüentemente, à frustração, insegurança e falta de motivação e de autonomia. Além de não ser eficiente, sob o ponto de vista da construção do sentido, também não parece ser eficiente sob o ponto de vista da aprendizagem de palavras, embora os resultados deste estudo não ofereçam evidências a esse respeito. O aprendizado natural de palavras, em situações de LM, parece necessitar de condições especiais, tais como exposição freqüente e continuada às palavras, que, a me ver são inexistentes nas situações de LE amostradas neste estudo.

Quando saliento a necessidade de um ensino de vocabulário mais sistemático não estou me referindo à volta ao ensino de língua tradicional que, em geral, caracteriza-se pela fragmentação, descontextualização ou a visão de vocabulário através de listas de palavras e glossários. Além de

evidências esparsas e muitas vezes conflitantes, não se encontra na literatura sobre vocabulário, como mostrei no Capítulo I, indicações mais precisas de como seria um ensino de vocabulário que conduza à compreensão. Embora o presente estudo não ofereça contribuições diretas com relação à abordagem a ser usada no ensino de vocabulário nas aulas de leitura, traz algumas indicações que poderão ser exploradas e avaliadas em um contexto de LE. Elas não dizem respeito apenas ao *como ensinar* ou que abordagem usar, mas também *o que ensinar* ou *que conteúdo* focalizar. Esse tipo de proposta, fundamentada nos resultados deste estudo, é consistente com a afirmação de Nagy & Herman (1987:33) quando afirmam que

O ensino do vocabulário que realmente faz uma diferença na compreensão geralmente tem algumas das seguintes características: exposições múltiplas às palavras ensinadas, exposição à palavras em contextos significativos, informação rica e variada sobre cada palavra, o estabelecimento de ligações entre as palavras ensinadas e a própria experiência e conhecimento prévio do aprendiz (...).

O ensino que gostaria de sugerir apresenta as características acima mencionadas, indispensáveis para a fluência de acesso às palavras ou um reconhecimento automático que parecem caracterizar bons leitores. Esse ensino extrapola a distinção receptivo/produtivo, uma vez que parece ser necessário muito mais conhecimento para a leitura do que a dicotomia receptivo/produtivo quer reconhecer. Ele não apenas incentiva o leitor a fazer uso dos recursos de que já dispõe, encorajando sua independência e participação ativa através de uma conscientização sobre a leitura e sobre sua experiência com o uso da LM, mas também procura ampliá-los, promovendo a aprendizagem de outras palavras de uma maneira mais abrangente.

Esse ensino, fundamentado em um conceito *rico* de vocabulário, pressupõe, portanto, uma visão de competência que vai além do significado das palavras, envolvendo os vários níveis de conhecimento fonético, sintático, lexical, morfológico, semântico, pragmático-discursivo. Esta seria uma competência definida não apenas quantitativamente, mas qualitativamente, ou em sua profundidade, possibilitando, ao aprendiz, uma integração das palavras aprendidas nas suas várias redes semânticas. Essa abordagem não seria baseada em uma visão atomista de vocabulário, que considera as palavras como unidades portadoras de significado absoluto, mas como elementos para a construção do discurso. Além de todos esses níveis de conhecimento, essa competência também incorpora uma dimensão de uso, que vai possibilitar o uso desses conhecimentos na atualização de significados potenciais na construção do sentido do texto. Em outras palavras, proponho um ensino do componente lingüístico centrado no léxico. Através de um foco em palavras, seriam trabalhados outros aspectos da língua, tais como sua posição na oração ou em sintagmas nominais, suas possibilidades ou funções dentro da oração, sua classe ou classes gramaticais, suas relações semânticas e de colocação, que considera também porções maiores do texto, dentre outros,

explicitando as relações existentes entre forma e significado. Esse tipo de ensino estaria mostrando ao leitor-aluno os vários recursos de acesso aos significados e à construção do sentido para o texto e que é necessário muito mais do que apenas significados para a leitura com compreensão. Os vários significados da palavra e suas conotações seriam enfatizados, uma vez que se mostram particularmente difíceis em LE, dado o distanciamento emocional ou a falta de vivência na língua que em geral caracteriza essas situações. Aspectos fonéticos também seriam salientados, uma vez que a pronúncia correta das palavras permite um acesso mais rápido ao significado em LE e também a possibilidade de intermediação via LM, através do reconhecimento de cognatos, como mostram os resultados deste estudo. Observações informais, a partir de depoimentos e observação de alunos/leitores no contexto foco desta investigação, parecem confirmar esse aspecto ou seja, que muitas dificuldades de inferência são solucionadas a partir de uma simples leitura do texto pelo professor. Não se pode deixar de salientar, entretanto, que, embora importante, a pronúncia correta das palavras não é o único elemento facilitador nesse caso, mas também a segmentação em componentes menores que é feita quando o texto é lido com compreensão.

Entretanto, um trabalho mais efetivo de pronúncia de palavras não parece fazer parte das abordagens que tratam a leitura como uma atividade ou habilidade independente, como é o caso das abordagens instrumentais de ensino da leitura em LE no Brasil. Uma justificativa dessa ausência é que a leitura silenciosa ou a leitura com compreensão é diferente da leitura oral e que a subvocalização incentivaria a leitura palavra por palavra que não é um comportamento desejável em leitura. Entretanto, é necessário salientar que a leitura em blocos maiores só é possível a partir de um processamento automático de palavras que não ocorre muito frequentemente nos protocolos dos leitores menos proficientes.

Não pretendo sugerir que se deva abandonar noções de cognatos ou ignorar semelhanças com a LM mas mostrar o efeito facilitador desses elementos de forma mais realista, ou como um recurso para levantamento de hipóteses, a serem confirmadas através de outros recursos. Também não elimino a possibilidade de utilização de *contextos ricos* ou *potencialmente ricos*, principalmente no início da aprendizagem, para exemplificar as várias possibilidades que o leitor teria de acesso ao texto, proporcionando-lhe mais segurança e motivação. Entretanto, o uso constante de contextos especialmente escolhidos daria, ao longo do tempo, uma falsa idéia das dificuldades envolvidas na inferência de palavras e na busca de compreensão. Concordo com a afirmação de Haarstrup (1989), já mencionada antes, ou seja, *"... enquanto o processamento ('intake') para a compreensão é sempre facilitado por um contexto rico, parece provável que o processamento para a aprendizagem é, correspondentemente, enfraquecido" (1c).*

Em outras palavras, o tipo de abordagem que proponho pressupõe deixar de lado uma visão de léxico subjugado à gramática, em que a estrutura da oração é preenchida com palavras. Esta abordagem pressupõe que o valor discursivo do léxico seja salientado através de suas colocações,

redes semânticas e cadeias coesivas. Nesse tipo de abordagem, trechos da língua são discutidos "lexicalmente", ou como propõe Nattinger (1980), dentre outros, como rotinas pré-fabricadas e não mais como produções sintáticas. As regras, nesse caso, seriam usadas quando houvesse necessidade de adaptação dessas unidades pré-formuladas às restrições sintáticas e aos requisitos do contexto. Uma visão equivalente, embora aparentemente restrita à gramática, é proposta por Widdowson (1989:136),

Esses aglomerados poderiam ser inicialmente simplesmente aglomerados lexicais, cujo significado seria aparente, em associação com um contexto, sem a necessidade de qualquer ajuste gramatical. Então, quando a associação do léxico e contexto se torna insuficiente para o estabelecimento do significado, regras gramaticais seriam usadas para ajudar, para adaptar e ajustar o léxico a quaisquer que sejam as adaptações sintáticas necessárias. Dessa forma, a subserviência das regras e sua função comunicativa estariam claras, e a análise prosseguiria da maneira e quando o acesso aos significados fosse necessário. Tal abordagem seria, eu acho, uma abordagem genuína, realmente uma (autêntica) abordagem comunicativa à pedagogia (2c).

Outra vantagem dessa abordagem, estaria, principalmente, em vir de encontro ao que os leitores percebem como sua maior dificuldade, ou seja, o vocabulário, sendo, portanto, bastante eficiente também sob o ponto de vista da motivação. Como salienta Allwright et alii (1986),

(...) se nós estudamos os processos pelos quais o léxico é adquirido, nós podemos nos aproximar da maneira como a gramática é adquirida. Quando eu observo os aprendizes em classe, eu os vejo, trabalhando com palavras, mas posso vê-los obtendo todos os tipos de ajuda com a gramática. De certa maneira, eles definem sua agenda para a lição através de perguntas sobre palavras. Portanto, nós poderíamos ver, no final, um processo natural de aquisição da gramática através de perguntas sobre o vocabulário" (3c).

Embora a qualidade da competência lexical seja um fator determinante para o acesso automático a significados, também inclui nesse ensino alguns exercícios de reconhecimento rápido de palavras que podem ser encontrados na literatura, tais como exercícios de segmentação, diferentes em LM e LE, principalmente envolvendo palavras de grafia e som semelhante. Uma segmentação mais eficiente permite evitar a leitura palavra por palavra e estimula o leitor a buscar unidades maiores de significado dentro do texto.

Esse tipo de conhecimento lexical mais profundo, voltado para um reconhecimento automático de palavras, não elimina a necessidade de inferência ou de compensações através de outras fontes, como o contexto, mas leva a inferência a outros níveis. O que quero dizer é que

leitores com conhecimento lingüístico vago e impreciso, como dos sujeitos de nível iniciante e intermediário deste estudo, tentam construir um contexto, que é usado com uma função muito mais ampla, não apenas para decidir quais dos significados de uma palavra conhecida, por exemplo, é a mais apropriada à situação em questão, mas para levantar hipóteses e inferir o significado da palavra, o que, quando o número de palavras desconhecidas é muito grande, se torna ineficiente.

Os dois estudos desta tese, além de oferecer algumas contribuições com relação à abordagem de ensino do vocabulário para leitura, sugere também alguns elementos relativos à natureza das palavras que seriam focalizadas nesse ensino. Uma base sólida de palavras de alta frequência parece ter sido fundamental para que o leitor intermediário avançado conseguisse inferir outras palavras menos frequentes. Embora não esteja propondo um controle das palavras a serem ensinadas ou a simplificação de textos, o que não é consistente com uma visão de leitura enquanto interação, considero importante, principalmente num primeiro estágio, salientar a construção de uma base lingüística ou mais especificamente lexical, nos moldes acima descritos, que se iniciaria com palavras mais frequentes, básicas ou centrais (*core words*). Elas constituiriam um vocabulário procedimental (*procedural*), através do qual seria possível se chegar a outras palavras de baixa frequência. A grande dificuldade de inferência dessas palavras é que elas não estão relacionadas a um tema ou tópico. Algumas sugestões nesse sentido podem ser encontradas em Robinson (1980) Hutchinson & Waters (1981) e Carter (1982, 1986, 1987a). Esse seria um vocabulário generalizável e estratégico, que poderia ser usado tanto na leitura de textos gerais como acadêmicos, facilitando a inferência de outras palavras e a construção do sentido do texto, mas também permitindo ao leitor através delas, aprender outras palavras, ampliando, assim, seu vocabulário nos moldes do que ocorre em situações de LM, de forma independente. Essa determinação de centralidade, mais do que as características formais das palavras, envolveria um critério pragmático ou de uso. Em outras palavras, oferece-se, nesse caso, a oportunidade de aquisição de um nível limiar.

Os resultados deste estudo também parecem sugerir que desse grupo fizessem parte palavras que compõem conjuntos fechados ou palavras de função, como pronomes, preposições, conjunções, dentre outras. Uma ênfase especial aos verbos e principalmente verbos frasais (*two-word verbs*), também parece muito importante para facilitar a segmentação de períodos mais longos. Sem dúvida alguma, pesquisas futuras seriam necessárias para confirmar a importância desses aspectos e sugerir outros.

Entretanto, é relevante ressaltar que, quando menciono a necessidade de construção de uma base lingüística não estou sugerindo que ela seria conduzida de forma isolada. A ênfase nessa questão não elimina a necessidade de se considerarem outras frentes, como já mencionei anteriormente. Além desses aspectos lingüísticos, os resultados desta tese também parecem indicar a necessidade de enfatizar a importância de recursos textuais, tais como a organização do texto, sua superestrutura, aparentemente não utilizados de forma explícita pelos sujeitos leitores. Todos esses aspectos,

entretanto, não eliminam a necessidade de se desenvolver no aluno uma postura mais ativa frente aos problemas que enfrenta no processo de leitura, instrumentalizando-o para identificar de forma mais eficiente a origem de seus problemas de compreensão — sejam eles lingüísticos ou extra-lingüísticos.

Uma última sugestão para a leitura que se pode depreender a partir dos resultados deste trabalho seria a necessidade de se criarem condições na sala de aula para que uma leitura crítica ou negociada possa ocorrer. Uma interpretação semântica nos moldes daquela alcançada pelo leitor intermediário avançado, embora dependente de uma competência lexical mais desenvolvida, não se restringe a ela. Essas condições também estão relacionadas à postura e abertura do professor para as contribuições dos alunos e às suas possíveis interpretações, incentivando uma atitude ativa frente ao texto. Entretanto, as condições para uma leitura adequada também seriam dependentes da escolha do texto e do propósito de leitura. Nesse caso, torna-se importante considerar objetivos para essa leitura que vão além da construção da idéia principal, incluindo outros níveis de compreensão. A elaboração do teste de leitura deste estudo, levando em conta um nível de compreensão de idéia geral, ou seja, o resumo, e também de compreensão de detalhes mostrou variações no desempenho dos sujeitos nas duas tarefas. As questões que incluem apenas a idéia principal são passíveis de serem respondidas com um menor conhecimento de vocabulário, mas dão uma idéia incorreta da proficiência em leitura do aluno-leitor, deixando de prepará-lo para uma leitura mais crítica necessária para o estudo acadêmico em geral.

Citações originais

(1c) "... whereas intake for comprehension is always facilitated by rich context, it seems likely that intake for learning is correspondingly weakened".

(2c) "These clusters might initially be simply lexical clusters the meaning of which was apparent, in association with a context, without the need of any grammatical adjustment. Then when the association of lexis and context became insufficient to establish meaning, grammatical rules would be called in to assist, to adapt and adjust the lexis with whatever syntactic fittings were required. In this way, the subservience of rules and their essential communicative function would be clear, and analysis would proceed as and when access to meanings was necessary. Such an approach would, I think, be a genuine indeed an (authentic) communicative approach to pedagogy".

(3c) "... if we study the processes by which the lexis is acquired, we may get closer to the way grammar is acquired. When I watch learners in class, I see them working on words, but I can see them getting all sorts of help with grammar in doing so. In a sense they put their agenda on the lesson through the questions they ask about words. So one might end up seeing a natural process of grammar acquisition through the natural process of enquiring about vocabulary".

APÊNDICE A

Questionários

A.1 – Questionário I

Instruções (apresentadas oralmente):

Vocês foram escolhidos para participarem como sujeitos de minha pesquisa de tese de doutorado. Para a coleta dos dados, que será realizada durante o presente semestre, desenvolvi uma bateria de instrumentos de avaliação. O instrumento de hoje é um questionário que deverão responder em português. Procurem ser claros e específicos em suas respostas. Não discutam as respostas com os colegas. Eu vou ler as instruções e os itens antes de vocês começarem. Se não entenderem alguma coisa, por favor, me perguntem.

1. Nome:

2. RA:

3. Idade:

4. Colegial:

Cidade:

5. Há quantos anos você estuda inglês?

6. Onde você estudou inglês?

a. Nome da escola:

Cidade:

b. Nome do curso particular:

7. Você lê em inglês?

8. Liste os tipos de leitura que você faz em inglês:

a. Na aula de inglês da UNICAMP:

b. Para o seu curso na UNICAMP:

c. Fora do curso:

9. Liste os tipos de leitura que você faz em português:

a. Fora do curso na UNICAMP:

10. Na sua opinião, o que é leitura?

11. Na sua opinião, o que é um bom leitor?

A.2 – Questionário II

Instruções (apresentadas oralmente):

As tarefas de hoje são três, e deverão ser completadas em uma sequência. Primeiro vocês vão ler o texto em anexo, depois responder em português às questões do questionário II e finalmente responder ao questionário III. Para responder a esse último questionário, prestem atenção às instruções impressas, depois da parte introdutória. Para respondê-lo, pensem nas coisas que vocês estiveram fazendo enquanto liam o texto. Se houver alguma coisa que não tenha ficado claro ou não entenderam, por favor, não hesitem em perguntar. Procurem ser claros em suas repostas.

1. Você entendeu o texto? Justifique sua resposta.

2. Quais as dificuldades que encontrou?
3. Achou o texto
 - a. Fácil
 - b. Mais ou menos
 - c. Difícil
4. Explique, em linhas gerais, o que entendeu do texto.
5. Você se considera um bom leitor em inglês? Por quê?
6. Você se considera um bom leitor em português? Por quê?

A.3 – Tarefa de leitura

THE TUNED-OUT GENERATION

A new survey reveals that young people are ignoring the news

Richard Zoglin

Television and radio news floods the airwaves; major events from across the globe pop instantly onto home screens; computers and fax machines relay information in a flash. But anyone who thinks the media boom has created a nation of news junkies needs to readjust his antenna. A sobering new study titled "The Age of Indifference, released last week by the Times Mirror Center for the People & the Press, reveals that young Americans are barely paying attention. The under-30 generation, it reports, "knows less, cares less and reads newspapers less" than any generation in the past five decades.

The sharp drop in newspaper readership is the survey's most dramatic, if least startling, revelation. Only 30% of Americans under 35 said they had "read a newspaper yesterday". That compares with 67% of young people who answered the question affirmatively in a 1965 Gallup poll. More surprisingly, TV has not filled the gap: only 41% of young people said they had watched a TV newscast the day before, down from 52% in 1965.

When it comes to major news events, young people are less interested and informed than their elders. Respondents between the ages of 18 and 29 were 20% less likely to say they had followed important news stories and 40% less likely to be able to identify a newsmaker like German Chancellor Helmut Kohl. Two exceptions: They showed high interest in sports and issues that affected them directly, such as abortion.

The generation gap has widened drastically in recent years. Surveys conducted in the 1940's, '50s and '60s showed that young people were just as interested as their elders in major stories like the McCarthy hearings and the Vietnam War. But since the mid-70's, the under-30 group has been tuning out. The result is a generation that votes less and is less critical of government and business. They are thus an "easy target of opportunity for those seeking to manipulate public opinion", the study warns.

Some news executives attribute this youthful apathy to information overload and explosion of media options. "We had one television in the house, and we had to watch the news when Daddy came home", recalls Steve Friedman, 43, executive producer of NBC's *Nightly News*; today's young people "have got their own TV and their own video systems". Friedman is trying to make the NBC newscast "more relevant" to young viewers by stressing family issues and adding touches of irrelevant humor. Louis Heldman, who is studying how to counteract declining readership for the Knight-Ridder newspaper chain, observes that people today, especially young working women, have less spare time for news. "Information needs to be delivered more efficiently", he says, "to people who are trying to get the kids dressed for school and who spend most of their time with the paper on the seat beside them in the car stalled on the freeway".

The Times Mirror study notes that the young audience has "buoyed the popularity of the new, lighter media forms", such as *People* magazine and TV's *A Current Affair*. The survey may give news executives a further excuse to soften and glitz up their products to try to woo the young. But that means walking a tricky tightrope: in trying to make the news more appetizing, they risk turning it into something other than the news.

Time, 9/7/1990

A.4 – Questionário I, II e III (parte A) preparados para a análise

1. Idade
2. Ano de ingresso na universidade
3. Curso na universidade
4. Anos estudando inglês
(5 anos ou mais)
(menos de 5 anos)
5. Já fez curso particular de inglês
(sim)
(não)
6. Faz curso particular de inglês
(sim)
(não)
7. Lê em inglês
(sim)
(não)
8. Tipos de leitura em inglês
(variadas)
(técnicas)
(nenhuma)
9. Tipos de leitura em português
(variadas)
(técnicas)
(nenhuma)
10. Auto-avaliação da idéia geral do texto
(bom)
(regular)
(ruim)
11. Avaliação do pesquisador da idéia geral do texto
(bom)
(regular)
(ruim)
12. Avaliação do nível de dificuldade do texto
(fácil)
(regular)
(difícil)
13. Avaliação das dificuldades do texto
(sem dificuldades)
(vocabulário)
(gramática)
(vocabulário e gramática)
(outras)
14. Auto-avaliação/sujeito se considera bom leitor em inglês
(sim)

(regular/alguns textos)
(não)

15. Auto-avaliação/sujeito se considera bom leitor em português
(sim)
(regular/alguns textos)
(não)

A.5 – Questionário III

Parte A

1. Nome: _____ Idade: _____

2. RA: _____

3. Turma de Inglês: _____

4. Você lê em outra(s) língua(s), além de português e inglês?

a. Indique: sim não

b. Que língua(s)?

5. Você faz outras leituras em inglês, além daquelas que faz na aula de inglês?

a. Indique: sim não

b. Que tipo?

- jornais

- revistas (Time, Newsweek)

- revistas na sua área

- livros na sua área

- livros (ficção, etc.)

- outras (especifique)

6. Número de anos estudando inglês

a. No Brasil

b. Em outro país

c. Especifique o país

7. Você faz outro curso de inglês, além de LA 122 ou LA 222?

a. Indique: sim não

b. Qual?

Parte B

As afirmações abaixo são sobre leitura silenciosa em inglês. Por favor, usando a escala a seguir, indique seu grau de concordância ou discordância com cada uma das afirmações, colocando um círculo no número apropriado.

1 *Concordo plenamente*

2 *Concordo*

3 *Neutro*

4 *Discordo*

5 *Discordo plenamente*

1. Quando eu leio em inglês, consigo fazer previsões sobre o que está para vir no texto

2. Quando eu leio em inglês, sou capaz de reconhecer a diferença entre os pontos principais e os detalhes do texto

1 2 3 4 5

3. Quando eu leio em inglês, sou capaz de relacionar informações que estão para vir no texto às informações que já apareceram 1 2 3 4 5
4. Quando eu leio em inglês, sou capaz de questionar a significação ou a veracidade do que o autor diz 1 2 3 4 5
5. Quando eu leio em inglês, sou capaz de usar meu conhecimento prévio e experiência para entender o conteúdo do texto que estou lendo 1 2 3 4 5
6. Quando eu leio em inglês, tenho consciência das coisas que eu entendo e das que não entendo 1 2 3 4 5
- Quando eu leio em inglês, se eu não entendo alguma coisa:
7. Eu continuo lendo para entender 1 2 3 4 5
8. Eu releio a parte problemática 1 2 3 4 5
9. Eu volto ao ponto imediatamente anterior à parte problemática e leio tudo de novo 1 2 3 4 5
10. Eu procuro as palavras desconhecidas no dicionário 1 2 3 4 5
11. Eu desisto e paro de ler 1 2 3 4 5
- Quando eu leio em inglês, as coisas que me ajudam a ler eficientemente são:
12. Falar mentalmente as partes das palavras 1 2 3 4 5
13. Entender o significado de cada palavra 1 2 3 4 5
14. Entender o significado global do texto 1 2 3 4 5
15. Ser capaz de pronunciar cada palavra 1 2 3 4 5
16. Focalizar estruturas gramaticais 1 2 3 4 5
17. Relacionar o texto ao que já sei do tópico 1 2 3 4 5
18. Procurar palavras no dicionário 1 2 3 4 5
19. Focalizar nos detalhes do texto 1 2 3 4 5
20. Focalizar na organização do texto 1 2 3 4 5
- Quando eu leio em inglês, as coisas que dificultam minha leitura são:
21. Os sons das palavras 1 2 3 4 5
22. A pronúncia das palavras 1 2 3 4 5
23. O reconhecimento das palavras 1 2 3 4 5
24. As estruturas gramaticais 1 2 3 4 5
25. O alfabeto 1 2 3 4 5
26. Relacionar o texto com o que eu já sei sobre o tópico 1 2 3 4 5
27. Entender o significado global do texto 1 2 3 4 5
28. A organização do texto 1 2 3 4 5

Na minha opinião, o que faz com que um leitor seja bom é sua habilidade em:

29. Reconhecer o significado de palavras	1	2	3	4	5
30. Pronunciar palavras	1	2	3	4	5
31. Entender o significado global do texto	1	2	3	4	5
32. Usar um dicionário	1	2	3	4	5
33. Adivinhar o significado de palavras	1	2	3	4	5
34. Integrar a informação do texto com a informação que eu já tenho	1	2	3	4	5
35. Focalizar nos detalhes do conteúdo	1	2	3	4	5
36. Apreender a organização do texto	1	2	3	4	5

A.6 – Questionário III preparado para a análise

Parte B

Confiança: 6 afirmações (*conf1*) a (*conf6*)

(<i>conf1</i>) Quando eu leio em inglês, consigo fazer previsões sobre o que está para vir no texto	1	2	3	4	5
(<i>conf2</i>) Quando eu leio em inglês, sou capaz de reconhecer a diferença entre os pontos principais e os detalhes do texto	1	2	3	4	5
(<i>conf3</i>) Quando eu leio em inglês, sou capaz de relacionar informações que estão para vir no texto às informações que já apareceram	1	2	3	4	5
(<i>conf4</i>) Quando eu leio em inglês, sou capaz de questionar a significação ou a veracidade do que o autor diz	1	2	3	4	5
(<i>conf5</i>) Quando eu leio em inglês, sou capaz de usar meu conhecimento prévio e experiência para entender o conteúdo do texto que estou lendo	1	2	3	4	5
(<i>conf6</i>) Quando eu leio em inglês, tenho consciência das coisas que eu entendo e das que não entendo	1	2	3	4	5

Correção ou reparo: 5 afirmações (*rep7*) a (*rep11*)

Quando eu leio em inglês, se eu não entendo alguma coisa,

(<i>rep7</i>) Eu continuo lendo para entender	1	2	3	4	5
(<i>rep8</i>) Eu releio a parte problemática	1	2	3	4	5
(<i>rep9</i>) Eu volto ao ponto imediatamente anterior à parte problemática e leio tudo de novo	1	2	3	4	5
(<i>rep10</i>) Eu procuro as palavras desconhecidas no dicionário	1	2	3	4	5
(<i>rep11</i>) Eu desisto e paro de ler	1	2	3	4	5

Eficiência: 17 afirmações (*efi12*) a (*efi20*)

Quando eu leio em inglês, as coisas que me ajudam a ler eficientemente são:

(efi12)	Falar mentalmente as partes das palavras	1 2 3 4 5
(efi13)	Entender o significado de cada palavra	1 2 3 4 5
(efi14)	Entender o significado global do texto	1 2 3 4 5
(efi15)	Ser capaz de pronunciar cada palavra	1 2 3 4 5
(efi16)	Focalizar estruturas gramaticais	1 2 3 4 5
(efi17)	Relacionar o texto ao que já sei do tópico	1 2 3 4 5
(efi18)	Procurar palavras no dicionário	1 2 3 4 5
(efi19)	Focalizar nos detalhes do texto	1 2 3 4 5
(efi20)	Focalizar na organização do texto	1 2 3 4 5
Dificuldade: 8 afirmações (dif21) a (dif28)		1 2 3 4 5

Quando eu leio em inglês, as coisas que dificultam minha leitura são:

(dif21)	Os sons das palavras	1 2 3 4 5
(dif22)	A pronúncia das palavras	1 2 3 4 5
(dif23)	O reconhecimento das palavras	1 2 3 4 5
(dif24)	As estruturas gramaticais	1 2 3 4 5
(dif25)	O alfabeto	1 2 3 4 5
(dif26)	Relacionar o texto com o que eu já sei sobre o tópico	1 2 3 4 5
(dif27)	Entender o significado global do texto	1 2 3 4 5
(dif28)	A organização do texto	1 2 3 4 5

Eficiência: (efi29) a (efi36)

Na minha opinião, o que faz com que um leitor seja bom é sua habilidade em:

(efi29)	Reconhecer o significado de palavras	1 2 3 4 5
(efi30)	Pronunciar palavras	1 2 3 4 5
(efi31)	Entender o significado global do texto	1 2 3 4 5
(efi32)	Usar um dicionário	1 2 3 4 5
(efi33)	Adivinhar o significado de palavras	1 2 3 4 5
(efi34)	Integrar a informação do texto com a informação que eu já tenho	1 2 3 4 5
(efi35)	Focalizar nos detalhes do conteúdo	1 2 3 4 5
(efi36)	Apreender a organização do texto	1 2 3 4 5

APÊNDICE B

Instrumentos/procedimentos para a avaliação do vocabulário

B.1 – Texto-base para testes de vocabulário

GALACTIC BIRTH? A surprise for scientists

Astronomers have long believed that galaxies, clusters that usually contain billions of stars, were all formed shortly after the Big Bang, the cataclysmic explosion some 15 billion years ago that spawned the universe. But that conviction was shaken last week when scientists announced that they had found evidence of a cosmic version of gestation: a galaxy preparing for birth. Said James Gunn, a Princeton University astrophysicist, after the announcement: "This is the Rosetta stone of galaxy formation." The apparent galactic embryo — actually a massive, disk-shaped cloud of hydrogen gas — was discovered fortuitously last spring by Cornell University astronomer Martha Haynes and her colleague Ricardo Giovanelli, when they were monitoring signals in outer space with the 1,000-ft. radio telescope at Arecibo, Puerto Rico. While focusing the telescope on what they thought was empty space in order to calibrate it, the astronomers picked up a signal pattern resembling that emitted by galaxies. The invisible cloud — estimated to be ten times as large as the Milky Way — loomed fairly close, astronomically speaking: 65 million light-years from earth. Since a light-year is the distance light travels in a year, the scientists were receiving signals from the cloud as it appeared 65 million years ago. Because it apparently contained no stars, the scientists concluded that they were observing a galaxy about to be born. Said Giovanelli: "This cloud indicates that galaxies can form slowly throughout the history of the universe and are not something that happened during some magical period in the distant past."

Time Magazine, 11/9/1989

B.2 – Procedimento para seleção de palavras para os testes de vocabulário

As palavras a serem incluídas nos testes de vocabulário foram selecionadas por um grupo de 28 alunos de diferentes áreas (Engenharia Elétrica, Estatística, Ciência da Computação, Matemática Aplicada e Engenharia Civil), cursando a disciplina Inglês Instrumental nos níveis I e II na UNICAMP no primeiro semestre de 1990, isto é, no semestre anterior ao semestre em que se realizou a coleta de dados. Esses alunos podem ser considerados uma amostra representativa dos alunos que participaram do estudo principal em termos de proficiência em leitura e, possivelmente, também em termos de conhecimento de vocabulário. Para a seleção das palavras, obviamente foi usado o mesmo texto do estudo principal.

Três testes (vide Apêndice B.3) foram usados para verificar a consistência das respostas, uma vez que se havia observado anteriormente que nem sempre os alunos grifam todas as palavras desconhecidas.

Teste 1

Texto apresentado aos alunos, com instruções para grifar todas as palavras desconhecidas.

Teste 2

Texto novamente apresentado, com instruções para grifar todas as palavras conhecidas.

Teste 3

Texto apresentado contendo todas as palavras, embaralhadas, com instruções para grifar todas as palavras desconhecidas.

Os resultados são apresentados no Apêndice B, Seção B.4.

20 dos 28 alunos fizeram primeiro o teste 1, depois o 2 e 3. Os outros alunos fizeram primeiro o teste 3, depois o 1 e 2. A diferença na sequência não parece ter afetado os resultados. Entretanto, houve, como já se previa, algumas inconsistências quando os resultados dos três testes foram comparados. Por exemplo, um número maior de palavras foram grifadas no teste 3 (palavras isoladas) do que no 1 e 2, o que é perfeitamente explicável se se considerar que os alunos poderiam estar usando o contexto para inferir palavras. Entretanto, há também inconsistências entre os resultados dos testes 1 e 2, que também pode ser explicado como algum tipo de "efeito de prática" (uma oportunidade extra de encontrar palavra).

Usando-se um critério arbitrário de 10% como ponto de corte, foram selecionadas 20 palavras. Seguindo-se o critério de corte, teria-se 24 palavras, portanto 4 palavras foram deixadas de lado. Elas foram: *embryo* (7%, 7%, 11%, cognata); *outer* (7%, 11%, 11%); *colleague* (0%, 0%, 22%), apenas sublinhada

quando ocorre isolada, também cognata; *announcement* (0%, 0%, 15%), mesmo caso da anterior.
Às 20 palavras foram adicionadas 10 palavras criadas, numa porcentagem de 50%.

B.3 – Instruções para testes de seleção de palavras

Teste 1

- Leia o texto *Galactic Birth* abaixo apresentado uma única vez e grife somente as palavras cujos significados você desconhece.

Teste 2

- Leia o texto mais uma vez e grife somente as palavras cujos significados você conhece, isto é, palavras que não seriam problemas para você durante a leitura.

Teste 3

- Grife somente as palavras cujos significados você desconhece:

surprise - astronomers - scientists - have - long - that - believed - galaxies - clusters - usually - billions - contain - of - were - stars - all - shortly - formed - after - Big - Bang - cataclysmic - some - explosion - ago - spawned - universe - but - conviction - shaken - was - last - when - week - announced - they - found - had - evidence - the - cosmic - a - version - gestation - preparing - said - for - university - this - astrophysicist - announcement - is - stone - formation - birth - galactic - apparent - embryo - actually - massive - disk-shaped - hydrogen - cloud - gas - discovered - fortuitously - spring - astronomer - colleague - and - by - signals - monitoring - in - outer - with - space - 1,000-ft. - telescope - at - radio - while - on - focusing - what - thought - empty - calibrate - in order to - picked up - pattern - emitted - resembling - invisible - estimated - to be - ten - large - times - Milky Way - loomed - close - astronomically - fairly - speaking - million - light-years - earth - from - since - light - distance - travels - year - receiving - appeared - because - contained - apparently - no - stars - concluded - observing - about - born - indicates - can - form - slowly - history - throughout - universe - just - something - happened - some - during - period - magical - past - distant.

B.4 – Número de alunos relatando palavras desconhecidas

Tabela B.1 – Número de alunos relatando palavras desconhecidas

Palavras	Teste 1		Teste 2		Teste 3	
	Nº de alunos	Percentual	Nº de alunos	Percentual	Nº de alunos	Percentual
Loomed	24	86%	25	89%	25	93%
Cluster	22	79%	20	71%	16	59%
Spawned	20	71%	21	75%	24	89%
Resembling	17	61%	21	75%	15	56%
Cloud	13	46%	12	43%	4	15%
Shaken	11	39%	13	46%	15	56%
Disk-shaped	9	32%	14	50%	10	37%
Fortuitously	7	25%	10	36%	10	37%
Fairly	7	25%	14	50%	10	37%
Massive	6	21%	7	25%	3	11%
Picked up	6	21%	10	36%	7	26%
Pattern	5	18%	6	21%	3	11%
Focusing	4	14%	5	18%	6	22%
Spring	4	14%	6	21%	5	19%
Empty	3	11%	4	14%	1	4%
Throughout	3	11%	1	4%	6	22%
Outer	2	7%	3	11%	3	11%
Signals	2	7%	2	7%	0	0%
Embryo	2	7%	2	7%	3	11%
Appeared	1	4%	4	14%	4	15%
Cataclysmic	1	4%	1	4%	2	7%
While	1	4%	2	7%	0	0%
Gestation	1	4%	1	4%	1	4%
Happened	1	4%	2	7%	2	7%
Born	1	4%	2	7%	2	7%
Believed	1	4%	1	4%	0	0%
Stone	1	4%	2	7%	2	7%
Slowly	1	4%	3	11%	0	0%
Receiving	1	4%	1	4%	3	11%
Calibrate	1	4%	0	0%	0	0%
Something	1	4%	1	4%	0	0%
Announced	0	0%	0	0%	2	7%
Stars	0	0%	0	0%	1	4%
Colleague	0	0%	0	0%	6	22%
Announcement	0	0%	0	0%	4	15%
Formed	0	0%	0	0%	1	4%
Concluded	0	0%	0	0%	1	4%
Monitoring	0	0%	0	0%	1	4%
During	0	0%	0	0%	1	4%
Close	0	0%	4	14%	0	0%
Astrophysicist	0	0%	0	0%	1	4%
Total de alunos	28	100%	28	100%	27	100%

B.5 – Teste de vocabulário I (Familiaridade)

- Indique, na escala de 1 a 4 que se segue, seu grau de familiaridade com cada uma das palavras abaixo.

1. *Eu nunca vi essa palavra antes*
2. *Essa palavra existe mas não sei o que quer dizer*
3. *Essa palavra existe mas estou em dúvida com relação ao que quer dizer*
4. *Eu conheço essa palavra. Eu sei o que quer dizer*

Exemplo: LOVE 1 2 3 4✓

1. LODINGS	1 2 3 4
2. CLUSTERS	1 2 3 4
3. SKAPPING	1 2 3 4
4. SPAWNED	1 2 3 4
5. SHAKEN	1 2 3 4
6. GLIM	1 2 3 4
7. LOOMED	1 2 3 4
8. DREALLED	1 2 3 4
9. RESEMBLING	1 2 3 4
10. CLOUD	1 2 3 4
11. LEAM	1 2 3 4
12. DISK-SHAPED	1 2 3 4
13. TARK	1 2 3 4
14. FORTUITOUSLY	1 2 3 4
15. FAIRLY	1 2 3 4
16. OBSERVEMENT	1 2 3 4
17. MASSIVE	1 2 3 4
18. TRIVEL	1 2 3 4
19. PICKED UP	1 2 3 4
20. FOCUSING	1 2 3 4
21. SPRING	1 2 3 4
22. EMPTY	1 2 3 4
23. THROUGHOUT	1 2 3 4
24. PATTERN	1 2 3 4
25. APPEARED	1 2 3 4
26. CLOSE	1 2 3 4
27. ADJUSTION	1 2 3 4
28. RECEIVING	1 2 3 4
29. RESKS	1 2 3 4
30. SLOWLY	1 2 3 4

B.6 – Teste de vocabulário II (Frequência)

- Indique, na escala de 1 a 3 que se segue, a frequência de cada uma das palavras abaixo nos textos que costuma ler.

1. *Não freqüente*
2. *Mais ou menos freqüente*
3. *Freqüente*

Exemplo: LOVE 1 2 3✓

1. LODINGS	1 2 3
2. CLUSTERS	1 2 3
3. SKAPPING	1 2 3
4. SPAWNED	1 2 3
5. SHAKEN	1 2 3
6. GLIM	1 2 3

7. LOOMED	1 2 3
8. DREALLED	1 2 3
9. RESEMBLING	1 2 3
10. CLOUD	1 2 3
11. LEAM	1 2 3
12. DISK-SHAPED	1 2 3
13. TARK	1 2 3
14. FORTUITOUSLY	1 2 3
15. FAIRLY	1 2 3
16. OBSERVEMENT	1 2 3
17. MASSIVE	1 2 3
18. TRIVEL	1 2 3
19. PICKED UP	1 2 3
20. FOCUSING	1 2 3
21. SPRING	1 2 3
22. EMPTY	1 2 3
23. THROUGHOUT	1 2 3
24. PATTERN	1 2 3
25. APPEARED	1 2 3
26. CLOSE	1 2 3
27. ADJUSTION	1 2 3
28. RECEIVING	1 2 3
29. RESKS	1 2 3
30. SLOWLY	1 2 3

B.7 – Subteste de vocabulário III-A (Derivação/flexão)

- Dê palavras derivadas de cada uma das palavras abaixo.

Exemplo: LOVE – lover, loving, beloved, lovemaking, lovesick, loverlike, etc.

- No caso de palavras já derivadas, dê as palavras das quais elas se derivaram.

Exemplo: BELOVED – love

1. CLUSTERS
2. SPAWNED
3. SHAKEN
4. LOOMED
5. RESEMBLING
6. CLOUD
7. DISK-SHAPED
8. FORTUITOUSLY
9. FAIRLY
10. MASSIVE
11. PICKED UP
12. FOCUSING
13. SPRING
14. EMPTY
15. THROUGHOUT
16. PATTERN
17. APPEARED
18. CLOSE
19. RECEIVING
20. SLOWLY

- Dê o singular/plural (quando for o caso) de cada uma das palavras abaixo.

Exemplo: LOVER – lovers

1. CLUSTER
2. SPAWNED
3. SHAKEN
4. LOOMED
5. RESEMBLING
6. CLOUD
7. DISK-SHAPED
8. FORTUITOUSLY
9. FAIRLY
10. MASSIVE
11. PICKED UP
12. FOCUSING
13. SPRING
14. EMPTY
15. THROUGHOUT
16. PATTERN
17. APPEARED
18. CLOSE
19. RECEIVING
20. SLOWLY

singular/plural

B.8 – Subteste de vocabulário III-B (Significado/definição)

- Dê o significado de cada uma das palavras abaixo. Você pode usar:
um sinônimo (em inglês)
um sinônimo (em português)
uma definição ou explicação (em português ou inglês)
- Dê outros significados adicionais, quando for o caso.

1. CLUSTERS
2. SPAWNED
3. SHAKEN
4. LOOMED
5. RESEMBLING
6. CLOUD
7. DISK-SHAPED
8. FORTUITOUSLY
9. FAIRLY
10. MASSIVE
11. PICKED UP
12. FOCUSING
13. SPRING
14. EMPTY
15. THROUGHOUT
16. PATTERN
17. APPEARED
18. CLOSE
19. RECEIVING
20. SLOWLY

B.9 – Subteste de vocabulário III-C (Associação)

- Dê todas as palavras em inglês que você conseguir imaginar que estão associadas com as palavras abaixo.
Exemplo: ACCIDENT – car, victims, hospital, ambulance, police, etc.

1. CLUSTERS
2. SPAWNED
3. SHAKEN

4. LOOMED
5. RESEMBLING
6. CLOUD
7. DISK-SHAPED
8. FORTUITOUSLY
9. FAIRLY
10. MASSIVE
11. PICKED UP
12. FOCUSING
13. SPRING
14. EMPTY
15. THROUGHOUT
16. PATTERN
17. APPEARED
18. CLOSE
19. RECEIVING
20. SLOWLY

B.10 – Subteste de vocabulário III-D (Colocação)

- Dê as palavras que freqüentemente ocorrem com as palavras abaixo.
- Exemplos: A palavra SUNNY ocorre freqüentemente com a palavra DAY.
A palavra TOO ocorre freqüentemente com as palavras MUCH, LITTLE, etc.

1. CLUSTERS
2. SPAWNED
3. SHAKEN
4. LOOMED
5. RESEMBLING
6. CLOUD
7. DISK-SHAPED
8. FORTUITOUSLY
9. FAIRLY
10. MASSIVE
11. PICKED UP
12. FOCUSING
13. SPRING
14. EMPTY
15. THROUGHOUT
16. PATTERN
17. APPEARED
18. CLOSE
19. RECEIVING
20. SLOWLY

B.11 – Subteste de vocabulário III-E (Contexto/classe gramatical)

- Marque com um X apenas os contextos que considera possíveis para cada uma das palavras abaixo.
- Dê a(s) classe(s) gramatical(ais) de cada uma das palavras abaixo.

Classe(s) gramatical(ais)

1. CLUSTERS:
1. It clusters
 2. They cluster
 3. Clusters of
 4. In clusters
 5. He clusters

6. They cluster together
7. The big clusters
8. The cluster books

2. SPAWNED:

1. They spawned
2. It spawned them
3. They spawned it
4. The spawned frog
5. It spawned the movement

3. SHAKEN:

1. By shaken
2. He was being shaken
3. Very well shaken
4. A good shaken
5. The earth has shaken
6. It has shaken him
7. He shaken it
8. The shaken milk
9. I drive shaken

4. LOOMED:

1. The large loomed
2. They loomed
3. It loomed
4. He drove loomed

5. RESEMBLING:

1. You resembling him
2. It is resembling that
3. The good resembling
4. The resembling sister
5. He is resembling him
6. By resembling

6. CLOUD:

1. The little cloud
2. A cloud of
3. In clouds
4. In the clouds
5. They cloud it
6. He comes cloud

7. DISK-SHAPED:

1. They disk-shaped
2. The disk-shaped mirror
3. The disk-shaped of
4. In disk-shaped
5. The good disk-shaped

8. FORTUITOUSLY:

1. A fortuitously encounter
2. He found it fortuitously
3. They fortuitously it
4. It was fortuitously found

9. FAIRLY:

1. It was fairly accurate

2. A fairly heavy book
3. It fairly flew up and down
4. It is fairly distributed
5. They fairly him
6. A fairly person

10. MASSIVE:

1. The good massive
2. To increase massive
3. She got massive fat
4. A massive man
5. They massive

11. PICKED UP:

1. He picked himself up
2. The picked up book
3. It was picked up
4. They picked him up
5. The good picked up

12. FOCUSING:

1. The focusing camera
2. They are focusing it
3. He is focusing them on
4. The focusing of

13. SPRING:

1. In the spring
2. Each spring
3. With a spring
4. They spring
5. They spring from
6. The spring box
7. The very spring box

14. EMPTY:

1. It was empty
2. The empty place
3. It is empty of
4. We empty it
5. the good empty
6. We drive empty

15. THROUGHOUT:

1. It goes throughout
2. They throughout
3. It was carpeted throughout
4. The throughout road

16. PATTERN:

1. The work pattern
2. A nice pattern of
3. To draw a pattern
4. Pattern carpets
5. Social pattern
6. To pattern on

17. APPEARED:

1. They appeared from

2. It appeared good
3. He appeared to be
4. The appeared man
5. She appeared

18. CLOSE:

1. They close it
2. The close doors
3. With windows close
4. We close at
5. The close book
6. They were close together
7. Very close friends
8. Close enough
9. She was close to
10. The good close

19. RECEIVING:

1. He receiving it
2. They have been receiving it
3. The receiving letter
4. By receiving
5. The good receiving

20. SLOWLY:

1. He goes too slowly
2. The slowly man
3. They slowly it

B.12 – Teste de vocabulário IV (Inferência em contexto)

- Leia o texto e dê o significado, mesmo que aproximado, das palavras em negrito.

GALACTIC BIRTH?
A surprise for scientists

Astronomers have long believed that galaxies, clusters that usually contain billions of stars, were all formed shortly after the Big Bang, the cataclysmic explosion some 15 billion years ago that spawned the universe. But that conviction was **shaken** last week when scientists announced that they had found evidence of a cosmic version of gestation: a galaxy preparing for birth. Said James Gunn, a Princeton University astrophysicist, after the announcement: "This is the Rosetta stone of galaxy formation." The apparent galactic embryo – actually a massive, **disk-shaped** cloud of hydrogen gas – was discovered **fortuitously** last spring by Cornell University astronomer Martha Haynes and her colleague Ricardo Giovanelli, when they were monitoring signals in outer space with the 1,000-ft. radio telescope at Arecibo, Puerto Rico. While focusing the telescope on what they thought was empty space in order to calibrate it, the astronomers picked up a signal **pattern resembling** that emitted by galaxies. The invisible cloud -- estimated to be ten times as large as the Milky Way -- **loomed fairly close**, astronomically speaking: 65 million light-years from earth. Since a light-year is the distance light travels in a year, the scientists were receiving signals from the cloud as it **appeared** 65 million years ago. Because it apparently contained no stars, the scientists concluded that they were observing a galaxy about to be born. Said Giovanelli: "This cloud indicates that galaxies can form slowly throughout the history of the universe and are not something that happened during some magical period in the distant past".

Time Magazine, 11/9/1989

1. CLUSTERS:
2. SPAWNED:
3. SHAKEN:

4. LOOMED:
5. RESEMBLING:
6. CLOUD:
7. DISK-SHAPED:
8. FORTUITOUSLY:
9. FAIRLY:
10. MASSIVE:
11. PICKED UP:
12. FOCUSING:
13. SPRING:
14. EMPTY:
15. THROUGHOUT:
16. PATTERN:
17. APPEARED:
18. CLOSE:
19. RECEIVING:
20. SLOWLY:

B.13 – Testes de vocabulário categorizados para a análise

1. (T1:FAMILIARIDADE)

- 1= nunca vi essa palavra
- 2= palavra existe mas não sei o significado
- 3= palavra existe mas estou em dúvida em relação ao significado
- 4= conheço a palavra, sei seu significado

2. (T2:FREQUÊNCIA)

- 1= não freqüente
- 2= mais ou menos freqüente
- 3= freqüente

3. (T3:FLEXÃO E DERIVAÇÃO)

- 0= em branco
- 1= errada
- 2= certa

4. (T4:DEFINIÇÃO E SIGNIFICADO)

- 0= em branco
- 1= errada
- 2= certa

5. (T5:ASSOCIAÇÃO)

- 0= em branco
- 1= 1 certa
- 2= 2 certas
- 3= mais de 2 certas

6. (T6:COLOCAÇÃO)

- 0= em branco
- 1= 1 certa
- 2= mais de 1 certa

7. (T7:CONTEXTO GRAMATICAL)

- 1= todas erradas
- 2= menos de meia certa
- 3= meia certa
- 4= certa

8. (T8:CLASSE GRAMATICAL)

- 1= todas erradas
- 2= meia certa
- 3= certa

9. (T9:INFERÊNCIA EM CONTEXTO)

- 0= em branco
- 1= errada
- 2= certa

B.14 – Testes de vocabulário recategorizados após primeira análise

1. (T1:FAMILIARIDADE)

- 1= não familiar
- 2= não familiar
- 3= mais ou menos familiar
- 4= familiar

2. (T2:FREQÜÊNCIA)

- 1= não frequente
- 2= mais ou menos frequente
- 3= frequente

3. (T3:FLEXÃO/DERIVAÇÃO)

- 0= não soube
- 1= não soube
- 2= soube

4. (T4:DEFINIÇÃO E SIGNIFICADO)

- 0= não soube
- 1= não soube
- 2= soube

5. (T5:ASSOCIAÇÃO)

- 0= não soube
- 1= soube
- 2= soube
- 3= soube

6. (T6:COLOCAÇÃO)

- 0= não soube
- 1= soube
- 2= soube

7. (T7:CONTEXTO GRAMATICAL)

- 1= não soube
- 2= soube
- 3= soube
- 4= soube

8. (T8:CLASSE GRAMATICAL)

- 1= não soube
- 2= soube
- 3= soube

9. (T9:INFERÊNCIA EM CONTEXTO)

- 0= não soube
- 1= não soube
- 2= soube

APÊNDICE C

Instrumento para a avaliação do conhecimento conceitual

C.1 – Teste de conhecimento conceitual

- Explique em português o que você entende por:

1. GALÁXIA
2. BIG BANG
3. PEDRA DE ROSETTA
4. EMBRIÃO GALÁCTICO
5. VIA LÁCTEA
6. ANO-LUZ

C.2 – Teste de conhecimento conceitual categorizado para análise

1. GALÁXIA
 - 1 (correto)
 - 2 (meio correto)
 - 3 (incorreto)
2. BIG BANG
 - 1 (correto)
 - 2 (meio correto)
 - 3 (incorreto)
3. PEDRA DE ROSETTA
 - 1 (correto)
 - 2 (meio correto)
 - 3 (incorreto)
4. EMBRIÃO GALÁCTICO
 - 1 (correto)
 - 2 (meio correto)
 - 3 (incorreto)
5. VIA LÁCTEA
 - 1 (correto)
 - 2 (meio correto)
 - 3 (incorreto)
6. ANO-LUZ
 - 1 (correto)
 - 2 (meio correto)
 - 3 (incorreto)

C.3 – Teste de conhecimento conceitual recategorizado após primeira análise

1. GALÁXIA
 - 1 (correto)
 - 3 (incorreto)
2. BIG BANG
 - 1 (correto)
 - 3 (incorreto)

3. PEDRA DE ROSETTA

1 (correto)

3 (incorreto)

4. EMBRIÃO GALÁCTICO

1 (correto)

3 (incorreto)

5. VIA LÁCTEA

1 (correto)

3 (incorreto)

6. ANO-LUZ

1 (correto)

3 (incorreto)

APÊNDICE D

Instrumentos para a avaliação da leitura em LM e LE e proficiência em LE

D.1 – Texto-base para Cloze I

O HOMEM VEZES O HOMEM

Calcula-se que até o final do século três crianças nascerão a cada segundo

A população atual da terra é de 5,3 bilhões de habitantes. Nos poucos segundos que se leva para ler essa frase ela já aumentou em 18 pessoas. A extraordinária velocidade não é, porém, um fenômeno de agora. Há 22 anos, quando a população do mundo estava estimada em 3,5 bilhões de pessoas, a lotação do planeta tirava o sono do biólogo americano Paul Ehrlich – e o levou a escrever o livro *A Bomba Populacional*, que previa sérias dificuldades para a humanidade se ela continuasse a crescer tão rapidamente. O livro causou polêmica e o autor foi acusado de alarmista. As críticas parecem não ter abalado a força de suas convicções. Ehrlich acaba de lançar seu segundo livro sobre o tema: uma espécie de desforra aos que duvidaram de suas previsões, intitulado *A Explosão Populacional*. Como o nome diz, o livro não se restringe a prever catástrofes. Agora, ele as mostra. Para o biólogo as falhas na camada de ozônio, o efeito-estufa que aquece a Terra e ameaça mudar seu mapa e a epidemia de Aids são alguns dos sintomas da grande tragédia que havia antecipado.

Ehrlich estuda as relações entre população e ecologia e é essa questão que fundamenta seus livros. O que o preocupa é o desequilíbrio dessas relações. Junto ao crescimento populacional aumenta o uso dos recursos naturais e o desgaste do ecossistema. Segundo Ehrlich, esse ecossistema, do qual os homens dependem para sobreviver, está em vias de exaustão. Ele se baseia em dados do Centro de Estudos Populacionais, um observatório populacional instalado em Washington, para afirmar que na última década do século o planeta ingressa no ponto máximo de seu crescimento demográfico em números absolutos: um ritmo estonteante de 93 milhões de novos habitantes por ano, o que significa três nascimentos por segundo.

A tese de *A Explosão Populacional* é a de que diminuir apenas o ritmo de crescimento não basta. Se bastasse, Ehrlich teria motivos de sobra para se tranquilizar. Demógrafos do mundo inteiro estão constatando que o crescimento populacional acabou se mostrando inferior ao que se havia profetizado décadas atrás. A demógrafa paulista Suzana Pasternac, professora de Demografia e Urbanização da Faculdade de Arquitetura e Urbanismo da Universidade de São Paulo, diz que esse tipo de projeção está fadado ao engano. "Nós sabemos que as projeções de população estão sempre meio erradas", diz Suzana. "Nessas projeções se usam dados do momento, é uma lógica conservadora em que não se levam em conta alterações de comportamento que possam vir a acontecer e que podem influir no crescimento demográfico", explica. Segundo ela, quando o biólogo americano escreveu o primeiro livro, em 1968, a explosão demográfica era um fantasma que aterrorizava todos os demógrafos. Hoje, é como se Ehrlich vivesse um pouco do passado. Na verdade, ele mesmo reconhece que o ritmo de crescimento demográfico está diminuindo sensivelmente graças aos programas de controle populacional aplicados em todos os cantos do mundo. Sua preocupação, porém, resiste.

"Cientistas do Primeiro Mundo sempre tiveram medo do crescimento populacional do Terceiro Mundo", teoriza Suzana. "Mas não precisam mais se preocupar. A famosa explosão populacional acabou", garante. Diz a demógrafa que no Brasil hoje com 145.762 milhões de habitantes mesmo que essa tendência se reverta, os filhos que deixaram de nascer, e portanto de procriar, já implicam uma geração menor. Suzana participou da elaboração do livro *Tendências Demográficas Recentes e Perspectivas para a Próxima Década*, no prelo desde a divulgação do Plano Collor, que será lançado pela Fundap e Fundação Secretaria Estadual de Análise de Dados (Seade). O livro mostra claramente a reversão da tendência à explosão no Brasil. A taxa de crescimento dos brasileiros, que em 1950 foi de 2,9%, em 1990 não deve ultrapassar 1,6%. Se os fatores projetados permanecerem, a previsão é de que até 1995 o índice caia para 1,5%.

Há basicamente cinco fatores que interferem no crescimento da população: opção pelo celibato; idade das pessoas que casam - na China, onde o controle populacional é rigoroso, é proibido o casamento de pessoas muito jovens e desestimulada a gravidez de mães solteiras; aumento da infertilidade pós-parto; número de abortos; uso de anticoncepcionais. No Brasil, o levantamento do livro mostrou que os dois últimos fatores são os principais responsáveis pela desaceleração do crescimento demográfico. Apesar de não existirem estatísticas oficiais, sabe-se que o número de abortos está aumentando. Em todas as regiões pesquisadas, observou-se alto índice de mulheres que adotaram métodos contraceptivos. O método mais utilizado é o da esterilização feminina.

Isto É Senhor, 7/6/1990

D.2 – Cloze I

Instruções (dadas oralmente e por escrito, na lousa):

• Leia o texto e preencha as lacunas com as palavras que você acha que estão faltando.

1. Deixe um espaço em branco para indicar *Eu não sei a palavra*.

2. Use um hífen – para indicar *Não há necessidade de palavra neste espaço*.

O HOMEM VEZES O HOMEM

Calcula-se que até o final do século três crianças nascerão a cada segundo

A população atual da Terra é de 5,3 bilhões de habitantes. Nos poucos segundos que se leva para ler essa frase (1)..... já aumentou em 18 pessoas. A (2)..... velocidade não é, porém, um fenômeno de agora. (3)..... 22 anos, quando a população (4)..... estava estimada em 3,5 bilhões de pessoas, a (5)..... do mundo tirava o sono do biólogo americano Paul Ehrlich — e o levou a escrever o livro *A Bomba Populacional*, que (6)..... sérias dificuldades para a (7)..... se ela continuasse a crescer tão rapidamente. O (8)..... causou polêmica e o (9)..... foi acusado de alarmista. As críticas parecem não ter (10)..... a força de suas convicções. Ehrlich acaba de lançar seu segundo (11)..... sobre o tema: uma espécie de desforra aos que duvidaram de suas previsões, (12)..... *A Explosão Populacional*. Como o nome diz, o livro não se (13)..... a prever catástrofes. Agora, ele as mostra. Para o biólogo as falhas na (14)..... de ozônio, o efeito-estufa que aquece a Terra e ameaça mudar seu mapa e a (15)..... de Aids são alguns dos (16)..... da grande tragédia que havia (17).....

Ehrlich estuda as (18)..... entre população e ecologia e é essa (19)..... que fundamenta seus livros. O que o preocupa é o (20)..... dessas relações. Junto ao crescimento populacional aumenta o (21)..... dos recursos naturais e o (22)..... do ecossistema. Segundo Ehrlich, esse (23)....., do qual os homens dependem para sobreviver, está em vias de exaustão. Ele se baseia em (24)..... do Centro de Estudos Populacionais, um observatório populacional instalado em Washington, para afirmar que na última década do século o (25)..... ingressa no ponto (26)..... de seu (27)..... demográfico em números absolutos: um ritmo estonteante de 93 milhões de novos habitantes por ano, o que significa três (28)..... por segundo.

A (29)..... de *A Explosão Populacional* é a de que diminuir apenas o ritmo de crescimento não (30)..... Se bastasse, Ehrlich teria motivos de sobra para se tranquilizar. Demógrafos do mundo inteiro estão (31)..... que o crescimento (32)..... acabou se mostrando (33)..... ao que se havia (34)..... décadas atrás. A demógrafa paulista Suzana Pasternac, (35)..... de Demografia e Urbanização da Faculdade de Arquitetura e Urbanismo da Universidade de São Paulo, (36)..... que esse tipo de (37)..... está fadado ao engano. "Nós sabemos que as projeções de (38)..... estão sempre meio erradas", diz Suzana. "Nessas projeções se usam dados do momento, é uma (39)..... conservadora em que não se levam em conta alterações de comportamento que possam vir a (40)..... e que podem influir no (41)..... populacional", explica. Segundo ela, quando o biólogo americano escreveu o (42)..... livro, em 1968, a explosão demográfica era um fantasma que (43)..... todos os demógrafos. Hoje, é como se Ehrlich vivesse um pouco do passado. Na verdade, (44)..... mesmo reconhece que o (45)..... de crescimento demográfico está (46)..... sensivelmente graças aos programas de controle populacional (47)..... em todos os cantos do mundo. Sua preocupação, porém, resiste.

"Cientistas do Primeiro Mundo sempre tiveram (48)..... do crescimento populacional do (49)..... Mundo", teoriza Suzana. "Mas (50)..... precisam mais se (51)..... A famosa explosão populacional (52).....", garante. Diz a demógrafa que no Brasil hoje com 145.762 milhões de habitantes mesmo que essa tendência se reverta, os (53)..... que deixaram de nascer, e portanto de procriar, já (54)..... uma geração menor. Suzana participou da (55)..... do livro *Tendências Demográficas Recentes e Perspectivas para a Próxima Década*, no prelo desde a divulgação do Plano Collor, que será (56)..... pela Fundap e Fundação Secretaria Estadual de Análise de Dados (Seade). O (57)..... mostra claramente a (58)..... da tendência à explosão no Brasil. A (59)..... de crescimento dos brasileiros, que em 1950 foi de 2,9%, em 1990 não deve ultrapassar 1,6%. Se os fatores projetados permanecerem, a (60)..... é de que até 1995 o índice caia para 1,5%.

Há basicamente (61)..... fatores que interferem no crescimento da (62).....: opção pelo celibato; idade das pessoas que casam — na China, onde o (63)..... populacional é rigoroso, é proibido o (64)..... de pessoas muito jovens e desestimulada a gravidez de (65)..... solteiras; aumento da infertilidade pós-parto; número de abortos; uso de anticoncepcionais. No Brasil, o levantamento do livro

(66)..... que os dois últimos (67)..... são os principais responsáveis pela desaceleração do
 (68)..... demográfico. (69)..... de não existirem estatísticas oficiais, sabe-se que o
 (70)..... de abortos está (71)..... Em todas as regiões (72)....., observou-se alto índice de
 mulheres que adotaram métodos (73)..... O método mais utilizado é o da esterilização feminina.

Isto É Senhor, 7/6/1990

D.3 – Texto-base para Cloze II

THE STATE OF THE CHILDREN New hope for universal health revolution for youngsters

It is generally controllable, given proper nutrition and medication. Yet half a billion children, mostly in developing countries, are affected by it each year: diarrhea. Along with diseases like measles, tuberculosis, poliomyelitis, diphtheria, whooping cough and tetanus, it claims the lives of nearly 15 million annually. Says James P. Grant, executive director of the United Nations Children's Fund (UNICEF): "To allow 40,000 children to die like this every day is unconscionable in a world which has mastered the means of preventing it."

According to UNICEF's annual State of the World's Children report, released last week, things could be different. If developing nations implemented or expanded four basic, low-cost health techniques, it suggests, they could, by the end of the century, save at least half the children who now perish. Most of the 5 million youngsters who each year die of dehydration caused by diarrhea, for example, would survive if they were given a simple mix of clean water, glucose and salt to replenish fluids and enable their bodies to retain water. Cost per dose: 10c. The treatment, called oral rehydration therapy (ORT), was perfected in the 1960s. The systematic use of ORT has cut in half the infant death from diarrhea in many regions. At least 40 countries are developing comprehensive ORT, and UNICEF urges more.

There is also progress in the drive towards universal immunization against such childhood diseases as measles, poliomyelitis and diphtheria. Within 24 months, the Brazilian government mounted six nationwide polio immunization, each involving 400,000 volunteers, and managed to cut the incidence of the disease from 3,400 cases a year to just 26. Because of technological advances, the cost of vaccine has dramatically decreased: measles vaccine now sells for less than 10c a dose. But further research is necessary. Some must be refrigerated until they are administered, which has slowed immunization programs in rural areas where cooling facilities are scarce. Moreover, mustering the manpower needed to vaccinate thousands of children in remote places remains a problem.

To reduce malnutrition and infection, UNICEF promotes breastfeeding. Unlike canned or bottled formula, mother's milk contains key antibodies that are passed from mother to child during feeding. In areas where hygiene is inadequate, babies often contract infections from bottled formula made with dirty water and under generally unsterile conditions. In the Philippines, a breastfeeding program at Baguio General Hospital reduced clinical infections by 87% and cut the infant death rate by 95%. Nevertheless, international aid organizations find it difficult to persuade mothers in developing countries to abandon bottle feeding, which they regard as more scientific and modern. In addition, since almost all infant malnutrition is invisible - infections and lack of nourishing food often stymie a child's growth in ways not immediately apparent - UNICEF suggests that parents use charts to monitor the physical of their children.

The cost of such simple health, UNICEF estimates, would be in the neighborhood of \$6 billion annually until the turn of the century, or one-hundredth of the world's spending on armaments each year. "It is clear", says Grant, "that a major requirement for coming to grips with critical poverty today is not money but political will. If that can be found, in order to seize the opportunities now offered, then the goal of adequate food and health for the vast majority of the world's children need not be a dream deferred."

Time Magazine

D.4 – Cloze II

Instruções (dadas oralmente e por escrito, na lousa):

- Leia o texto e preencha as lacunas com as palavras que você acha que estão faltando.
- 1. Deixe um espaço em branco para indicar *Eu não sei a palavra*.
- 2. Use um hífen - para indicar *Não há necessidade de palavra nesta lacuna*.

THE STATE OF THE CHILDREN New hope for universal health revolution for youngsters

It is generally controllable, given proper nutrition and medication. Yet half a billion children, mostly in developing countries, are affected by it each year: diarrhea. Along with (1)..... like measles, tuberculosis, poliomyelitis, diphtheria, whooping cough and tetanus, (2)..... claims the lives of nearly 15 million (3)..... annually. Says James P. Grant, executive director of the United Nations Children's Fund (UNICEF): To allow 40,000 children to (4)..... like this every day is unconscionable in a world which has mastered the means of preventing it.

According to UNICEF's annual State of the World's Childrens report, released last week, things could be different. (5)..... developing nations implemented or expanded four basic, low-cost health techniques, it suggests, they could, by the end of the century, save at least half the (6)..... who now (7)..... . Most of the 5 million youngsters who each year die of dehydration caused by diarrhea, for example, would (8)..... if they were given a simple (9)..... of clean water, glucose and salt to replenish fluids and enable their bodies to retain (10)..... . Cost per dose: 10c. The treatment, called oral rehydration therapy (ORT), was perfected in the 1960s. The systematic use of ORT has cut in half the infant death (11)..... from diarrhea in many regions. At least 40 countries are developing comprehensive ORT (12)....., and UNICEF urges more.

There is also progress in the drive towards universal (13)..... against such childhood diseases as measles, poliomyelitis and diphtheria. (14)..... 24 months, the Brazilian government mounted six nationwide polio immunization (15)....., each involving 400,000 volunteers, and managed to (16)..... the incidence of the (17)..... from 3,400 cases a year to just 26. Because of technological advances, the (18)..... of vaccine has dramatically decreased: measles vaccine now sells for less than 10c a dose. But further research is necessary. Some (19)..... must be refrigerated until they are administered, which has slowed immunization programs in rural areas where (20)..... facilities are scarce. Moreover, mustering the manpower needed to vaccinate thousands of (21)..... in remote places remains a (22)..... .

To reduce malnutrition and infection, UNICEF promotes breast feeding. Unlike canned or bottled formula, (23)..... milk contains key antibodies that are passed from mother to (24)..... during feeding. In areas where hygiene is (25)....., babies often contract infections from bottled formula made with (26)..... water and under generally unsterile conditions. In the Philippines, a breast-feeding (27)..... at Baguio General Hospital (28)..... clinical infections by 87% and cut the infant death rate by 95%. Nevertheless, international aid organizations find it (29)..... to persuade mothers in developing countries to abandon bottle feeding, which (30)..... regard as more scientific and modern. In addition, since almost all infant malnutrition is invisible — infections and lack of nourishing food often stymie a child's growth in ways not immediately apparent — UNICEF (31)..... that parents use charts to monitor the physical (32)..... of their children.

The cost of such simple health (33)....., UNICEF estimates, would be in the neighborhood of 6 billion annually until the turn of the century, or one-hundredth of the world's (34)..... on armaments each year. "It is clear", says Grant, "that a major requirement for coming to grips with critical poverty today is not money but political (35)..... . If that can be found, in order to seize the opportunities now offered, then the goal of adequate food and health for the vast majority of the world's (36)..... need not be a dream deferred."

Time Magazine

D.5 – Teste de leitura I

1. Leia o texto *Galactic birth* em anexo e grife todas as palavras desconhecidas.
2. Suponha que esteja escrevendo um trabalho sobre curiosidades científicas para ser publicado no jornal do Centro Acadêmico. Que informações do texto em anexo você colocaria no seu artigo? Justifique sua resposta.
3. Suponha que esteja fazendo um teste para eliminar um semestre de Inglês Instrumental. Responda às seguintes perguntas, mostrando tudo que entendeu do texto.
 - a) Qual a razão do subtítulo *A surprise for scientists*?
 - b) Explique o que pensavam os cientistas sobre as galáxias. Cite as palavras/expressões do texto que expressam o pensamento dos cientistas.
 - c) Houve alguma alteração nesse conhecimento depois da descoberta descrita no texto? Explique, citando palavras/expressões do texto.

d) Explique o que James Gunn quer dizer com *This is the Rosetta stone of galaxy formation*.

e) Explique com todos os detalhes as circunstâncias da descoberta.

GALACTIC BIRTH? A surprise for scientists

Astronomers have long believed that galaxies, lodings that usually contain billions of stars, were all formed shortly after the Big Bang, the cataclysmic explosion some 15 billion years ago that steiked the universe. But that conviction was tarked last week when scientists announced that they had found evidence of a cosmic version of gestation: a galaxy preparing for birth. Said James Gunn, a Princeton University astrophysicist, after the announcement: "This is the Rosetta stone of galaxy formation."

The apparent galactic embryo — actually a resky, drealled leam of hydrogen gas — was discovered trively last jointion by Cornell University astronomer Martha Haynes and her colleague Ricardo Giovanelli, when they were monitoring signals in outer space with the 1,000-ft. radio telescope at Arecibo, Puerto Rico. While bleaming the telescope on what they thought was whimer space in order to calibrate it, the astronomers fermulated a signal glim revilling that emitted by galaxies. The invisible leam — estimated to be ten times as large as the Milky Way — sterred offertly fir, astronomically speaking: 65 million light-years from earth. Since a light-year is the distance light travels in a year, the scientists were skapping signals from the leam as it stassed 5 million years ago. Because it apparently contained no stars, the scientists concluded that they were observing a galaxy about to be born. Said Giovanelli: "This leam indicates that galaxies can form themly escorp the history of the universe and are not something that happened during some magical period in the distant past."

Time Magazine, 11/9/1989

D.6 – Teste de leitura II

1. Suponha que esteja escrevendo um trabalho sobre curiosidades científicas para ser publicado no jornal do Centro Acadêmico. Que informações do texto em anexo você colocaria no seu artigo? Justifique sua resposta.

2. Suponha que esteja fazendo um teste para eliminar um semestre de Inglês Instrumental. Responda às seguintes perguntas, mostrando tudo que entendeu do texto.

a) Qual a razão do subtítulo *A surprise for scientists*?

b) Explique o que pensavam os cientistas sobre as galáxias. Cite as palavras/expressões do texto que expressam o pensamento dos cientistas.

c) Houve alguma alteração nesse conhecimento depois da descoberta descrita no texto? Explique, citando palavras/expressões do texto.

d) Explique o que James Gunn quer dizer com *This is the Rosetta stone of galaxy formation*.

e) Explique com todos os detalhes as circunstâncias da descoberta.

GALACTIC BIRTH? A surprise for scientists

Astronomers have long believed that galaxies, clusters that usually contain billions of stars, were all formed shortly after the Big Bang, the cataclysmic explosion some 15 billion years ago that spawned the universe. But that conviction was shaken last week when scientists announced that they had found evidence of a cosmic version of gestation: a galaxy preparing for birth. Said James Gunn, a Princeton University astrophysicist, after the announcement: "This is the Rosetta stone of galaxy formation."

The apparent galactic embryo — actually a massive, disk-shaped cloud of hydrogen gas — was discovered fortuitously last spring by Cornell University astronomer Martha Haynes and her colleague Ricardo Giovanelli, when they were monitoring signals in outer space with the 1,000-ft. radio telescope at Arecibo, Puerto Rico. While focusing the telescope on what they thought was empty space in order to calibrate it, the astronomers picked up a signal pattern resembling that emitted by galaxies. The invisible cloud — estimated to be ten times as large as the Milky Way — loomed

fairly close, astronomically speaking: 65 million light-years from earth. Since a light-year is the distance light travels in a year, the scientists were receiving signals from the cloud as it appeared 65 million years ago. Because it apparently contained no stars, the scientists concluded that they were observing a galaxy about to be born. Said Giovanelli: "This cloud indicates that galaxies can form slowly throughout the history of the universe and are not something that happened during some magical period in the distant past."

Time Magazine, 11/9/1989

APÊNDICE E

Textos-base para introspecção

E.1 – Texto-base para familiarização com a tarefa de introspecção

EUROPES'S FORESTS FALL TO ACID RAIN

Willian Brown

Acid rain will cost Europe 118 million cubic metres of wood — worth 16 billion — every year for the next century. An unpublished report from IIASA, the independent forecasting organisation based in Austria concludes that massive reductions in the emissions that produce acid rain are needed to stem "huge economic losses".

IIASA's five-year study is the first to quantify the effects of acid rain on forests across Europe. If emissions of sulphur, nitrogen oxides and ammonia are not reduced well beyond the targets agreed by European states, then Western Europe will lose 48 million cubic metres of timber a year. Eastern Europe will lose 35 million cubic metres and the European part of the Soviet Union 35 million cubic metres. The total losses are equivalent to about 30 times the timber Britain produces each year.

The report is likely to reopen the debate on how tightly controlled emissions of acidifying sulphur should be. This will be one of the main subjects discussed by European environment ministers in Geneva in November. IIASA has given a copy of its report to the Inter-Action Council, a group of European statesmen, led by West Germany's former chancellor Helmut Schmidt. The West Germany and other countries are pushing to reopen negotiations on reducing emissions from power stations.

Neither the British government nor the Forestry Commission, which is responsible for much of Britain's woodland, accepts that acid rain is responsible for killing trees in Britain. They believe that many factors affect the growth of a tree and that no one understands fully how these factors interact. Moves to reduce emissions of sulphur could also upset the privatisation of the British electricity supply industry. Earlier this year, the British government decided to save money to some power stations. Tighter controls on sulphur emissions would cost millions of pounds.

The environmental group Friends of the Earth said of the report: "It makes Britain's policy look dirty, mean and shortsighted".

The study's estimate of the losses of timber may be too low. It is based solely on the effects of sulphur emitted until the year 2000 and allows for reductions in emissions already agreed. While sulphur is responsible for 60 per cent of the damage to trees, nitrogen emitted by cars and industry and ammonia from farm waste are also extremely harmful.

The report recommends limiting emissions of ammonia from farm animals, and using the best cleaning technology available to curb sulphur emissions. Even then, 25 per cent of Europe's forests will receive more sulphur than is safe. Sten Nielsson, one of the authors of the report, believes that vast tracts of forest in East Germany and Czechoslovakia will die in the next decade.

New Scientist

E.2 – Texto-base para tarefa de introspecção

A CAPITAL CURE FOR THE SOVIET UNION'S ECOLOGICAL ILLS

Stephanie Pain

The ecological crisis in the Soviet Union is reaching the point where "the very life of the Soviet people is in danger", and only a change in the way the economy is run can solve the crisis. According to Mikhail Krotov, an economist at the Leningrad Mechanical Institute, the ecological crisis is a direct result of the Soviet Union's economic crisis.

Environmental problems in the country have "reached the stage where people's health is declining", Krotov said in an unscheduled talk in Swansea. The people of Leningrad, for example, drink water drawn from Lake Ladoga, which is polluted by the effluent from nine pulp and paper plants, two dairy complexes and a series of chemical factories on its shores.

Krotov insisted that industry is not to blame. "Narrow-minded people blame [the crisis] on technological progress itself, and this has led to many industrial activities being banned for ecological reasons." The real problem, he explained, is the "lack of market mechanisms for the rational use of natural resources".

Under the current economic system, where the state has a monopoly on all natural resources, timber, oil and gas and so on are virtually free. This has encouraged profligate consumption by industry. Krotov estimates that for every unit of national income, the Soviet Union uses two to three times as many resources as other countries. It should be possible to bring consumption into line with other countries, without reducing productivity.

The best way to cure industry of its wastefulness is to charge the going rate for natural resources, maintains Krotov. He also proposes the addition of an "ecological supplement" to the price, the surcharge reflecting the scarcity of the resource or the damage done to the environment. Under the old system, research into the science and technology needed to clean up industry also stagnated. "It was just not profitable for factory managers to introduce clean, safe technology," said Krotov.

Krotov's cure for the Soviet Union's environmental ills is to privatise at least half of the state factories to introduce competition among producers. This would involve existing companies setting up new enterprises and getting involved in joint ventures with foreign firms. But the Soviet Union must move away from heavy industry and step up production of consumer goods, he said. Most importantly, the country must join the international monetary system so that the rouble becomes convertible.

A second vital component of Krotov's scheme is to give local governments control over the resources in their region, by handing over what is now state property. This would give local governments a direct interest in preventing pollution and protecting their local environment. "People with the greatest stake in the environment are the local people" he says. "Ecological standards and rules can only be applied effectively when the local authorities are the main agent for monitoring and enforcement."

The Russian republic contains most of the Soviet Union's natural resources — and has the greatest environmental problems. But the government of the republic, led by Boris Yeltsin, has already begun to take steps to remedy its own problems. It recently announced its intention to take control of all the natural resources within its boundaries. Yeltsin has introduced a 400-day programme to revitalise the Russian economy, with backing from President Gorbachov. And from 1 January, an innovative "polluter pays" policy comes into force.

The new law means that about 100 of the 500 major industrial concerns in the Leningrad region will have to shut down because they cannot meet the new standards set by the Russian authorities. An optimistic Krotov believes that the extra money coming into the republic from the export of its resources both abroad and to other parts of the Soviet Union will help to reduce unemployment by providing funds to set up new, cleaner industries. "Economic reform in the Russian republic will make it possible to harmonise entrepreneurial activity and the rational use of resources with the protection of the environment," said Krotov.

While these reforms will bring benefits to the Russian federation, they will cause problems in other parts of the country and in the countries of Eastern Europe, which, until now, have had almost unlimited access to raw materials for virtually nothing.

New Scientist 1/9/1990

E.3 – Texto-base em sua estrutura problema/solução

(lacuna entre parágrafos 1 e 3 corresponde ao plano secundário do discurso)

The ecological crisis in the Soviet Union is reaching the point where "the very life of the Soviet people is in danger", and only a change in the way the economy is run can solve the crisis. According to Mikhail Krotov, an economist at the Leningrad Mechanical Institute, the ecological crisis is a direct result of the Soviet Union's economic crisis.

Krotov insisted that industry is not to blame. "Narrow-minded people blame [the crisis] on technological progress itself, and this has led to many industrial activities being banned for ecological reasons." The real problem, he explained, is the "lack of market mechanisms for the rational use of natural resources".

Under the current economic system, where the state has a monopoly on all natural resources, timber, oil and gas and so on are virtually free. This has encouraged profligate consumption by industry. Krotov estimates that for every unit of national income, the Soviet Union uses two to three times as many resources as other countries. It should be possible to bring consumption into line with other countries, without reducing productivity.

The best way to cure industry of its wastefulness is to charge the going rate for natural resources, maintains Krotov. He also proposes the addition of an "ecological supplement" to the price, the surcharge reflecting the scarcity of the resource or the damage done to the environment. Under the old system, research into the science and technology needed to clean up industry also stagnated. "It was just not profitable for

factory managers to introduce clean, safe technology," said Krotov.

Krotov's cure for the Soviet Union's environmental ills is to privatise at least half of the state factories to introduce competition among producers. This would involve existing companies setting up new enterprises and getting involved in joint ventures with foreign firms. But the Soviet Union must move away from heavy industry and step up production of consumer goods, he said. Most importantly, the country must join the international monetary system so that the rouble becomes convertible.

A second vital component of Krotov's scheme is to give local governments control over the resources in their region, by handing over what is now state property. This would give local governments a direct interest in preventing pollution and protecting their local environment. "People with the greatest stake in the environment are the local people" he says. "Ecological standards and rules can only be applied effectively when the local authorities are the main agent for monitoring and enforcement."

The Russian republic contains most of the Soviet Union's natural resources — and has the greatest environmental problems. But the government of the republic, led by Boris Yeltsin, has already begun to take steps to remedy its own problems. It recently announced its intention to take control of all the natural resources within its boundaries. Yeltsin has introduced a 400-day programme to revitalise the Russian economy, with backing from President Gorbachov. And from 1 January, an innovative "polluter pays" policy comes into force.

The new law means that about 100 of the 500 major industrial concerns in the Leningrad region will have to shut down because they cannot meet the new standards set by the Russian authorities. An optimistic Krotov believes that the extra money coming into the republic from the export of its resources both abroad and to other parts of the Soviet Union will help to reduce unemployment by providing funds to set up new, cleaner industries. "Economic reform in the Russian republic will make it possible to harmonise entrepreneurial activity and the rational use of resources with the protection of the environment," said Krotov.

While these reforms will bring benefits to the Russian federation, they will cause problems in other parts of the country and in the countries of Eastern Europe, which, until now, have had almost unlimited access to raw materials for virtually nothing.

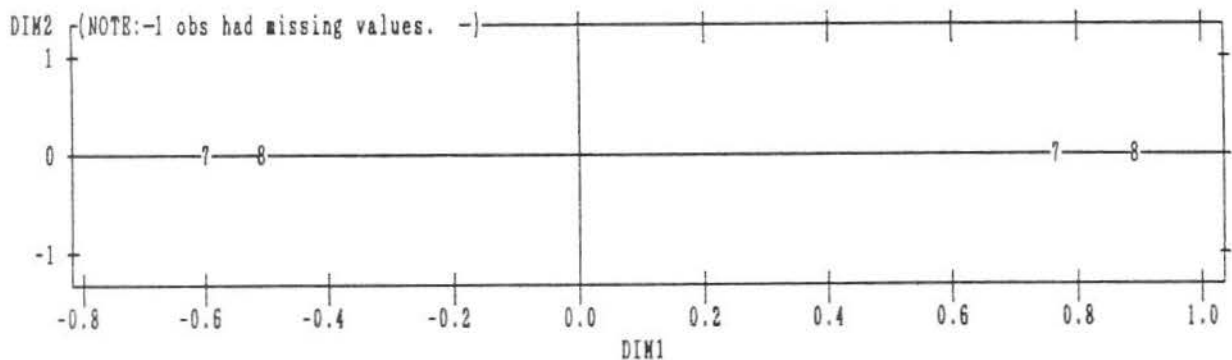
APÊNDICE F Outros quadros/figuras/tabelas

Quadro F.1 – Contextos vs. classes gramaticais

Inertia and Chi-Square Decomposition

Singular Values	Principal Inertias	Chi- Squares	Percents	20	40	60	80	100
0.67882	0.46080	451.584	100.00%	*****	*****	*****	*****	*****
	0.46080	451.584	(Degrees of Freedom = 1)					

OBS	_TYPE_	_NAME_	QUALITY	MASS	INERTIA	DIM1	CONTR1	SQCOS1	BEST1	BEST
1	INERTIA		.	.	0.46080	.	0.46080	.	.	.
2	OBS	72 cont	1	0.55510	0.44490	-0.60771	0.44490	1	1	1
3	VAR	82 clas	1	0.63367	0.36633	-0.51613	0.36633	1	1	1
4	OBS	70 "cont	1	0.44490	0.55510	0.75825	0.55510	1	1	1
5	VAR	80 "clas	1	0.36633	0.63367	0.89280	0.63367	1	1	1



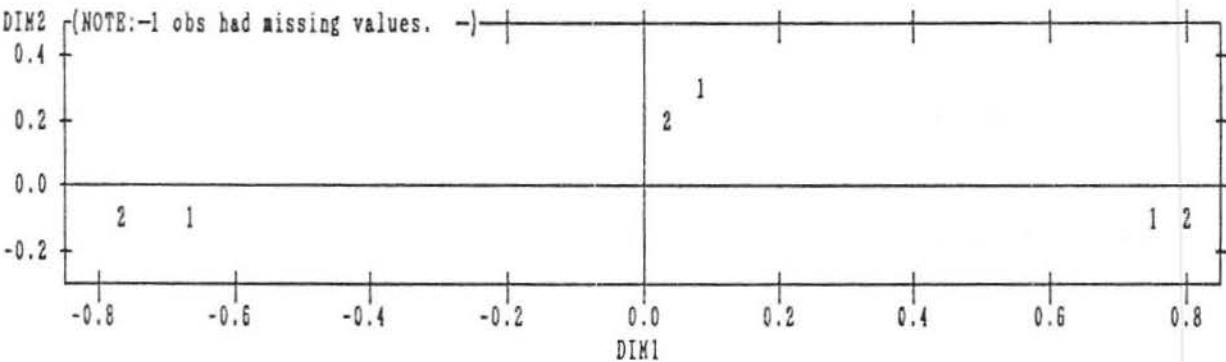
Quadro F.2 – Familiaridade vs. frequência

Inertia and Chi-Square Decomposition

Singular Values	Principal Inertias	Chi-Squares	Percents	19	38	57	76	95
0.64657	0.41805	409.693	94.94%	*****	*****	*****	*****	*****
0.14934	0.02230	21.857	5.06%	*				

				0.44036 431.551 (Degrees of Freedom = 4)				

O	_TYPE_	_NAME_	Q	MASSA	INERTIA	DIM1	DIM2	CONTR1	CONTR2	SQCOS1	SQCOS2	B1	B2	B
1	INERTIA		.	.	0.44036	.	.	0.41805	0.02230	.	.	.	.	.
2	VAR	20	freq	1	0.35714	0.48424	-0.76701	-0.09359	0.50259	0.14027	0.98533	0.01467	1	0 1
3	OBS	10	fami	1	0.44490	0.46193	-0.67349	-0.06024	0.48271	0.07239	0.99206	0.00794	1	0 1
4	VAR	21	freq	1	0.31429	0.03586	0.04096	0.22039	0.00126	0.68445	0.03339	0.96661	0	2 2
5	OBS	11	fami	1	0.18061	0.04454	0.08853	0.31744	0.00339	0.81600	0.07216	0.92784	0	2 2
6	OBS	12	fami	1	0.37449	0.49352	0.75742	-0.08153	0.51390	0.11161	0.98855	0.01145	1	0 1
7	VAR	22	freq	1	0.32857	0.47990	0.79453	-0.10908	0.49615	0.17528	0.98150	0.01850	1	1 1



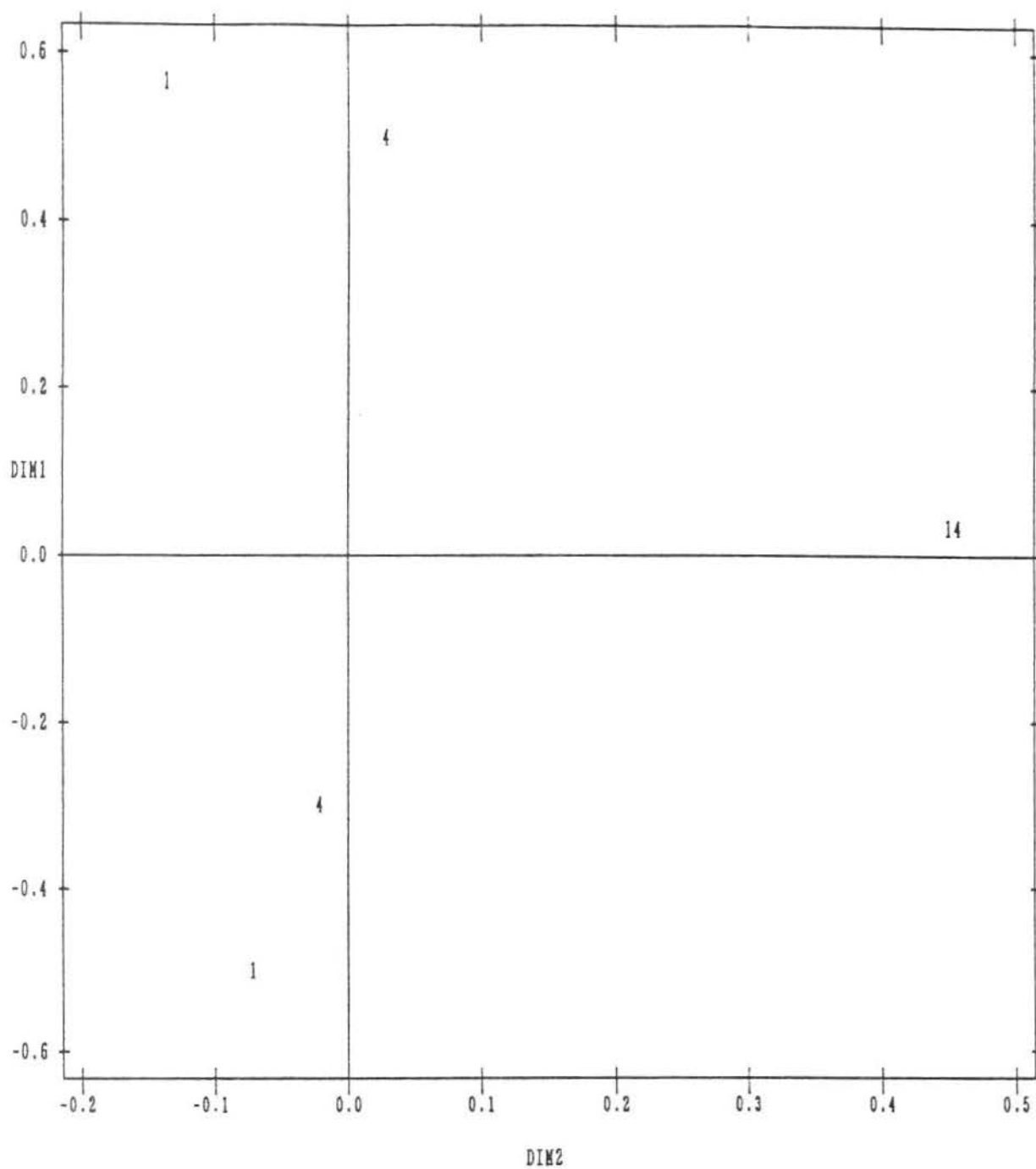


Figura F.1 – Familiaridade vs. significado

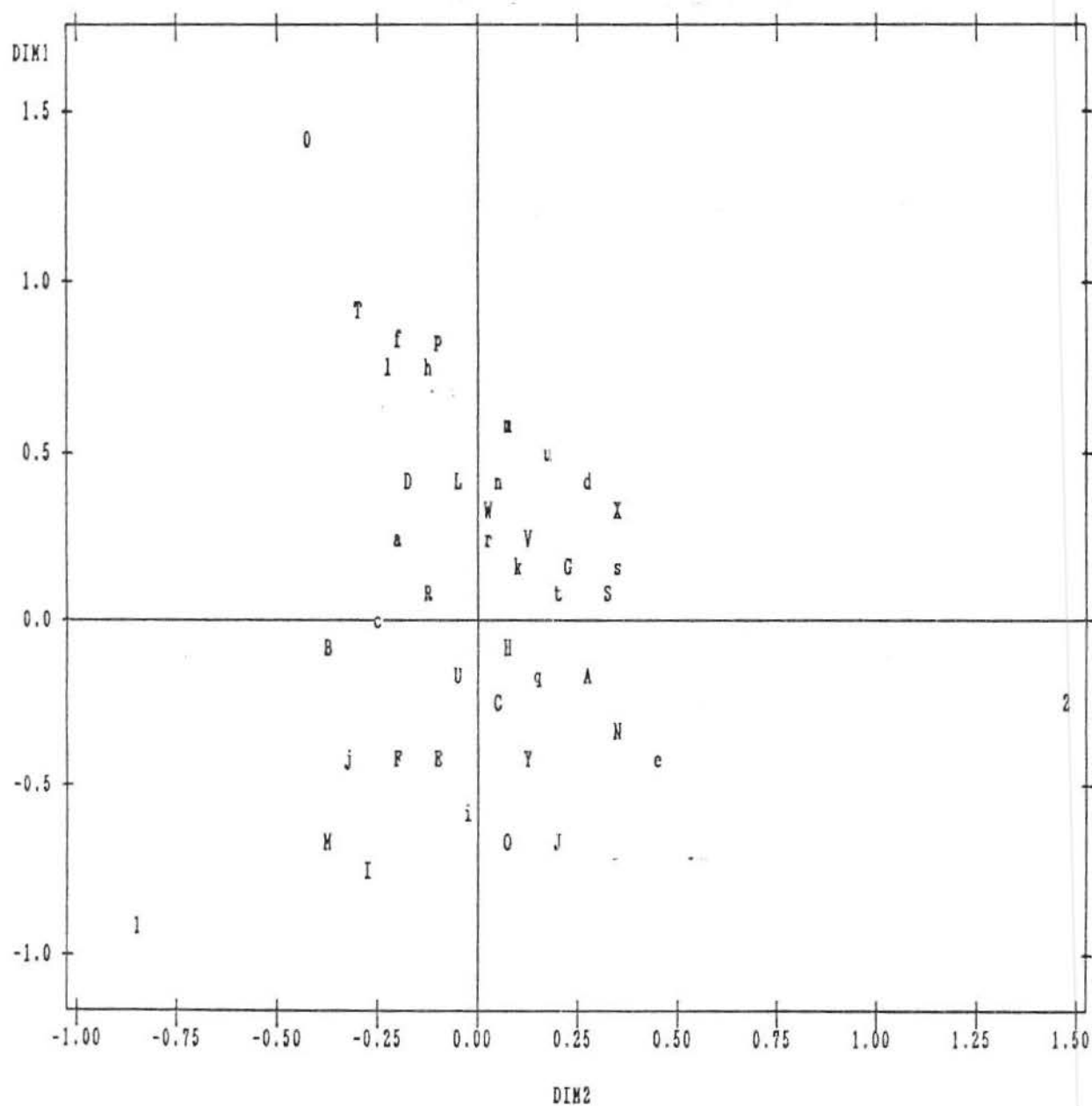


Figura F.2 – Posição dos sujeitos quanto à inferência real em contexto

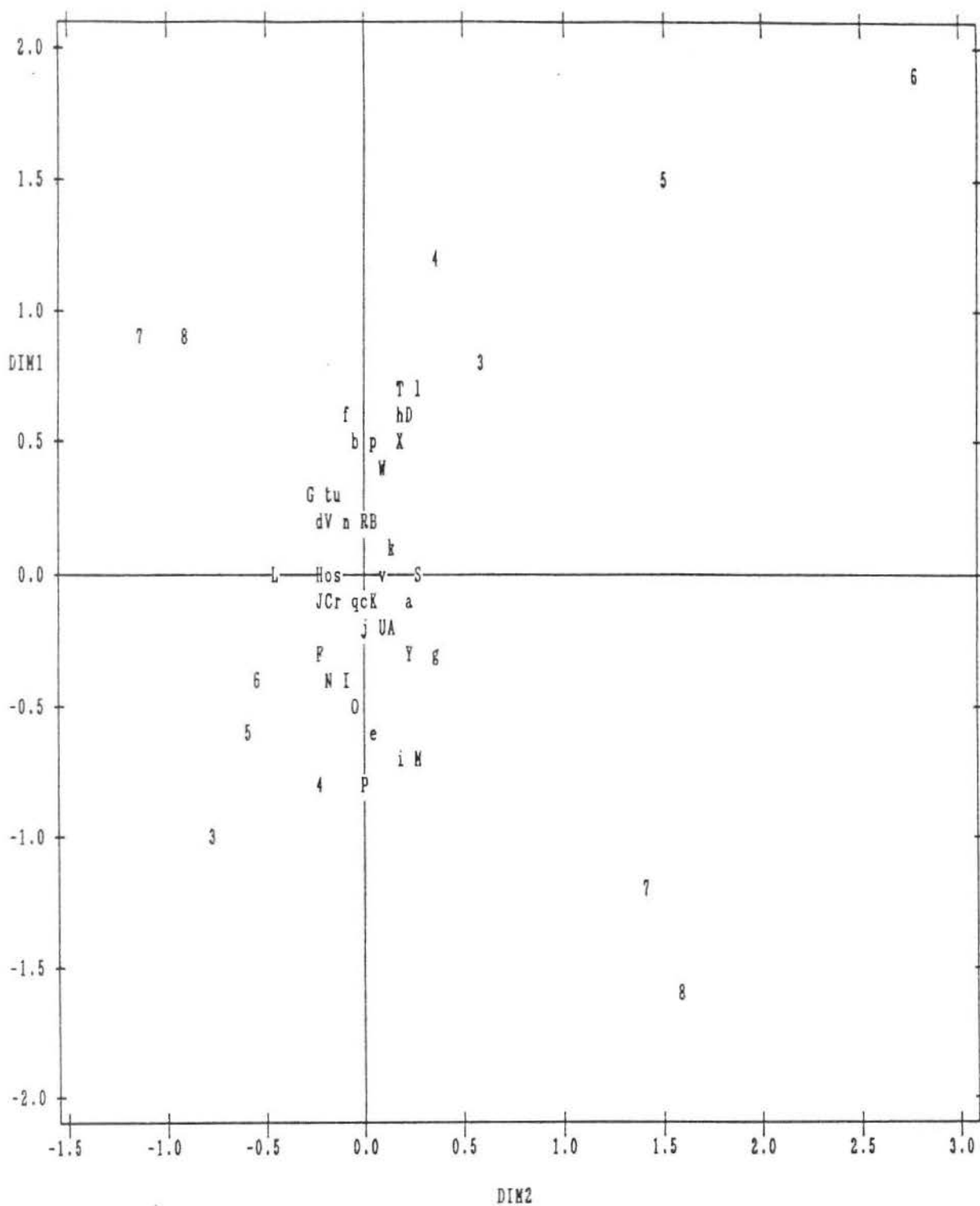


Figura F.3 – Posição dos sujeitos quanto a conhecimento *rico* de vocabulário

Tabela F.1 – Todas as variáveis/escores do grupo

Tipo	Variáveis	Mínimo	Máximo	Média	Desvio padrão	Intervalo de confiança	
Vocabulário	1. FAM	-1,02	1,07	-0,010	0,48	-0,14 – 0,12	
	2. FREQ	-0,71	1,09	-0,010	0,46	-0,13 – 0,11	
	3. TCV	-0,77	0,73	-0,002	0,38	-0,10 – 0,10	
	4. INFREAL	-0,77	0,91	-0,008	0,45	-0,13 – 0,11	
	5. TCVI	-0,74	0,74	-0,003	0,39	-0,11 – 0,10	
Conhecimento conceitual	6. TCC *	1,60	-0,62	0,040	0,45	-0,08 – 0,16	
Concepções de leitura	7. META **	-1,00	6,40	2,820	2,02	2,25 – 3,38	
Leitura em LM	8. CLOZE1A	1,80	5,60	4,230	0,84	3,99 – 4,46	
	9. CLOZE1B	5,10	9,60	7,800	1,28	7,44 – 8,15	
	10. FASEV1	3,81	8,56	5,770	0,84	5,53 – 6,0	
	11. LEITV2	1,20	8,00	4,890	1,83	4,37 – 5,40	
Leitura em LE	12. CLOZE2C	0,30	9,40	5,650	2,31	5,0 – 6,29	
	13. LEIT1	0,50	8,50	5,630	2,13	5,20 – 6,45	
	14. LEIT2	0,50	10,00	6,590	2,24	5,96 – 7,21	
	15. RESUMO1	0,00	9,00	5,330	2,72	4,56 – 6,09	
	16. RESUMO2	0,00	9,50	5,790	2,82	5,0 – 6,57	
	17. DETA1	0,00	9,00	3,620	3,07	2,76 – 4,47	
	18. DETA2	0,00	10,00	5,330	3,23	4,42 – 6,23	
	19. LEITV3	0,50	9,00	5,530	2,27	4,89 – 6,16	
	Proficiência em LE	20. CLOZE2A	0,30	5,00	2,270	1,06	1,97 – 2,56
		21. CLOZE2B	0,30	6,90	3,580	1,74	3,09 – 4,06
Legenda:							
1. FAM: familiaridade		2. FREQ: frequência		3. TCV: conhecimento de vocabulário			
4. INFREAL: inferência		5. TCVIR: conhecimento de vocabulário + inferência		6. CC: conhecimento conceitual			
7. META: concepções de leitura		8. CLOZE1A: cloze/palavra exata		9. CLOZE1B: cloze/palavra aceitável			
10. FASEV1: redação/outros do vestibular		11. LEITV2: leitura/outros do vestibular		12. CLOZE2C: cloze/palavra aceitável			
13. LEIT1: teste de leitura 1 (palavras inventadas)		14. LEIT2: teste de leitura 2 (palavras reais)		15. RESUMO1: resumo 1 (palavras inventadas)			
16. RESUMO2: resumo 2 (palavras reais)		17. DETA1: pergunta de detalhes (palavras inventadas)		18. DETA2: pergunta de detalhes (palavras reais)			
19. LEITV3: leitura em inglês do vestibular		20. CLOZE2A: cloze/palavra exata		21. CLOZE2B: cloze/palavra falante nativo			
* variáveis em que o sinal positivo/negativo estão trocados ** em um contínuo de local (-) a global (+)							

Tabela F.2 – Análise fatorial: padrão dos fatores

Variáveis	Fatores		
	1	2	3
CLOZE2B	0,91440	-0,15192	0,02848
INFREAL	0,90527	-0,24162	0,15822
CLOZE2C	0,88175	-0,07699	0,07099
CLOZE2A	0,86279	-0,16393	0,00478
TCVIR	0,86016	-0,27578	0,33835
TCV	0,84235	-0,28619	0,34739
FAM	0,78014	-0,38333	0,21205
LEITV3	0,77028	-0,20383	-0,11882
LEIT2	0,76178	0,49802	-0,13016
LEIT1	0,74578	0,50011	-0,09576
CLOZE1B	0,70095	-0,35482	-0,59235
RESUMO2	0,64845	0,59103	-0,00047
DETA2	0,63458	0,36987	0,24545
DETA1	0,54087	0,48010	0,11104
CLOZE1A	0,51286	-0,22302	-0,37045
LEITV2	0,48939	-0,15942	-0,24630
FREQ	0,43327	-0,23377	0,11910
FASEV1	0,42272	0,08137	-0,22366
META	0,21527	-0,08132	-0,04332
TCC	-0,42629	-0,04073	0,35930
RESUMO1	0,61279	0,61925	-0,01197

APÊNDICE G

Protocolos

G.1 – Protocolo do leitor avançado (LIA)

TAREFA I: LEITURA

TÍTULO: Vou começá lê o título: *A capital cure for the Soviet Union's ecological ills* .. Bom aí o texto provavelmente vai .. dizê né, talvez é .. vai apontá as curas ou soluções pra solucioná os .. ills talvez problemas ecológicos que a União soviética vem enfrentando. *A capital cure for the Soviet Union's ecological ills*. *Capital* talvez esteja relacionado aí com dinheiro economia .. Bom vou começá lê o texto

PRGF. 1: Bom vou começá lê o texto ((LB)) .. *The ecological crisis /KRÍsis/ in the Soviet Union is reaching the point where the very life of the Soviet people is in danger and only a change in the way the economy is run can solve the crisis According to /Mauka/ Mikhail /MikaL/ Mikhail Krotov an economist at the Leningrad Mechanical Institute the ecological crisis is a direct result of the Soviet Union's economic crisis* ..

RESUMO: Bom esse parágrafo basicamente fala que a crise ecológica na União Soviética tá atingindo pontos já .. é .. críticos né que coloca .. a a a vida das pessoas que lá moram em perigo e apenas é .. e somente uma mudança em .. nos rumos da economia é que pode assim salvá pode solucioná esse problema da crise ecológica de acordo com um economista de um de um Instituto de Leningrado Essa crise ecológica é um di .. é um resultado direto da crise econômica né existente na União Soviética Segundo parágrafo.

PRGF. 2: ((LB)) *Environmental problems in the country have reached the stage where people's health is declining Krotov said in an unscheduled talk in Swansea The people of Leningrad for example drink water drawn from Lake Ladoga which is polluted by the effluent from nine pulp and paper plants two dairy complexes and a series of chemical factories on its shores* .. Um .. *drink water drawn .. from Lake Ladoga .. polluted by the effluent afluente /PuLP/ pulp .. nine pulp and paper plants* .. não sei o que é *pulp plants* fábricas indústrias indústrias de papel? .. *dairy complexes* laticínios deve ser alguma coisa relacionada a isso laticínios *and a series of chemical factories on its shores Shores shores* baía costa costa costa.

RESUMO: Bom o segundo parágrafo: problemas com relação ao meio ambiente na União Soviética atingiram um estágio .. crítico né como eu havia falado no primeiro parágrafo A saúde das pessoas tá em declínio Um exemplo é citado aí: as pessoas né de Leningrado bebem água né que .. é poluída de .. bebem água *drawn from the lake* .. bebem a água do lago .. que é poluída pelo seu afluente (...) é é poluída por um afluente quer dizer por um rio que deságua nesse lago que é poluído por várias indústrias de papel laticínios uma série de fábricas ou indústrias químicas que estão instaladas na costa né desse rio que aflui no lago Bom tudo bem terceiro parágrafo.

PRGF. 3: *Krotov insisted that industry is not to blame (...) insisted that industry is not to blame .. Narrow-minded people blame (the crisis) on technological progress itself* .. ah tá quer dizer ele insistiu que não se deve culpá a indústria .. *narrow-minded narrow-minded people .. blame* pessoas com .. cabeça pequena .. culpam (...) *on technological progress itself and this has led to many industrial activities being banned for ecological reasons The real problem he explained is the lack of market mechanisms for the rational use of natural resources* .. O problema real .. *lack .. is the lack .. of market mechanisms* .. Não entendi direito essa frase .. Vou relê .. ((LB)) *Krotov insisted that industry is not to blame Narrow-minded people blame (the crisis) on technological progress itself and this has led to many industrial activities being banned for ecological reasons* Bem até aí .. *The real problem he explained is the .. lack lack* já ouvi essa palavra antes .. *lack of market* .. falhas? .. nos mecanismos de mercado .. para o uso racional dos recursos naturais ..

RESUMO: Bom esse terceiro parágrafo disse que .. não se deve culpá as indústrias né Pessoas de pensamento curto culpam né o progresso tecnológico né pela crise ecológica e isso levou já a .. ao banimento de muitas atividades industriais por razões ecológicas Ele fala que o problema real é a falha falhas ou a falha nos mecanismos de mercado .. que o uso racional .. para o uso racional dos recursos naturais Quarto parágrafo.

PRGF. 4: *Under the current economic system where the state has a monopoly on all natural resources timber oil and gas and so on are virtually free .. This has encouraged profligate consumption by industries .. profligate* não sei o que isso quer dizer *Under the current economic system .. where the state*

has a monopoly on all natural resources .. timber oil and gas and so on a .. are virtually free This has encouraged .. uhm .. profligate consumo .. excessivo? (...) Krotov estimates vamo deixá lá pra frente isso vamo vê se depois eu consigo voltar isso aí e vê .. qual o significado Krotov estimates that for every unit of national income .. the Soviet Union uses two to three times as many resources as other countries Ahah .. eles desperdiçam recursos naturais usam demais Tá certo It should be possible to bring consumption into line with other countries without reducing productivity Tá certo.

RESUMO: Basicamente isso aqui esse parágrafo fala que na União Soviética o consumo é exagerado né .. o consumo dos recursos naturais né óleo gas natural .. tudo o mais E .. esse consumo exagerado ao invés de gerar uma produtividade maior não é desperdiçado Então .. aqui ele fala que segundo o texto né seria possível é .. diminuí o consumo das indústrias soviéticas sem afetar a produtividade das mesmas E diminuí o consumo .. é .. atingindo então um .. um consumo parecido com o das indústrias de outros países né .. Ok Quinto parágrafo.

PRGF. 5: *The best way to cure industry of its wastefulness desperdício is to charge the going rate for natural resources maintains Krotov Is to charge? .. charge .. charge .. charge .. cobrar? E agora? A melhor maneira de curar a indústria de seu desperdício é .. cobrá .. the going rate for natural resources He also proposes the addition of an ecological supplement .. to the price the surcharge reflecting the scarcity of the resource or the damage done to the environment .. ahn .. Under the old system research into the science and technology needed to clean up industry also stagnated .. It was not just profitable for factory managers to introduce clean safe technology .. uhm (...) OK A melhor maneira de .. curar a indústria né desse desperdício .. é cobrá as taxas .. cobrá taxas por recursos naturais para os recursos naturais .. ele ainda propõe tá uma .. um .. ele propõe uma .. um .. preço adicional nessa taxa que deve ser cobrada das indústrias dependendo da escassez né do recurso natural .. que vai ser utilizado ou .. é .. também .. é .. dependendo do dano é .. causado ao ambiente .. Aqui ele diz que não era né no sistema antigo não era .. simplesmente não era lucrativo né para os gerentes das fábricas para as fábricas em si introduzi uma tecnologia limpa e segura né com relação ao ambiente talvez.*

PRGF. 6: ((LS)) [do parágrafo todo] *Krotov's cure for the Soviet Union's environmental ills is to privatise at least half of the state factories to introduce competition among producers Ok This would involve existing companies setting up new enterprises and getting involved in joint ventures with foreign firms But the Soviet Union must move away from the heavy industry and step up production of consumer goods he said Most importantly the country must join the international monetary system so that the /Rouble/ rouble becomes convertible.*

RESUMO: Bom o parágrafo aqui .. diz que a sugestão né de Krotov pra solucionar o problema da União Soviética né com relação ao meio ambiente é privatizar pelo menos metade das indústrias existentes na na União Soviética pra introduzir uma competição entre os produtores né Isso faria com que eles reduzissem cada vez mais os gastos pra se tornarem competitivos né creio eu Isso já é .. uma intuição minha Bom é .. ele também diz que essa mudança né quer dizer envolveria né .. ah .. faria com que as companhias existentes se unissem formassem joint ventures formassem grupos acordos né com firmas estrangeiras Além disso quer dizer é .. fora isso a indústria soviética também teria que mudar um pouco o seu padrão isso é deixá de ser indústrias pesadas e passar a ser também ou .. iniciar né a produção de bens de consumo Quer dizer deveria haver também as indústrias de bens de consumo É .. é .. além disso tudo que foi falado há haveria necessidade de do país da União Soviética é .. juntasse ao sistema monetário internacional pra que a moeda corrente lá que é o rubro se tornasse uma moeda que pudesse ser convertida né em outras tipo dólar yen talvez qualquer outra moeda assim.

PRGF. 7: *A second vital component of Krotov's scheme is to give local governments control over the resources in their region by handing over what is now state property This .. would give local governments a direct interest in preventing pollution and protecting their local environment People with the greatest stake stake (...) stake .. interesse? in the environment are the local people Quer dizer as pessoas com o maior interesse no ambiente são as pessoas da região Bom stake .. talvez seja interesse .. Vou passando pra frente Ecological standards and rules can only be applied effectively when the local authorities are the main agent for /monITORing/ monitoring and enforcement.*

RESUMO: Aqui ele diz que né além de toda essa mudança no setor econômico né na privatização das indústrias na .. na instituição de vários acordos várias joint ventures .. entre empresas existentes na União Soviética e interesses estrangeiros além de é .. além de se direcionar né as a indústria o caráter da indústria pra indústria de bens de consumo além de torná de fazer com que o país é .. fizesse parte do sistema monetário internacional Esse segundo parágrafo diz que .. um componente .. um segundo componente quer dizer uma outra uma outra peça vital do do esquema desse Krotov desse economista né de Leningrado do

Instituto de Leningrado é .. dá mais poder dá mais controle dá um maior poder aos governos locais .. né poder esse sobre os recursos sobre a exploração sobre o aproveitamento desses recursos na região né .. Ele diz que isso daria né aos governos locais um maior interesse na .. em em se prevenir a poluição em se protegê o seu próprio ambiente local .. E'.. aqui ele diz né que as regras e os padrões ecológicos que .. só podem ser aplicados efetivamente quando as autoridades locais são o principal agente né monitorando quer dizer controlando esse esforço um esforço.

PRGF 8: Bom *the Russian republic contains most of the Soviet Union's natural natural resources — and has the greatest environmental problems But the government of the republic led by Boris Yeltsin .. has already begun to take steps to rem remedy .. its own problems It /hɛSɛNtɪl/ recently announced its intention to take control of all the natural resources within its /baundə-rɪs/ boundaries .. /baundə-rɪs/ boundaries .. fronteiras? O que é /baundə-rɪs/ boundaries? It recently announced its intention to take control of all natural resources within its /baundə-rɪs/ boundaries Deve ser fronteiras .. /baundə-rɪs/ boundaries .. Yeltsin has introduced a 400-day programme to revitalise the Russian economy with backing from President Gorbachov .. And from first January an innovative "polluter pays" policy comes into force.*

RESUMO: Aqui nesse parágrafo diz que a Rússia a república da Rússia contém a maior quantidade de recursos naturais e também tem os maiores problemas com relação ao ambiente Mais esse governo .. ou seja o governo dessa república desse .. da .. da Rússia já tomou alguns já vem tomando né alguns passos em direção ao .. à solução desses problemas né .. Esses passos .. podem ser os seguintes: recentemente o governo anunciou sua intenção de controlá né é .. de tê um controle de todos os recursos naturais dentro de suas fronteiras O governo já anunciou um programa de revitalização da economia russa com uma política é .. de .. de multas né aos às pessoas que poluem né quer dizer eles vão passar a aplicar multas às pessoas que provocam a poluição que poluem o ambiente Quer dizer em resumo isso deve ter ficado meio confuso A Rússia que é uma república da União Soviética contém né o .. a maior quantidade de recursos naturais da União Soviética e também contém os maiores problemas com relação ao meio ambiente Mais a Rússia já vem tentando .. é solucionar seus problemas através de várias medidas né descritas aí. Bom ..

PRGF. 9: Bom .. *The new law means that about 100 of the 500 major industrial concerns in the Leningrad region will have to shut down because they cannot meet the new standards set by the Russian authorities The new law means that about 100 .. of the 500 major industrial concerns .. will have to shut down .. concerns? .. Que que que é concerns? .. Tá estranha essa colocação a colocação dessa palavra aí*

.. Bom vamos lá .. *An optimistic Krotov believes that the extra money coming into the republic from the export of its resources both abroad and to other parts of the Soviet Union will help to reduce unemployment by providing funds to set up new cleaner industries ..Economic reform in the Russian republic will make it possible to harmonise .. /ɛntɹɛprɛNsɪrɪəl/ entrepreneurial .. Essa palavra eu não sei o que que é não /ɛntɹɛprɛNsɪrɪəl/ entrepreneurial .. Economic reform in the Russian republic will make it possible to harmonise /ɛntɹɛprɛNsɪrɪəl/ entrepreneurial activity and the rational use of resources with the protection of the environment*

RESUMO: Esse último parágrafo .. harmonizá (...) ah .. tá difícil essa palavra (...)

PE: Alguma idéia alguma hipótese?

LIA: Não .. a principio não .. A reforma econômica na Rússia tornará possível harmonizá .. /ɛntɹɛprɛNsɪrɪəl/ /entrepreneurial .. deixa eu voltá lá pra cima vamo vê se eu consigo .. vê .. esclarecê alguma coisa aqui 'Pera aí *The new law means that about 100 of the 500 major industrial concerns in the Leningrad region will have to shut down because they cannot meet the new standards set by the Russian authorities An optimistic Krotov believes that the extra money coming into the republic from the export of its resources both abroad and to other parts of the Soviet Union will help to reduce unemployment by providing funds to set up new cleaner industries Economic reform in the Russian republic will make it possible to harmonise entrepreneurial activity .. ahn ahn não sei não sei mesmo O parágrafo diz que essa nova lei .. que a Rússia que o governo russo implantou instituiu provavelmente fará com que uma grande quantidade de fábricas né sejam obrigadas a fechar por não conseguirem .. atingir os novos padrões estabelecidos pelas autoridades russas Mas ele acredita mas o Krotov não .. o economista é .. acha que .. acredita que o dinheiro extra que vai entrá na república né através da exportação dos recursos dessa república pro pro exterior e pra outras partes da União Soviética vai ajudá a reduzir o desemprego né é .. assim é .. através é .. da .. é.. de de através da .. criação de fundos que ajudem a estabelecê construí novas indústrias né novas indústrias que estejam de acordo com a .. a .. os padrões dessa nova lei instituída pelo pelas autoridades russas. Agora o que tá falando nesse último parágrafo que é um comentário dele eu num essa palavra tá é meio importante aí eu não sei ..*

PE: Isso que você acabou de falá é o último parágrafo ou é o penúltimo?

LIA: Penúltimo

PE: Agora essa palavra do último parágrafo 'tá interferindo com todo o parágrafo?

LIA: Não da última frase!

PE: Ah entendi

LIA: Da última frase

PE: Ah tá bom

LIA: Do penúltimo parágrafo da última frase

PE: Mas tudo bem você não tem a menor idéia do que seja você não quer fazer uma hipótese?

LIA: Não tenho a menor idéia não que eu não queira fazer

PE: Eu sei! Você tem uma idéia da classe gramatical dessa palavra? Pode ajudá?

LIA: É um adjetivo

PE: É. Pode deixá de lado?

LIA: Não não dá acho que não dá pra se deixá de lado. Tá tá tá .. dando uma característica dessa atividade. Essa atividade tá sendo comp .. tá sendo .. comparada aí com o uso racional dos recursos .. quer dizer ..

PE: Ah ah tudo bem

LIA: Bom último parágrafo ((LB)) [do parágrafo todo] *While these reforms will bring benefits to the Russian federation they will cause problems in other parts of the country .. and in the countries of Eastern Europe which until now have had almost unlimited access to raw materials for virtually nothing (...)* Uhm .. Enquanto essas reformas trarão benefícios .. à Rússia .. elas trarão elas causarão problemas em outras partes do país e em outros países da Europa do Leste que até agora .. tiveram quase tiveram um acesso quase ilimitado *to raw materials* .. *raw materials* seria matéria prima .. *for virtually nothing* por quase nada .. Bom (...) essas reformas elas tá trazendo benefício pra Rússia mas tá causando problemas em outras partes do país .. porque o estado compra quer dizer ele vai comprar da Rússia e .. a Rússia tá exportando a .. essa matéria prima esses recursos naturais tanto pros outros países como pras outras partes da União Soviética né. Quer dizer esse dinheiro que tá entrando na Rússia vai se benéfico para ela. Mais enquanto isso .. os outros países os outros estados e os outros países vão se prejudicá com isso .. Por que?

PE: Alguma coisa não tá fazendo sentido pra você?

LIA: É aí é que tá eu não entendi porque essas reformas .. trarão benefícios pra Rússia mas causarão problemas pras outras partes .. do país (...) Tá essas reformas algumas dessas reformas poderiam .. aumentá ainda mais o acesso a essas matérias primas .. às outras partes (...) porque .. bom .. seria talvez .. bom .. a Rússia exportando pra pras outras partes facilitaria então o acesso dessas partes às matérias primas Quer dizer é .. voltando aqui a um problema discutido logo no começo o problema na União Soviética é que .. como tudo é estatal tudo é monopolizado o acesso a essas matérias primas é muito fácil quer dizer então eles eles consomem muito eles desperdiçam matéria prima E isso .. é prejudicial pro ambiente quer dizer .. eles poderiam diminuir o consumo da da matéria prima sem afetar a produtividade Né agora então deve ser mais ou menos isso quer dizer essas reformas de uma .. de alguma forma facilitariam ainda mais esse .. o o acesso .. às matérias primas e fariam com que .. então .. é .. eles .. num .. meçam direitinho a quantidade certa pra pra consumo talvez consumam em em .. excesso continuem consumindo em excesso ..

TAREFA II: GRIFAR PALAVRAS

Aqui .. até agora .. /PuLP/ *pulp*? *pulp* aqui né *which is polluted by the effluent from nine /PuLP/ pulp and paper plants* indústria fábrica *paper plant* tá fábrica de papel /PuLP/ *Pulp* .. algum tipo de indústria mas que tipo não sei Não deu pra inferi também aí acho que aí né foi uma palavra que eu não soube dizê qual o significado e não sei até agora *Narrow-minded people blame (...) lack is the lack* Ai eu pensei talvez como falha deficiência né mais ou menos assim acho que deu algum sentido pelo contexto Eu já tinha tido contato com essa palavra antes acho que em algum texto que eu já tinha lido nas aulas de .. instrumental Mas não lembro Daí eu acho que é isso [continua ((LB))] *Timber* .. *timber* é algum tipo de recurso natural eu não sei bem qual é também Ai ele só tá citando então acho que não tem tanto .. tanto problema em não saber Não sei ainda o que que é .. *profligate* .. exagerado demasiado excessivo alguma coisa nesse sentido Não sabia Quer dizer ainda não sei né pode ser até que não seja isso aí né ..[continua lendo em voz baixa] .. *wastefulness* .. analisando os sufixos aí *ness* é um sufixo formador de substantivo isso aí é substantivo Então *waste to waste* é gastar desperdiçar *Wasteful* é um adjetivo relacionado né a isso *wastefulness* .. desperdício [continua ((LB))] *the surcharge surcharge* Bom eu vou sublinhar aqui mas .. Bom aqui *he also proposes the addition of an ecological supplement to the price the surcharge* quer dizer ele tá voltando aí um pouco a .. *surcharge* seria isso esse *ecological supplement to the price* né quer dizer então uma taxa adicional aí alguma coisa assim Quando eu tava lendo nem cheguei a pará na palavra foi automático foi uma coisa automática não sei se certa mas foi [continua ((LB))] Eu não sei o que quer dizer realmente assim tipo *enterprise joint ventures* mas são termos de economia mesmo né .. *joint ventures*? *cartel trust* deve ser coisa assim tipo parecida né acordos tá falando que com firmas estrangeiras algum tipo de acordo união aliança

.. [continua ((LB))] Bom aqui eu não sabia o que que é mas deve ser moeda acho que não precisa nem .. /Robol/ *ruble Stake* .. *people with the greatest stake in the environment are the local people* .. *stake*? Eu não sei se a pronúncia é essa .. interesse? talvez As pessoas com um maior .. interesse .. no ambiente são as pessoas da região Não sei se interesse caberia muito bem aí não mas .. inclusive eu falei na hora em que eu tava aqui mas eu fiquei em dúvida Não sei se é isso mesmo [continua ((LB))] /baundə'ris/ *boundaries* .. creio que isso aqui seja fronteira limite Também não sei o que que é isso aqui Não sabia né? Não sei o que quer dizer *entrepreneurial Raw materials* aqui matérias primas né Inferi também *Raw* .. eu já ouvi essa palavra antes mas não guardei qual o significado dela.

G.2 – Protocolo do leitor intermediário (LIO)

TAREFA I: LEITURA

TÍTULO: Eu já deparei com um problema aqui Eu não sei o que é /KəR/ *cure* e .. essa palavra aqui eu não sei o que é (...) [referindo-se a *ills*] Bem acho que vai falar sobre alguma coisa em relação à .. União Soviética mas eu não sei bem o que que é porque acho que a palavra .. /KəR/ *cure* ou *ills* .. tá? tem que saber o que que é Mas deve falar alguma coisa sobre .. é .. ecologia na União Soviética Certo ou alguma coisa nesse sentido Tem a ver com o verde .. Tá vou ler o texto agora (... ..)

PRGF. 1: Bom já deparei com um problema aqui que eu não sei o que é /Hiʃʌn/ *reaching* certo? Mas .. vou continuar lendo pra ver se eu descubro (...)

RESUMO: Bom no primeiro parágrafo ele fala que a crise ecológica tá? na União Soviética tá? *is reaching* que eu não sei o que que é tá? é .. um ponto é .. um ponto onde as muitas .. a vida das pessoas dos soviéticos tá? são colocadas em perigo alguma coisa nesse sentido certo? E somente tem um um caminho tá? .. para resolver essa crise .. certo? no caso relacionada também com a crise econômica Então de acordo com .. com o economista de Leningrado certo? chamado Mikail Kro Krotov certo? do Instituto de .. Mecânica de Leningrado tá? a crise ecológica é o resultado direto da .. da .. crise econômica na União Soviética (... ..)

PRGF. 2: Deparei com a mesma .. no segundo parágrafo deparei com o mesmo problema .. certo? a palavra /Hiʃʌ/ *reach* eu não sei que que é (...) Bom já deparei também com outras palavras /ənʃə'dledʊ/ *unscheduled* e /Sənsiə/ *Swansea* .. certo? São duas palavras também que eu não sei (...) e /Pʌlp/ *pulp* também é .. outra palavra que eu não sei .. e /ʃoʊz/ *shores* também não sei o que que é certo?

RESUMO: Bom o segundo parágrafo .. fala sobre os problemas do meio ambiente no país certo? .. E' .. parece-me aqui que ele fala que os problemas econômicos no país .. chegam num estágio onde a população é .. certo? tá .. alguma coisa os problemas os problemas os problemas .. sobre o meio ambiente estão aumentando e .. alguma coisa nesse sentido né? .. Certo? que os problemas do meio ambiente estão aumentando e alguma coisa relacionada as pessoas vem declinando O que eu não sei o que que é devido a palavra /Hiʃedʊ/ *reached* e /HeLtu/ *health* tá? .. Bom segundo (...) Depois ele diz aqui *Krotov said in an unscheduled talk in Swansea* não entendi o que que está escrito aqui devido a palavra *unscheduled* e *Swansea* certo? Depois ele relata que as pessoas de Leningrado por exemplo é .. bebem água de um do lago Ladoga tá? é .. local é poluído por um por .. o local é poluído .. por .. alguma coisa .. tá? Deve ser deve ser fábrica de papel. É isso mesmo tá? Então ele .. é o seguinte. No lago Ladoga são jogados puluentes certo? vindos tá de nove .. /Pʌlp/ *pulp* eu não sei o que é de nove fábricas de papel que .. no caso essas são as fábricas de papel e as fábricas de algum .. alguma outra coisa que é o /Pʌlp/ *pulp* no caso certo? Essas esses dois complexos tá? de fábricas jogam diariamente uma série de .. fatores químicos certo? e o finalzinho do parágrafo que eu não sei o que que é [referindo-se a *shores*] (...)

PRGF. 3: Bom já notei .. nesse parágrafo já já empaquei na palavra *blame* tá? e logo seguinte tem a /nə'roumɪndəd/ *narrow minded* que eu não sei o que que é (...) Deparei tá? com outra palavra aqui também chamada /Bə'nɛd/ *banned* (...) Tem outra também aqui /Lək/ *lack* que eu não sei o que que é (...) Bom no terceiro parágrafo ele diz o seguinte Eu não entendi muito bem vou tentar aqui explicar alguma coisa certo?

RESUMO: Bom parece que o o .. Krotov que é o economista lá do Instituto de Leningrado ele insistiu que a indústria não é alguma coisa acho que é .. *blame* deve ser problema né? culpado alguma coisa assim certo? (...) Bom pelo que eu entendi aqui do que tá entre aspas no terceiro parágrafo certo? é .. as pessoas .. certo? .. não o seguinte Acho que o progresso tecnológico tá? é .. são produzidos pelas próprias pessoas tá? e isto é .. tem feito a atividade industrial lá (...) deve ser abaixar ou diminuir por razões ecológicas tá? Mas ele é .. o Krotov explica que o problema real certo? é que o /Lək/ *lack* eu não sei o que é /Lək/ *lack* certo? o /Lək/ *lack* não sei os mecanismos de mercado fariam um uso racional de .. é .. fontes naturais certo? Deve ser alguma coisa relacionado tá? que ou seja a indústria ou a atividade industrial aumente certo? mas que ela

use o as fontes naturais racionalmente não poluindo alguma coisa acho que .. o que o terceiro parágrafo diz é isso .. certo? (... ..)

PRGF. 4: Bom deparei com outra palavra aqui é *timber* (...) Tem então *profligate* outra palavra que eu não sei (... ..) Eu não sei o que quer dizer *bring* também esqueci Eu conheço essa palavra mas eu esqueci o significado (... ..)

RESUMO: Bom pelo texto aqui eu entendi que *timber* deve ser algum .. algum combustível fóssil certo? .. Bom o quarto parágrafo diz o seguinte: bom dentro do sistema econômico corrente onde o estado tem o monopólio de todas as reservas naturais *timber* óleo gás então são .. então o sistema econômico é virtualmente livre certo? Isto foi encorajado tem encorajado tá .. a .. concepção *profligate* que eu não sei o que que é pela indústria certo? Então Krotov ele estima que é .. para cada unidade .. nacional certo? /inKōM/ *Income* .. outra palavra que eu esqueci de relatar que é *income* que eu não sei o que que é. *In* é dentro *come* é vir certo? *Income* vir para dentro certo? Deve ser indústria dentro do país alguma coisa assim certo? Então a União Soviética usa de duas a três vezes as .. as .. muitas fontes tá .. e .. as mesmas .. as .. as muitas fontes tá? em outros países também certo? Bom ele diz que isso pode ser possível tá? *to bring* eu não sei o que que é .. a concepção dentro de uma linha com outros países é .. na qual podem vai reduzindo a produtividade alguma coisa assim (... ..)

PRGF. 5: Bom no quinto parágrafo eu já deparei com outra palavra que eu não sei é (...) /wæstɪfʊlnəs/ *wastefulness* certo? Eu não sei o que que é /ʃɑRg/ *Charge* também não /Hætt/ *Rate* também não /mæɪns/ *Maintains* também não sei o que que é /sɜːsɪtɪ/ *scarcity* também não sei o que que é (... ..) Bom eu sei o que é .. *clean* *Clean* se não me engano é apagar Agora *clean up* *up* é pra cima apagar pra cima eu não entendi (...) Aparece aqui a .. o .. novamente a palavra .. não não é novamente não é diferente da anterior lá no .. parágrafo anterior que é /prɒfɪtəbəl/ *profitable* Não sei o que que é *profit* *able* é capaz mas *profit* eu não sei o que que é (... ..)

RESUMO: Bom lendo aqui o quinto parágrafo certo? diz o seguinte: que o melhor caminho tá? pra (...) /KəR/ *cure* deve ser cura não sei tô chutando certo? que o melhor caminho pra curar a indústria certo? .. deve ser alguma alguma .. coisa /Hætt/ *rate* deve ser ratear não não é ratear acho que usar de maneira mais racional os fontes naturais certo? Ele também propõe tá? a adição de um suplemento ecológico no preço certo? (...) A sentença onde ele fala *the surcharge reflecting the scarcity of the resource or the damage done to the environment* eu só entendi a última parte que é a fonte ou o gerenciamento feito no meio ambiente certo? E a primeira parte eu não entendi o que é certo? Bom dentro do velho sistema é.. pesquisas dentro da ciência e tecnologia precisou tá? é .. sei lá deve ser apagar a indústria também apagar uma indústria também estagnada alguma coisa assim certo? Isso foi justamente não sei lá aproveitável para .. o gerenciamento de fábricas .. para introduzir o *clean* que eu não sei o que que é O *clean* é apagar mas aqui não tem esse sentido de apagar certo? Não deve ser acabar né? Não acabar também não encaixa (...) Bom enfim eu não consegui inferir a palavra .. quer dizer um sentido ..pro português pro sentido da frase aqui pra *clean* certo? (...) Bom vou ler o sexto parágrafo (... ..)

PRGF. 6: Bom no sexto em .. eu deparei com a palavra /ʒɔɪnt/ *joint* /ʒɔɪntvɛntʃər/ *joint venture* que eu não entendi muito bem o que significa *joint venture* tá? e /fɔRɪn/ *foreign* que eu também não sei o que que é (... ..) Eu não entendi também o que é /Huubl/ *rouble* num .. num sei o significado dessa palavra.

RESUMO: Bom nessa .. nesse sexto parágrafo a cura que o Krotov tá? propôs para a União Soviética sobre o meio ambiente tá? acho que fará a privatização de pelo menos metade tá? provocará a privatização de pelo menos metade das fábricas do estado pra introduzir uma competição entre os produtores certo? Isso isso envolverá tá? companhias existentes é .. isso envolverá novas companhias existente certo? e também envolveu /ʒɔɪntvɛntʃər/ *joint venture* em .. deve ser com outras firmas certo? ou seja vai entrar novas firmas no mercado certo? juntamente fazendo uma ligação com outras firmas certo? Mais ou menos nesse sentido certo? Mas .. a União Soviética ela precisa ir em frente certo? Ir em frente tá? pra ter uma boa um bom parque industrial alguma coisa assim certo? e ter um consumo de produção bom diz o Krotov o Krotov né certo? (...) Bom também o mais importante é que o país ele precisa de um sistema monetário internacional que .. sustente acho que isso essa estrutura de mercado certo? É .. praticamente isso que eu entendi do texto Do texto não do parágrafo seis (... ..)

PRGF. 7: Não sei o que é *scheme* [soletra em português s-c-h-e-m-e] Não sei o que que é (... ..) Eu não entendi a a .. expressão *handing over* certo? *Over* é sobre *handing* *hand* é mão botando a mão sobre? .. não sei (... ..) Bom eu encontrei outro problema aqui .. que é .. quando ele fala assim .. *people with the greatest stake* (...) não sei o que é *stake* certo? Aliás eu não sei o que é a expressão *greatest stake* certo? Não sei mesmo mas vou ver se eu descubro (... ..)

RESUMO: Bom o que ele fala aqui no .. no .. sexto parágrafo sext .. não primeiro segundo terceiro quarto quinto sexto sétimo no oitavo certo? no oitavo parágrafo é o seguinte: que o segundo componente vital na proposta dever ser proposta de Krotov certo? E um controle gover governamental local tá? sobre as fontes e .. naquela região certo? .. É agora eu não entendi como fica *handing over* ne (...) tá mas deve ser algum sobre colocando a mão colocando esse controle governamental sobre propriedades as (...) as agora as propriedades do estado certo? Bom isto dará ao governo local o direito .. um direito .. no .. um interesse no controle tá? da poluição e proteção dos seus locais de meio ambiente certo? Então depois ele fala aqui que .. pessoas com *greatest stake* que eu não sei o que é do meio ambiente são as pessoas locais ele diz né Bom os padrões ecológicos e as regras podem ser bem aplicadas efetivamente quando as autoridades locais são o principal agente de monitoramento e .. é.. de .. ne (...) Bom (...) ele fala assim *monitoring and enforcement* certo? é a força e monitoramento naquela região certo? na região local Ta bem vou passar para o proximo parágrafo o nono (... ..)

PRGF. 8: Bom acho que agora eu descobri *greatest* Deve ser alguma coisa muito (...) o maior problema (...) bom eu deparei com uma palavra que eu não conheço é /aNaunSɛdɪ/ *announced* (...) não sei o que que é (...) Bom tem uma palavra aqui chamada assim /baunDə-rɪs/ *boundaries* certo? Bom eu sei que *boundary* deve ser contorno certo? *boundaries* eu não sei o que que é Pode ser um falso cognato certo? (...)

RESUMO: Bom o que ele diz aqui no nono parágrafo [referindo-se ao oitavo] é o seguinte: que a república russa é .. contém a maioria dos fontes naturais da União Soviética e devido a isso é .. tem os maiores problemas do meio ambiente ambientais ne? Mas o governo da república certo? Não sei o que é que é /Lə-du/ *led* Sei que *led* é a luzinha que acende no *drive* so isso Ta tem é .. /alRə-du/ *already* Sei que /Rə-du/ *ready* é pronto tá? pronto certo? *already* não sei o que que é Começou com pequenos com passos pra remediar esse problema certo? Quer dizer são os problemas ambientais tá? Ele recentemente é .. anunciou sua intenção de fazer um controle de todas as fontes naturais certo? com *boundaries* que eu não sei o que que é certo? Esse Yeltsin também introduziu é.. um programa diário para revitalizar a economia russa con .. conjuntamente com o presidente Gorbachev e .. em primeiro de janeiro um .. pagamento ..dever ser um imposto novo né certo? E ele fala *innovative "polluter pays"* .. eu não sei o que é "*polluter pays*" /poLaɪs/ *policy comes into force* .. certo? Eu não entendi bem o final disso aí o final do parágrafo (... ..)

PRGF. 9: Bom o próximo parágrafo que eu 'tou lendo aqui eu não entendi aqui o que é a palavra /ʃut/ *shut* é .. /ʃutDauN/ *shut down* certo? (... ..) Uma palavra aqui também eu bom .. também eu (...) Bom eu já vi essa palavra mas eu não lembro o significado dela que é /Bout/ *both* certo? E tem a palavra seguinte que é também /aBRouDu/ *abroad* que eu também não sei o que que é eu nunca vi na vida (... ..) Tem outra palavra aqui também que eu não não entendi nunca vi na vida: *entrepreneurial* certo? Não sei o que que é (...)

RESUMO: Bom esse décimo [referindo-se ao nono] diz o seguinte: que as novas as novas leis tá? que .. foram seriam impostas de 100 das 500 maiores indústrias de na região de Leningrado terão é .. devido a isso certo? eles não podem colocar colocar novos padrões .. pelas autoridades russas Não ficou muito legal certo? (...) Acho que é o seguinte: Acho que de 100 dessas indústrias maiores indústrias da região de Leningrado eles não conseguiram eles não obedeceram as autoridades russas colocando esses novos padrões de .. de .. de .. lei contra o proteger o meio ambiente certo? Bom o otimista Krotov acredita que o dinheiro o dinheiro extra certo? vindo tá pra dentro da república devido a exportação de seus .. de suas fontes certo? precisa ser acho que aplicado em outras partes da União Soviética e .. e que ajudarão a reduzir *unemployment* certo? a reduzir deve ser a diminuição de .. fundos a diminuição de fundos pelo .. não não é reduzindo .. reduzindo alguma coisa certo? Depois ele fala pelo fornecimento de fundos pra .. *set up new cleaner industries* as ..cu não sei o que é *cleaner* certo? Depois ele diz aqui que uma reforma econômica da república russa tá? fará .. é .. será possível harmonizar o *entrepreneurial* que eu não sei o que que é atividade atividade e o uso racional dos das fontes com proteção do meio ambiente diz Krotov certo?

PRGF. 10: Tou lendo agora o último parágrafo (... ..) Eu não sei o que é a palavra /Hau/ *raw* /tuHau/ *to raw*

RESUMO: Bom o último parágrafo diz o seguinte: enquanto essas reformas tá? são benéficas para a federação russa elas causam elas causarão problemas em outras partes do país do país certo? E nos países da Europa externa deve ser no oeste ne certo? Europa Bom países da Europa aí grudados ali na União Soviética .. na qual ate agora eles eles não tem .. tido um acesso limitado não limitado ne ilimitado de materiais para (...) /tuHau/ *to raw* /tuHau/ *to raw* eu não sei o que que é Depois ele fala materiais para virtualmente nada certo? É praticamente .. eu não entendi o finalzinho .. devido o que eu não sei a palavra o que significa a palavra *to raw*.

TAREFA II: GRIFAR PALAVRAS

Bom /Hɪʃ/ *reach* é uma palavra que eu não sei o que que é .. Tem uma derivada de *reach* aqui também *reached* conseqüentemente também eu não sei o que que é (...) /Hɛlθ/ *health* também não sei (...) Ah sim essa eu nunca vi na vida /ənʃə-dlɛd/ *unscheduled* eu sei o que é um /ʃə-dlɛ/ computacionalmente mas eu não sei no texto aqui como é que fica E a outra é Swansea Eu não lembro Eu devo ter visto essa palavra mas eu não lembro não lembro o significado dela (...) Ah sim outra palavra que eu não sei é /Pʊlp/ *pulp* .. certo mas eu consegui inferi que *pulp* deveria ser um tipo de fábrica fábrica de alguma .. fábrica de *pulp* não sei o que é *pulp* .. é fábrica de alguma coisa mas .. no texto aqui eu não consegui saber o que era *pulp* .. Ah aqui tem outra palavra aqui é *shores* Eu também não sei o que que é .. Tem também *narrow-minded* /MʌNdɛd/ *Blame* não sei o que que é também .. *led* .. *banned*? .. *lack* .. Também não sei o que que é .. *timber* .. *timber* eu não sei o que que é mas eu consegui .. eu acho que deve ser algum derivado das fontes naturais que o texto diz .. depois ele diz óleo e gás certo? *Timber* acho que deve ser algum derivado disso aí .. *income* também eu já vi essa palavra mas eu não lembro o que que é .. *bring* também não sei o que é .. /wɛɪstɪfʊlɪnɪs/ *wastefulness* nunca vi essa palavra na minha vida .. /Hɛt/ *hate* .. *maintains* não sei se é assim que pronuncia .. *surcharge* .. *scarcity* não entendi o sentido de *clean up* aqui .. sei que *clean* é apagar mas *clean up* eu não sei o que que é .. /prɒfɪtəbəl/ *profitable* também .. eu sei o que é *able* mas não sei o que é *profit* Então não adianta nada .. o *clean* de novo aqui e o *safe* .. *ills* *ills* também não sei o que que é .. /Kɜr/ *cure* que eu não sabia no .. inicialmente agora eu descobri o que que é E' cura mesmo certo? Descobri lendo o texto marco ou não marco? [pergunta para a pesquisadora, que responde positivamente] *Among* também não sei mas no texto aqui acho que deu pra inferi o que que era certo? Deve ser alguma coisa entre produtores alguma coisa assim .. /ɔɪntvɛntʃɪz/ *joint-ventures* .. eu não sei o que que é e .. /fɔrɪn/ *foreign* também não sei o que é certo? Mas *joint ventures* deve ser .. eu sei que isso aí é um termo de mercado certo? que deve ser você juntar duas ou mais empresas pra formar uma firma só alguma coisa nesse sentido .. Então eu imaginei que era isso .. /Hɛvɪ/ *heavy* .. outra palavra que eu não lembro o que que é .. *join* .. *rouble* .. não sei o significado também .. a palavra *convertible* que sei lá deve ser alguma coisa em relação à conversão .. ser capaz de converter alguma coisa assim .. Uma palavra que eu consegui inferi mas não sei não ficou muito legal no texto .. que eu não sei o que que é .. não entendi o significado de *handing over* certo? Não sei o que que é. Deve ser colocar a mão sobre .. alguma coisa nesse sentido .. sobrepor .. não sei não se encaixa muito legal no texto .. *the greatest* é o maior. Eu não sabia o que que era depois mais na frente eu descobri consegui inferi certo? Eu não sei o que é *stake* .. Eu não sei o que que é *enforcement* mas deve ser .. alguma coisa relacionada no texto aqui com as autoridades locais certo? .. Não consegui inferi o que que é .. o /lɛd/ *led* aparece novamente e eu não sei o que que é .. *announced* também não sei o que que é .. *bon* .. *boundaries* *boundaries* .. eu sei que *boundaries* é contorno né .. *boundaries boundaries* eu sei que *boundaries* é contorno né .. *boundaries* eu não sei o que que é .. *polluter* .. essa palavra aqui eu não consegui inferi .. pode ser imposto no caso aqui no texto sei lá Mas não tem esse sentido porque está entre aspas não deve ter esse sentido .. *concerns* .. aqui .. *shut down* também eu não consegui uma tradução boa pra coloca no texto .. /əbrʊd/ *abroad* .. *unemployment* .. também não sei o que que é .. *cleaner industries* também não sei .. o significado de *cleaner* .. *entrepreneur* .. sei não sei se é assim que se fala .. *entrepre* .. também não sei o que que é .. *to raw* .. eu não entendi muito bem o final dess .. o final do texto aqui ..

G.3 – Protocolo do leitor iniciante (LIE)

TAREFA I: LEITURA

TÍTULO: Vou vê se entendo esse .. o título primeiro .. vê o assunto .. se eu sei alguma coisa do assunto (...) *capital* .. *capital* /Kʊr/ *cure* para a União Soviética .. então *capital* para a União Soviética *ecological* *ills* .. *ills* .. *ills* .. não sei o que é nem /Kʊr/ *cure* .. também não sei .. *Soviet Union's capital* então *capital* a mais né para a União soviética dinheiro a mais pra ajudá a ecologia na União Soviética .. Esse subtítulo aqui .. *Stephanie Pains* ah! é a mulher que escreveu Então o que eu sei do título .. a ecologia (...) tá muito (...) a ecologia .. é muito .. dia a dia .. tal tem gente falando de ecologia .. então acho que deve ser alguma coisa que 'tão fazendo lá que 'tá .. acabando com alguma mata alguma coisa é .. uma poluição .. poluição fábrica .. deve sê .. alguma coisa .. poluição nas cidades nas floresta vamo vê o que que é então

PRGF. 1: *The* /eKɒlɒgɪkəl/ *ecological* /KRɪzɪs/ *crisis in the Soviet Union* é .. ecologia .. ecologia .. /KRɪzɪs/ *crisis* .. não sei não sei o que que é /KRɪzɪs/ *crisis* não sei .. na União Soviética .. *is* .. /sɒm/ *some* .. um ponto .. /ʃeɪŋdʒ/ *change* não sei /KRɪzɪs/ *crisis* também não na União Soviética .. vou voltá né esse texto 'tá difícil *The* /eKɒlɒgɪkəl/ *ecological* .. a ecologia na União Soviética .. a um ponto .. como que .. foi muito .. a vida

do povo na União Soviética .. a um ponto .. vão falar de ecologia .. a União Soviética é um ponto que tem muito a ver com a vida .. do povo soviético certo? *In danger danger* .. deve sê o perigo deve sê o perigo o perigo o perigo é .. está em perigo O povo está em perigo o povo soviético está em perigo Então um ponto .. é muito .. o soviético povo está em perigo .. o povo soviético está em perigo e .. a troca não .. homem .. próprio não sei .. uma troca não .. 'pera aí economia correndo? *run run chuva chuva can* pode /SOLv/ *solve* *solve* não sei .. é isso solver solver a /KRÍZÍs/ *crisis* .. *crisis* é chave .. você tem que descobrir o que é *crisis* .. *crisis* é palavra chave .. aqui volta lá *ecological crisis* .. *crisis ah crisis* .. *crisis* de crise *ahn crise* .. crise política crise ecológica crise econômica Ah agora sim *ecological crisis* .. ecologia a crise na ecologia na União Soviética Ah sei descobri o que é crise agora .. a um ponto .. onde .. ela .. eles .. muito .. o povo .. perigo Então o povo soviético em perigo .. e cada troca de seus .. economia chuva .. pode resolver a crise Então econômica o modo econômico pode diminuir a crise né pode resolver a crise .. Ah eu estou aqui crise .. de acordo com Mikhail Krotov um economista de Leningrado Mechanical Institute Instituto Mecânico a ecologia .. a crise ecológica é .. um dire .. é um resultado direto da econ .. da crise econômica na União Soviética Então a crise na ecologia é um resultado direto da crise econômica da União Soviética certo?

RESUMO: Então o primeiro parágrafo diz o seguinte: primeiro parágrafo entendi que .. a ecologia a crise na União Soviética está em crise ecológica certo? E tem um ponto que onde que o povo corre perigo por causa dessa crise e .. em troca .. e também essa crise é acompanhada de que? Pela crise econômica De acordo com o car .. aqui o Mikail Krotov economista de Leningrado Instituto Mechanical a crise ecológica .. né .. é um resultado direto da .. crise econômica da União Soviética .. certo? Desse parágrafo eu não entendi /Héʃɪŋ/ *reaching* .. e .. algumas umas aí que eu tentei botá .. sabê o que que é ..

PRGF. 2: /envɪrəməntəl/ *Environmental problems in the* /KauNtr/ *country have* .. então meio ambiente .. os problemas ambientais meio ambiente no .. país .. tem .. *country* é país oh .. beleza hein! Lembrei bem! Tem /Hɪʃɛd/ *reached reached* ah acho que é reagido .. reagido tem reagido estágio onde pessoas /Helt/ *health* quentes? .. estão declinando Pessoas crentes quentes *health* quentes será? Não sei acho que é quentes Pessoas quentes não tem sentido .. pessoas *health* quentes Será pessoas quentes poderosas então né .. meio ambiente problema ambiental no país .. tem .. foi .. reagido .. um estágio onde pessoas quentes .. Krotov disse *say* é disse na .. em uma a /ʌnʃɛdlɛd/ *unscheduled* /TaLk/ *talk in* /sanStə/ *Swansea*. Ah de novo *Swansea* então ele disse em uma não sei o que que é *unscheduled* .. não sei o que que é .. *talk* .. falou em *Swansea* .. povo de Leningrado por exemplo bebem água /DRauN/ *drawn* (...) *drawn* (...) não sei deixa eu ver .. problemas no meio ambiente no país .. os problemas no meio ambiente no país tem reagido /Hɪʃɛd/ *reached* /Hɪʃɛd/ *reached* tem reagido um estágio onde pessoas quentes estão declinando .. Krotov disse em uma *unscheduled talk in Swansea* disse falou in *Swansea talk* disse em uma reunião sei lá *unscheduled* deve ser .. uma convenção em *Swansea* ah então Krotov disse que (...) tem pessoas que estão estão declinando as pessoas aí estão declinando não sei o que é *help* O povo de Leningrado por exemplo bebe água .. /DRɪN/ *drawn* para Lake Ladoga .. bebe água .. /DRɛuN/ /DRɪN/ *drawn* dirigida né que vem acho que vem do Lago .. Ladoga Ladoga deve ser o nome do lago 'tá maiúscula com sua poluição ah .. tá! .. poluição do rio aqui .. então não vai ter água Com sua poluição .. pelo afluente .. pelo seu afluente .. né .. pelo afluente para nove .. para nove /PuLp/ *pulp* .. *pulp pulp* não sei que que é .. *nine and paper plants* .. e papel .. plantas que faz papel acho que é nove .. para nove *pulp* and plantas /Pəpər/ *paper* e dois /Daɪ/ *dairy complexes and* /Sɛrɪs/ *series of* /ʃɛmɪkəl/ *chemical* /FəKtɔrɪs/ *factories on its* /ʃəRS/ *shores Dairy dairy* não sei o que é então dois complexos diários ah não não é *dairy* complexos é .. florestais e uma série de químicos ah .. pode ser alguma coisa poluente .. *dairy* complexos poluentes uma série de *chemical* fatores .. no setor nos /ʃəRS/ *shores* lá no lugar *shores* .. *shore* .. *shore* eu não sei *shore* não sei *shores* eu não sei Meio ambiente problemas ambientais então voltando aqui problemas ambientais no país tem .. tido uma reação uma reação .. estado .. *state* estado onde pessoas estão declinando .. o povo de Leningrado por exemplo bebe água .. poluídos.

RESUMO: Então o segundo parágrafo é assim: Krotov .. o problema ambiente na União soviética para Krotov aquele cara o economista Krotov diz: o problema ambiental no país tem o que? tem reag reagido um *stage* onde pessoas estão declinando estão acho que caindo né morrendo sei lá Então ele disse isso numa .. convenção em falou em *Swansea* Então ele dá um exemplo Por exemplo o povo de Leningrado bebe água /DRɛuN/ *drawn* rio dirigida de um rio que vem dum rio não de uma lagoa aliás de uma lagoa Ladoga e como a água é poluída por causa que caiu complexos .. com a poluição da água o seu afluente por seu afluente de nove *pulp* e .. plantas .. duas .. cada planta assim nove plantas que tão movendo por dia não sei .. ele polui a lagoa por causa dos dois complexos e uma série complexos /Daɪrɪ/ complexos *dairy* e uma série de químicos fatores químicos no lugar lá certo? Então ele falou o que? .. o problema no país 'tá matando .. tá acabando com muitas pessoas e deu um exemplo de Leningrado um exemplo sobre a água da lagoa e a lagoa poluiu né foi poluída pelo afluente pelo afluente né seu efluente afluente é..

PRGF. 3: Esse parágrafo é o terceiro Krotov insistiu que a indústria não /Blɛm/ *blame* não polui não /Blɛm/

blame .. /BLɛM/ *blame* .. não sei Krotov insistiu que a indústria não *bl* .. *bl* .. não sei o que que é .. Krotov insistiu que industriais indústrias os industriais não .. não sei esse *bl* .. tem que saber .. Krotov insistiu que indústrias .. foi .. não foi .. /nərouMNdɛd/ *narrow-minded* povo .. /BLɛM/ *blame* crises então que os industriais .. não nem ligam acho que .. para isso .. para *narrow-minded* não sei o que que é .. povo ah! deve ser o povo né .. /BLɛM/ .. a crise a tecnologia no .. *itself* suas .. têm /Lə-d/ *lead* para muitas atividades industriais .. /BɛNG/ *being* /Bə-nɛd/ *banned* não sei reações ecológicas .. /BɛNG/ *being* /BɛNG/ .. *being* /Bə-nɛd/ *banned* não sei /BɛNG/ *being ecological reactions* reações ecológicas Então Krotov insistiu que a indústria não polue Não .. não despolue acho que não faz nada é o povo .. a crise .. a crise na tecnologia no progresso tecnológico também e suas .. tem .. com muitas atividades industriais As indústrias né tem processos industriais fazem .. causam.. reações ecológicas O problema real ele /ɛsplənɛd/ *explained* /ɛsplənɛd/ *explained* ele explanou ele explanou né é .. o lago de .. é o grande .. /Lɛk/ *lack* não é lago não /Lɛk/ de mercado .. mecanismo *mech* é o grande mercado mecanismo para a reação para usar o recurso natural .. Então o real problema ele explanou explanou .. ele quem o Krotov lá Krotov explanou .. é o .. é seu .. o .. por causa .. é que .. tão poluindo .. os os mercados os mecanismos de mercados para .. para a racionalização o uso dos recursos naturais .. e .. então ele falou que o problema é que .. tem que começar racionar .. os recursos naturais certo então começar racionar os recursos naturais.

RESUMO: Então no terceiro parágrafo Krotov insiste que as indústrias não .. *to* /BLɛM/ *blame* não sei .. *narrow-minded* também não .. povo né /BLɛM/ *blame* .. povo de *narrow-minded* .. a *crisis* .. fala que a crise ecológica tá na tecnologia industrial no progresso tecnológico porque .. é .. muitas atividades industriais provocam as reações né .. as reações ecológicas muitas atividades industriais provocam as reações .. ecológicas certo e o problema real falou Krotov ele explanou né falou explanou escreveu é .. foi .. está no .. está no .. na racionalização da água dos recursos naturais na racionalização dos recursos naturais certo? Então no terceiro parágrafo então insistiu a indústria não .. não *bl*.. então e o povo diz que a crise a crise fala de crise e .. por causa do grande avanço né tecnológico né tecnologia então esse progresso afeta as muitas indústrias em ação essas indústrias em atividades né formam as reações ecológicas E o problema real está no reacionamento da água dos recursos naturais .. Certo agora quarto quarto parágrafo.

PRGF. 4: *Under the* /Kʊrɛnt/ *current* /ɛkɒnɒmɪk/ *economic* /sɪstɪm/ *system* o corrente o o sistema economico corrente onde o estado tem o monopólio nos recursos naturais .. onde o sistema economico .. *under* não sei deve ser onde onde não onde é .. onde não sei o que que é também *Under the current economic system under* .. o corrente o corrente .. o sistema economico lá vigente né Rússia o sistema economico deles lá onde o estado é tem o monopólio nos recursos naturais .. *timber* também não não é nos recursos naturais .. *timber timber* apareceu noutro texto e eu não sei que que é isso *timber* .. e.. óleo *oil* e gás e óleo e *timber* deve ser tiner alguma coisa química tiner gás e também .. uhm .. são uhm .. virtualmente livres Ah! .. então .. o sistema corrente economico lá onde .. o estado tem o monopólio dos recursos naturais óleo *timber* óleo e gás *timber* óleo e gás também .. são virtualmente livres Se .. seus .. estes têm encorajado .. *timber oil and gas* .. *and so* .. também são virtualmente livres Então .. o estado tem o monopólio né .. dos recursos naturais Ah! e *timber* óleo e gás são virtualmente livres E porque isso aí vem de recurso natural né Se .. estes têm encorajados *profligate* proliferação /kɒnsʊnpɪʃən/ *consumption* para a indústria .. /kɒnsʊnpɪʃən/ *consumption* .. estes.. fatores eles têm encorajados proliferação da consumição para a indústria .. não sei o que é *profligate* nem /kɒnsʊnpɪʃən/ *consumption* mas dá pra ver aqui Suas ..tem encorajado a proliferação *profligate consumption* concepção concepção já lembrei Tá então estes têm encorajados .. a proliferação aumentado a proliferação da consumição pelas indústrias uhm! então as indústrias têm uma certa concepção .. certo Então esses fatores aí .. tem certa concepção do que acontece Krotov estimou que para .. por muitos muitos união internacional /ɪnkoʊm/ *income* união internacional entre /ɪnkoʊm/ *income* não sei *income* a União Soviética usa dois para três tempos muitos .. /rɛzɔrsɛs/ *resources* (...) a União Soviética usa dois pra *three* não é 3 de 3 não deve ser 2 para 3 tempos .. muitos recursos *resources resources* não .. né e não tem nada a ver *resource* ah! *resources* então acho que é recurso sim ... recursos naturais recursos .. *was* cada um *countries* países cada país *Showed* mostrar mostrou ele mostrou né ele mostrou ele quem? O cara lá o Krotov mostrou a possibilidade a possibilidade para trazer para levar a concepção até *line* .. com outros países .. até *line* .. não sei até *into* até acho que é até até com outros países *without* reduzindo a produtividade .. produtividade assim né dessa maneira reduzindo a produtividade.

RESUMO: Quarto parágrafo vamo ver de novo Tá O problema econômico corrente o sistema econômico corrente na União Soviética onde o monopólio onde o estado onde o estado tem o monopólio dos recursos naturais *timber* óleo e gás são virtualmente livres Isso encoraja a proliferação da concepção na indústria Krotov estimou que muitos da união .. muita união muitos nacional entre a União soviética seus dois ou três dos muitos recursos do seu país . A possibilidade de se trazer a concepção entre entre .. o que .. onde com outros países ah .. então aqui no quarto parágrafo fala do do .. sistema econômico na Rússia eles o sistema economico russo tem o monopólio do .. dos recursos naturais né naturais recursos e *timber* e o óleo o gasolina

gás estão também virtualmente livres são livres né E isso faz o que? Isso encoraja .. encorajou né tem encorajado a a profiliação numa concepção das indústrias pelas indústrias e o Krotov fala de novo O Krotov estima o que? O Krotov estimou que para .. que por muitos nacional .. União Soviética os dois os dois para três dos grandes dos muitos recursos têm seu país cada país seu próprio país .. e .. ele .. né o Krotov (...) 'tô voltando .. então falava o quarto parágrafo .. que a .. autonomia política economica o sistema economico na União Soviética né onde o Estado tem o monopólio dos recursos naturais óleo gás e timber né .. também são virtualmente livres né e .. isto tem encorajado a profiliação da concepção pelas indústrias certo? E o Krotov fala a mesma coisa Ele estima que para cada .. que para todos .. a união nacional né .. que tem na União soviética a União soviética usa o que? dois para dois e três de muitos recursos que têm em seu país então ele usa (...) certo então tem poucos né ele tem poucos recursos usa dois ou três dos muitos recursos próprio país e mostrou a possibilidade de trazer a concepção entre cada país entre o próprio país e a redução produtividade Certo? Este é o quarto parágrafo Aqui não entendi as palavras *under timber profligate* /konSuNpʃɪn/ *consumption income line* .. certo?

PRGF. 5: Quinto parágrafo .. *The best way* o melhor o melhor way não sei o que é way eu sei é mas não sei .. *the best way* o melhor para .. /KuR/ *cure* curar indústria para *cure* eu já vi atrás aqui eu consegui deixa eu vê (...) o melhor way para curar industrialmente indústrias de seu .. /Kures/ *cures* eu não sei o que é *uasterFULnes/ wastefulness* também não O melhor método para jeito maneira curar a indústria .. de seu de .. seus problemas sei lá é trocar .. trocar .. o .. os recursos .. trocar o .. *going rate* por recursos naturais *Going rate* deve ser recursos não naturais né Trocar o *going rate* por recursos naturais A melhor maneira para curar a indústria é trocar os recursos sei lá elétricos sei lá financeiros por recursos .. naturais certo? *Maintains Krotov* né falou *Krotov* lá Ele também propôs uma adição de um suplemento ecológico ele propôs a adição de um suplemento ecológico para o price para o /PRɪs/ *price* eu não sei o que que é a /sɜrʃɛɪɡ/ *surcharge* a *surcharge surcharge change* trocar não *change* é trocar .. refletindo a a escassez do recursos refletindo .. refletindo a escassez do recurso .. ou a troca *dam .. a /DEɪmɛɡ/ damage damage* não sei .. para o no meio ambiente *Under* o velho sistema antes no velho sistema *research .. research ..* experiência pesquisa pesquisa entre os cientistas entre a ciência e a tecnologia mostrou *needed* deve ser mostrou não tecnologia *needed ..* para .. /KLɪən/ *clean needed* isso aparece em todos os textos tem que saber essa palavra vou decorar No velho sistema pesquisa entre a ciência e .. entre a ciência mostrou a ciência e tecnologia mostrou *clean clean* claro para a indústria também estagnado certo *It* quase estava justo não /prɒfɪtəbəl/ *profitable* proliferando para os fatores *managers* próprios para as indústrias de *cleaner* clara safe tecnologia disse Krotov Então Krotov disse mais uma coisa aqui (...) o melhor meio o melhor meio para sair da indústria então o quinto parágrafo fala do que? O melhor meio da indústria para sair para curar a indústria é trocar os métodos atuais pelos métodos .. pelos recursos naturais atuais pelos recursos naturais disse Krotov Ele também propôs a edição de um suplemento ecológico para primeiro para .. para *price* e *surcharge* refletindo a escassez falta de recurso .. refletindo a falta de recurso um suplemento ecológico refletindo a falta de recurso pra mostrar a falta de recurso no meio ambiente e o velho sistema .. e no velho sistema né .. feito uma pesquisa né de cientistas e tecnologias de ciência e tecnologia mostrando a indústria também estagnou que a indústria também está parada está falha estagnada e foi justamente proliferada por fatores próprios das indústrias livres A indústria livre também preocupada com as indústrias livres fazendo *safe* eu não sei o que que é .. tecnologia disse Krotov.

RESUMO: Então o quinto parágrafo fala como que pode curar lá a indústria e também com propostas do Krotov de criar um suplemento ecológico para que? Para alertá para mostrá a falta de recursos naturais no meio ambiente E fala também do sistema antigo do velho sistema onde onde a pesquisa ciência e tecnologia mostrou a indústria também estagnada e .. foi justamente não proliferado por fatores próprios das indústrias livres não *clean* das indústrias *clean* não é livre não das indústrias fatores das indústrias da .. tecnologia Então por causa dessa poluição aí as indústrias desenvolveram tecnologicamente pesquisas descobriu isso né (...) Agora .. ah ah seis ..

PRGF. 6: *Krotov's* curou curou o que é *cure* /KuR/ *the cure* os curadores não .. *Krotov's* curou para a União Soviética meio ambiente curou *cure* não é curou não trouxe disse ou melhor .. *Krotov's* /KuR/ *cure cure cure* /Kure/ para a União Soviética meio ambiente *ills* então Krotov disse para o meio ambiente da União Sovi .. meio ambiente da União Soviética .. *ills* não sei já falei né para privatizar /lɔːst/ *least* /Hɔːlfi/ o passado .. /lɔːst/ *least* semana /lɔːst/ *least* dos fatores menos /Hɔːlfi/ *half* num sei do estado dos fatores estaduais para introdução da competição *among* entre os produtores Ah certo! então Krotov disse que a União soviética .. então Krotov cura curou curou pra cura né pra União Soviética no ambiente Soviético .. curar no ambiente soviético Krotov disse privatizar menos nos .. que os fatores est .. privatizar menos todos .. de os estados os fatores de introdução na competição entre produções Suas .. seus involvimento .. isto envolveria envolveria *would* envolveria existência existindo companhias *set* por pondo novas /ɪntɜːprɪtɪsɪs/

enterprises novas enterprises e get get got got pegar não .. get got gotten (...) get dar? não .. pegar acho que é pegar .. pegar isso envolveria a existência de existindo companhias pondo novos interesses e .. pegando envolvendo pegando um /ʒɔɪnvenTʃəz/ joint ventures com .. /fɔɪŋɡ/ foreign firmas com várias firmas /ʒɔɪnvenTʃəz/ joint venture joint venture é fazer .. unir firmas .. unir firmas uma faz trabalho pra um outra faz o trabalho pra outra joint venture isso aí eu sei .. tá vendo? joint ventures sistema japonês como no Japão joint venture mas a União Soviética mostrou mais .. mostrou .. mos .. mostrou .. /mʊvʉ/ move away não sei para heavy para sua indústria é /stʌp/ step up a produção de consumers os bons consumos ele disse .. ele disse Quem? O Krotov Mas a União Soviética mostra mais um .. heavy metal heavy indústria e a produção de consumer goods ele disse Mais importante importante no país must join international monetary sistema também que o /hɒbəl/ rouble becomes /kɒnvertəbəl/ convertible .. convertido convertible become became becom become became become saber conhecer é .. A União mas a União Soviética o maior serão a indústria heavy pesada indústria pesada e uma produção de consum.. de bom de consumos de bons consumidores consumos ele disse Mais importantemente o país mostrar intern .. moneta .. join .. não sei mostrar sistema monetário internacional do que saber converter /hʌbl/ não sei.

RESUMO: Então o quinto parágrafo Krotov fala o que? Como a União Soviética /kʊr/ cure Krotov's cure da União Soviética meio ambiente privatizar menos porque .. que .. de .. os .. estado os fatores estaduais para introduzir a competição entre os produtores Então .. privatizar .. menos .. os menos os fatores .. do meio ambiente porque pra ver maior competição entre o que? Entre os produtores certo? E isto envolveria a existência de companhias novos interesses com novos interesses e pegando envolvendo /ʒɔɪnvenTʃəz/ joint ventures com novas firmas Então joint ventures sistema japonês certo? a união de novas .. união de firmas sistema japonês Um faz o trabalho pra outro uma troca de favores joint ventures A União .. mas a União mas ele fala mas a União Soviética tem uma indústria muito pesada de consumo bom de bom consumo indústria pesada certo? De bom consumo ele disse o coisa aí Mas mais importantemente que o país mostrar a inter monetary .. o sistema monetário internacional do que saber conhecer convertido né fazer isso daí Mais importante o país tá bem com o sistema monetário internacional né porque fazer .. isso nas joint ventures essas coisas novas certo? Então esse é o quinto parágrafo Quinto não sexto Nesse parágrafo não sei o que é /kʊr/ cure /hʌlf/ half(...) /mʊvʉ/ move away from .. move away .. /stʌp/ step .. consumer goods sei mais ou menos vão vê né joint .. must joint então venture é .. sistema japonês então é joint venture .. the international monetary system rubro é /hɒbəl/ rouble isso daí .. certo .. no segundo (...) agora o sexto .. certo (...) Em seguindo um segundo vital componente de Krotov's /sɛlm/ scheme não sei é para dar local governadores o local over governador governador controlar os recursos nos seus nas suas regiões Então um segundo componente vital de Krotov é o governador controlar os recursos na sua região por handing hand é mão dando mão por hand é mão handing .. mão mão .. mão por que .. o que .. agora própria o que proprietário do estado .. Isso daria governadores locais uma .. uma direção um direto interesse na prevenção da população e proteção do ambiente local certo? Então isso daria aos governadores locais um direto interesse na prevenção prevenção prevenimento não no prevenimento prevenção da população e no .. proteção .. dos ambientes dos seus ambientes dos seus ambiente local Povos com a greatest com grande .. stake no meio ambiente são o povo local Ah então eles vão tratar do que é deles mesmo então eles não vão querer certo estragá o ambiente perto deles ele disse. Ecological standards ecológicos e /hʉl/ /hʉl/ rules não sei podem podem próprios apelidados eficientemente quando as autoridades locais são muitos agentes para monitorando e "enforçamento" e "esforçamento" "enfortemento" não sei Então sexto parágrafo sétimo aqui.

PRGF. 7: Vital .. segundo .. vital o segundo componente vital de Krotov's /sɛlm/ scheme não sei é trazer é dar né é eleger acho que é é botar governadores .. então um segundo componente vital para ele ele Krotov é botar governadores .. dar gover .. controle aos governadores certo? pro local .. pros locais de suas próprias regiões certo? Então o governador vai controlar o meio ambiente de sua própria região certo? de sua própria região Assim fazendo agora sua própria proprietário por handing handing não sei que .. o que agora estado proprietário Isto .. isto daria .. o que .. os governadores locais uma direta faria com que os governadores locais tivessem um direto interesse em proteger o ambiente porque é da sua região certo? Porque é o local de seu povo ele disse tá porque .. direta proteção prevenindo o meio ambiente e protegendo o local meio ambiente Certo? então isso daria uma direta .. prevenção da poluição da população de poluição né e um e a proteção do meio ambiente .. tá então os .. então os .. os com grandes ... gostam de seus próprios meio ambiente né não vai querer que sujam que .. seu local e a ecologia .. ecological standards tá próprio ... então o seu próprio efetivamente quando o local das autoridades são agentes monitorando eficientes Então vai ficar mais eficiente o controle sobre a população porque o governador vai tomar conta de sua própria .. região acho que é isso tá.

RESUMO: Esse é o sétimo parágrafo Então o segundo .. o segundo componente vital pra Krotov é que os governadores controlem os recursos de suas regiões .. e' novas propriedades novos estados propriedades certo

de suas próprias regiões porque isso dá o que? Isso dá aos governadores os recursos na .. um direto interesse na prevenção da população da poluição e proteção do meio ambiente e para os grandes não querem que o meio ambiente seus .. né vantagens pro povo ele disse Ecologia .. é própria tem mais .. eficiência quando as autoridades .. agentes são agentes de seu próprio ambiente certo? Certo Esse é o sét .. parágrafo sétimo parágrafo Agora o oitavo .. parágrafo .. oitavo parágrafo.

PRGF. 8: A Rússia a república russa contamina *contains contains* contamina mais do que a soviética União dos recursos naturais *Contains* não é contamina .. a república russa contamina muitos .. dos seus recursos naturais dos seus recursos naturais .. e tem grande .. *greatest* meio ambiente grande problema ambiental Mas o governador da república *led* .. por Bóris Yeltsin tem *already* começado? .. *begin began begun* .. *begin* começar tem começado .. para *take take taken take took /ToK/ taken* pegar *take control* não *take* pegou *book* pegar livro não tá .. acho que é pegar sim tem pegado */STtPS/* para remédio remdiar os problemas esses problemas esses próprios problemas Recentemente anun .. anunciou a extensão controle o controle dos recursos naturais */uuDtN/* seus *within /bonDarts/ boundaries* Yeltsin tem introduzido 400 programa de 400 dias para revisar a economia russa com *backing* com *backing* para Presidente Gorbatchov E para 1º de janeiro em inovação poluição *pays pay* pagar pagar a poluição policia *come came come* dá uma força Então policia dão a força ..

RESUMO: Certo então o parágrafo aqui A república russa continha muitos de seus recursos naturais e tem um grande problema ambiental Mas o governador da república *led* não sei .. *led* descrito por Bóris Yeltsin falando por Bóris Yeltsin tem */aRə-di/ already* começado a pegar */STtPS/ steps* para remediar seus problemas *already* não sei *remedy* acho que é remediando seus problemas diminuindo os problemas é remediar é remediar acho que é remediar remediar não prevenir seus problemas remediar E recentemente anunciou a extensão do controle dos recursos naturais */uuDtN/ within* não sei *its bound its /BONdarts/ boundaries /bonDarts/ boundaries* também não Yeltsin tem introduzido esse deve ser o cara lá Tem introduzido ..[leitor interrompe a tarefa de leitura].

TAREFA II: GRIFAR PALAVRAS

PE: Agora né você vai sublinhar as palavras e vai falar ao mesmo tempo tá? Então vai (...)

LIE: E' .. */KRizis/ crisis* mas eu consegui inferir .. É */Htʃɪŋg/ reaching* (... ..)

PE: Grifa todas mesmo que você não soubé.

LIE: E' tem umas que eu sei mas não tenho certeza ..

PE: Essas você grifa também.

LIE: Tenho que grifar?

PE: Isso .. tá.

LIE: Então tá (...) *way .. solve (... ..) /Htʃɛd/ reached* (...)

PE: *Stage* você sabe?

LIE: Eu não sei totalmente não.

PE: Tem uma idéia?

LIE: É.

PE: Que você acha que que é?

LIE: *Estágio*.

PE: Ah ah!

LIE: É isso mesmo?

PE: Ah ah!

LIE: */Hə-lt/ health* aqui eu acho que é quente mas num (...) */uanʃə-dlɛd/ Unscheduled* (...) */DRɛuN/ drawn* .. */PuLp/ pulp* .. */DəR/ there* (...) *shores* .. acho que mostrar .. num sei .. */BLɛM/ blame* .. */narouMɪNdɛd/ narrow-minded* .. aqui acho que é ..

PE: *People people* você não grifou né?

LIE: Não eu sei .. *blame* eu já grifei antes .. *itself* .. eu sei mais ou menos eu não sei ..

PE: Fala o que você acha que é.

LIE: Eu já vi mas esqueci isso aqui eu vi ...

PE: ah ah ..

LIE: */BɛNG/ /Bə-NG/ being .. banned* (...) *under* (...) *timber* (...) */ɛnKoUraʒ/ encourage* acho que é encorajar .. não sei .. *encourage .. encourage* (...) *profligate /kɒnsə-npʃɪən/ consumption income* (...) *line* .. */uɪfəut/ without /uɪfəut/ without /KuR/ cure* .. */uastɜFulnɛs/ wastefulness wastefulness /Goɪŋg/ going /Hɛt/ rate?*

PE: *Resources* você conseguiu?

LIE: *Resource* pra mim eram recursos .. naturais

PE: ah ah.

LIE: /PRaɪs/ Price .. /SuRʃɑːʒ/ surcharge .. /ɛskærstɪ/ scarcity .. /dæɪmɛɪʒ/ damage .. /Don/ done (...) needed de novo tenho que decorar essa palavra todo texto tem /KLɪən/ clean ..

PE: Clean?

LIE: É /KLɪn/ clean .. /stɑːɡneɪtɪd/ stagnated (...) managers /Saf/ safe /seɪf/ safe (...) o ill aqui do título eu não entendi o que ..

PE: Ah ah apareceu de novo.

LIE: /Haɪf/ Half (... ..) /ɪntərpraɪz/ enterprise pra mim era interesse eu não sei ..

PE: Ah ah!

LIE: /foʊraɪn/ foreign .. /MOVUː/ move away .. step .. join .. /Rublo/ rublo .. /ʃeɪm/ scheme.. outro parágrafo.

PE: Ah ah!

LIE: /ʃeɪm/ scheme.. /over/ over over handing eu sei que hand é mão handing não sei (...) /GREɪtɛst/ greatest .. standards /Hʊlɪs/ .. /rɛɪnfɔːsmənt/ reinforcement .. contains (...) /Ləːdi .. /aRəːd/ already .. /rɛmɛd/ remedy..

PE: To take steps você entendeu?

LIE: Onde 'ta?

PE: Um pouquinho antes led by Boris Yeltsin has already begun to take steps ..

LIE: Ah step não sei take é .. take /ToK/ took take .. take take .. recente recentemente .. /uɪDɪn/ within its /baʊndəɪrɪs/ boundaries ..

PE: É tudo isso que você não sabe? Então grifa tudo ..

LIE: Backing .. pays .. pay pagar deve ser .. parei aqui

PE: Você parou aí né? Policy .. polluter pays .. policy ..

LIE: Policy ..

PE: Policy você sabe o que é?

LIE: Policy .. acho que é policia não sei polluter pays parei um pouco antes até aqui.

PE: Ah ah tá bom. Marca então põe uma cruzinha aí onde você parou.

BIBLIOGRAFIA

- ABRAHAM, R.G. & C.A. CHAPPELLE. The meaning of cloze test scores: an item difficulty perspective. *The Modern Language Journal*, 76(4), 1992:468-79.
- ADAMS, S.J. Scripts and the recognition of unfamiliar vocabulary: enhancing second language reading skills. *The Modern Language Journal* 66(2), 1982:155-159.
- AISENSTADT, E. Restricted collocations in English lexicology and lexicography. *ITL* 53, 1981:53-61.
- AITCHINSON, J. *Words in the mind*. Oxford: Blackwell, 1987.
- ALDERSON, J.C. The cloze procedure and proficiency in English as a foreign language. *TESOL Quarterly* 13(2), 1979:219-27.
- _____. Reading: a reading problem or a language problem? In J.C. Alderson & A. Urquhart, eds., *Reading in a foreign language*. New York: Longman, 1984:1-27.
- _____. S. BASTIEN, & A. MADRAZO. A comparison of reading comprehension in English and Spanish. *Research and Development Unit: report no. 9*. Mexico City: Universidad Nacional Autonoma de Mexico, Centro de Ensenanza de Lenguas Extranjeras, 1977a.
- _____. & S. RICHARDS. Difficulties which students encounter when reading texts in English. *Research and Development Unit Report no. 8*. Mexico City: Universidad Nacional Autonoma de Mexico, Centro de Ensenanza de Lenguas Extranjeras, 1977b.
- ALLEN, V.F. *Techniques in teaching vocabulary*. New York: Oxford University Press, 1983.
- ALLWRIGHT, R., S. PIT-CORDER & R. ROSSNER. Talking shop: language teaching and Applied Linguistics. *English Language Teaching Journal* 40(3), 1986:185-191.
- AMES, W.S. The development of a classification scheme of contextual aids. *Reading Research Quarterly* 2(1), 1966:57-82.
- ANDERSON, J.R. Acquisition of cognitive skill. *Psychological Review* 89(4), 1982:369-406.
- ANDERSON, R.C. & Z. SHIFFRIN. The meaning of words in context. In R.J. Spiro, B.C. Bruce & W.F. Brewer, eds., *Theoretical issues in reading comprehension*. Hillsdale, NJ: Erlbaum, 1980:330-348.
- _____. & P. FREEBODY. Reading comprehension and the assessment and acquisition of word knowledge. In R. Hutson, ed., *Advances in reading/language research: a research annual*. Greenwich CT: JAI Press, 1983.
- ARON, H. Comparing reading comprehension in Spanish and English by adult Hispanics entering a two-year college. In Proceedings of the Third International Conference on Frontiers in Language Proficiency and Dominance Testing, R. O. Silverstein ed., *Occasional Papers on Linguistics*, Carbondale, Ill.: Southern Illinois University, Department of Linguistics, 1980:165-173.

- ARROJO, R. & K. RAJAGOPALAN. O ensino da leitura e a escamoteação da ideologia. *Revista Letras* 10(1/2). Campinas: Pontificia Universidade Católica de Campinas, 1991:45-49.
- ATKINSON, R.C. Mnemotechnics in second-language learning. *American Psychologist* 30, 1975:821-28.
- _____ & RAUGH, M. An application of the mnemonic key word method to the acquisition of Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory* 1 (2), 1975:126-33.
- BAKER, L. & A.L. BROWN. Cognitive monitoring in reading. In J. Flood, ed., *Understanding reading comprehension*. Newark, Delaware: International Reading Association, 1984:21-44.
- BALTRA MONTANER, A.H. *Reading for academic purposes: an eclectic exploration into reading theories and practical classroom applications*. São Paulo: Pontificia Universidade Católica. Tese de doutorado, 1982.
- BARBOSA, A.C.G. "Gramatiquice" ou "estrategice" — é esta a questão? Comunicação apresentada no I Intercâmbio de Pesquisas em Linguística Aplicada. São Paulo: Pontificia Universidade Católica, 1990.
- BARIK, H. & M. SWAIN. Three-year evaluation of a large-scale early grade French immersion program: The Ottawa study. *Language Learning* 25, 1975:1-30.
- BARNARD, R. et alii. Reading strategies. In H. Brookes, G. Pozzo & T. Rowe, eds., *Reading in a foreign language*. Milan: Giuseppina Cortese, 1980:400-15.
- BARNETT, M.A. Syntactic and lexical/semantic skill in foreign language reading: importance and interaction. *Modern Language Journal* 70, 1986:343-349.
- _____ Reading through context: how real and perceived strategy use affects L2 comprehension. *The Modern Language Journal* 72, 1988:150-62.
- BARRERA-VAZQUEZ, A. The Tarascan project in Mexico. In *The use of the vernacular in education. Monographs in Fundamental Education* 8. Paris: UNESCO, 1953:77-86.
- BARTLETT, F.C. *Remembering*. Cambridge: Cambridge University Press, 1932.
- BEAUGRANDE, R. Information, expectation and processing: on classifying poetic texts. *Poetics* 7, 1978:3-44.
- BECK, I.L., C.A. PERFETTI & M.G. McKEOWN. The effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology* 74, 1982: 506-21.
- BENSOUSSAN, M. & B. LAUFER. Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading* 7(1), 1984:15-32.
- BENZÈCRI, J.-P. *L'Analyse des Données* 1-2. Paris: Dunod, 1973.
- BIALYSTOK, E. & M. SHARWOOD-SMITH. Interlanguage is not a state of mind in an evaluation of the construct for second-language acquisition. *Applied Linguistics* 6(2),

- BIBER, D. *A model of textual relations within the written and spoken modes*. University of Southern California. Tese de doutorado, 1984.
- BLOCK, E. The comprehension strategies of second language readers. *TESOL Quarterly* 20, 1986:463-94.
- BORDINI, M.G. *Literatura: a formação do leitor*. Porto Alegre: Mercado Aberto, 1988.
- BRAGA, D. & J. BUSNARDO. A contribuição da ensino da leitura para a deformação do leitor. *Trabalhos em Lingüística Aplicada* 3. Campinas: Universidade Estadual de Campinas, 1984:23-36.
- _____, J. BUSNARDO, M.G. MORAES & M.V.R. SCARAMUCCI. O ensino de leitura em inglês em um contexto de língua estrangeira. Campinas: Universidade Estadual de Campinas. Manuscrito não publicado, 1992.
- BRIGHT, J.A & G.P. MCGREGOR. *Teaching English as a Second Language*. London: Longman, 1970.
- BROUGHTON, M.M. The ESP component in Applied Linguistics: A British view. *The ESPECIALIST* 1. São Paulo: Pontifícia Universidade Católica, 1980.
- BROWN, P. & G. YULE. *Discourse analysis*. Cambridge: Cambridge University Press, 1983.
- BROWN, D.F. Advanced Vocabulary Learning: the problem of collocation. *RELC Journal* 5(2), 1974.
- BROWN, A.L. Metacognitive development & reading. In R.J. Spiro, B.C. Bruce & W.F. Brewer, eds. *Theoretical issues in reading comprehension*. Hillsdale, NJ.: Lawrence Erlbaum, 1980:453-81.
- BRUCE, B. & A. HUBIN. Readability formulas: matching tool and task. In A. Davison & G.M. Green, eds., *Linguistic complexity and text comprehension. Readability issues reconsidered*. Hillsdale: Lawrence Erlbaum Associates, 1988.
- BUSNARDO, J. & M.G. MORAES. Negociando o sentido (elogio da ignorância). *Trabalhos em Lingüística Aplicada* 1. Campinas: Universidade Estadual de Campinas, 1983:9-33.
- CAMPION, M.E. & W.B. ELLEY. *An academic vocabulary list*. Wellington: New Zealand Council for Educational Research, 1971.
- CANALE, M. From communicative competence to communicative language pedagogy. In J. Richards & R. Schmidt, orgs., *Language and Communication*. London: Longman, 1983.
- _____, & M. SWAIN. Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1(1), 1980:1-47.
- CARRELL, P. Second language acquisition in classroom instruction: reading. In *Annual Review of Applied Linguistics* 9. Cambridge: Cambridge University Press, 1988:223-242.
- _____. Metacognitive awareness and second language reading. *The Modern Language*

Journal 73(2), 1989:121-134.

_____. Second language reading: reading ability or language proficiency? *Applied Linguistics* 12(2), 1991:159-179.

CARROLL, J.B., P. DAVIES & B. RICHMAN. *The American heritage word frequency book*. New York: Houghton Mifflin, Boston American Heritage, 1971.

CARTER, R. A note on core vocabulary. *Nottingham Linguistic Circular* 2(2), 1982:39-51.

_____. Core vocabulary and discourse in the curriculum: a question of subject. *RELC Journal* 17(1) 1986:52-67.

_____. Is there a core vocabulary? Some implications for teaching. *Applied Linguistics* 11, 2(8), 1987a.

_____. Vocabulary and second/foreign language teaching. *Language Teaching* 20(1), 1987b:3-16.

_____. & M. McCARTHY. *Vocabulary and language teaching*. Longman: London, 1988.

CAVALCANTI, M. do C. *The pragmatics of FL reader-text interaction: key lexical items as source of potential reading problems*. Lancaster: Universidade de Lancaster. Tese de doutorado, 1983.

_____. *Interação leitor-texto: aspectos de interpretação pragmática*. Campinas: Editora da UNICAMP, 1989.

_____. Intersubjetividade e pluralidade na leitura de uma lacuna de fábula. In M.S.Z. de Paschoal & M.A. Celani, orgs., *Linguística aplicada: da aplicação da linguística à linguística transdisciplinar*. São Paulo: EDUC, Pontifícia Universidade Católica, 1992:223-229.

CELANI, M.A. O ensino de Inglês Instrumental em universidades brasileiras. *The ESPecialist* 3. São Paulo: Pontifícia Universidade Católica, 1981.

_____. et alii. *The Brazilian ESP project: an evaluation*. Paulo: CEPRIL/EDUC, Pontifícia Universidade Católica, 1988.

CHALL, J. *Readability: an appraisal of research and application*. Ohio State Bureau of Education Research, 1958.

CHANELL, J. Applying semantic theory to vocabulary teaching. *ELT Journal* 35(2), 1981:115-122.

CHOMSKY, N. *Syntactic structures*. The Hague: Mouton & Co, 1957.

_____. *Aspects of the theory of syntax*. Cambridge: M.I.T Press, 1965.

CICOUREL, A. *Cognitive sociology: language and meaning in social interaction*. The Free Press, 1974.

CLARKE, M. The short-circuit hypothesis of ESL reading - or when language competence interferes with reading performance. In P. Carrell, J. Devine and D. Eskey, eds., *Interactive*

- approaches to second language reading*. Cambridge: Cambridge University Press, 1988:114-124.
- CLARKE, D. & I.S.P. NATION. Guessing the meaning of words from context: strategy and techniques. *System* 8(3), 1980:211-20.
- COADY, J.M. A psycholinguistic model of the ESL reader. In R. Mackay et alii, eds., *Reading in a second language*. Rowley, Mass.: Newbury House, 1979:5-12.
- COHEN, A.D. Recent uses of mentalistic data in reading strategy research. *DELTA* 3(1). São Paulo: Pontificia Universidade Católica, 1987:57-84.
- COLLERSON, J.W. *An approach to the analysis of semantic cohesion in text*. Lancaster: University of Lancaster. Dissertação de mestrado, 1974.
- COOPER, M. Linguistic competence of practised and unpractised nonnative readers of English. In J.C. Alderson & A. Urquhart, eds., *Reading in a foreign language*. New York: Longman, 1984:122-135.
- CORACINI, M.J.R. *Um fazer persuasivo: o discurso subjetivo da ciência*. São Paulo: Pontes/EDUC, 1991.
- _____ et alii. *O enigma do rei*. Anais do III Encontro da Anpoll. Rio de Janeiro, 1988.
- CORSON, D.J. The Corson measure of passive vocabulary. *Language and Speech* 26, 1983:3-20.
- COWAN, J.R. & Z. SARMAD. Reading performance of bilingual children according to type of school and home language. *Language Learning* 26, 1976:353-379.
- COWIE, A.P. EFL dictionaries: past achievements and present needs. In R.R.K. Hartmann, ed., *LEXeter '83 Proceedings*. Tübingen: Niemeyer, 1984.
- CRONBACH, L.J. An analysis of techniques for diagnostic vocabulary testing. *Journal of Educational Research* 36, 1942:206-217.
- CROTHERS, E. & P. SUPPES. *Experiments in second-language learning*. New York: Academic Press, 1967.
- CROW, J.T. Receptive vocabulary acquisition for reading comprehension. *The Modern Language Journal* 70(2), 1986:242-250.
- _____ & J.R. QUILEY. A semantic field approach to passive vocabulary acquisition for reading comprehension. *TESOL Quarterly* 19(3), 1985:497-513.
- CUMMINS, J. Cognitive/academic language proficiency, linguistic interdependence, and optimum age question and some other matters. *Working Papers on Bilingualism* 19, 1979:197-205.
- CZIKO, G. Language competence and reading strategies: a comparison of first and second-language oral reading errors. *Language Learning* 30(1), 1980:101-116.
- DALE, E. Vocabulary measurement: techniques and major findings. *Elementary English* 42, 1965:895-901.

- _____. R. RAZIK. *Bibliography of vocabulary studies* (2nd edition). Ohio: Ohio State University, Bureau of Educational Research and Service, 1963.
- DANK, M.E. *A study of the relationship of miscues to the model of formal reading instruction received by selected second graders*. University of Massachusetts. Tese de doutorado, 1976.
- DASCAL, M. Una critica reciente a la noción de significado literal. In *Crítica* 18(53), México, 1986.
- _____. Contextualism. In H. Parret, ed., título não disponível, 1981:153-177.
- DAVIES, F.B. The interpretation of frequency ratings obtained from "The teacher's word book". *Journal of Educational Psychology* 35, 1944:169-174.
- _____. Psychometric research on comprehension in reading. *Reading Research Quarterly* 7, 1972: 628-678.
- DELAWARE, J.A. The relationship of beginning reading instruction and miscue patterns. In W.D. Page, ed., *Help for the reading teacher*. Urbana, Ill.: National Council for Teachers of English, 1975.
- DEVINE, J. The relationship between general language competence and second language reading proficiency: implications for teaching. In P. Carrell, J. Devine & D. Eskey, eds., *Interactive approaches to second language reading*. Cambridge: Cambridge University Press, 1988:260-78.
- DEYES, T. Towards an authentic "discourse cloze". *Applied Linguistics* 5(2), 1984:128-137.
- ELLEY, W. & F. MANGUBHAI. The long-term effects of a book flood on children's language growth. *Directions* 7, 1981:15-24.
- ERICKSON, F. Advantages and disadvantages of qualitative research design on foreign language research. In Barbara Freed, org., *Foreign language acquisition research*. D.C. Heath & Company, 1991:338-353.
- ESKEY, D.E. A model program for teaching advanced reading to students of English as a foreign language. *Language Learning* 23(2), 1973:169-184.
- _____. Holding in the bottom: an interactive approach to the language problems of second language readers. In P. Carrell, J. Devine and D. Eskey, eds., *Interactive approaches to second language reading*. Cambridge: Cambridge University Press, 1988:93-100.
- _____. D. & W. Grabe. Interactive models for second language reading: perspective on instruction. In P. Carrell, J. Devine and D. Eskey, eds., *Interactive approaches to second language reading*. Cambridge: Cambridge University Press, 1988:223-238.
- FAERCH, C. & G. KASPER. Procedural knowledge as a component of foreign language learners' communicative competence. In H. Bolte & W. Herrlitz eds., *Kommunikation im Sprachunterricht*. Utrecht: Rijksuniversiteit, 1985:169-99.
- _____. From product to process - introspective methods in second language research. In C. Faerch & G. Kasper, eds., *Introspection in second language research*, Clevedon, 1987:5-23.

- FAIRCLOUGH, N.L. Critical and descriptive goals in discourse analysis. *Journal of Pragmatics* 9, 1985:739-763.
- FARR, R. *Reading: what can be measured?* Newark, DE: International Reading Association, 1969.
- FISHER, R. A. The precision of discriminant functions. *Ann Eugene London* 10, 1940:422-429.
- FLAVELL, J.H. Cognitive monitoring. In W.P. Dickson, ed., *Children's oral communication skills*. New York: Academic Press, 1981:35-60.
- FOTOS, S.S. The cloze test as an integrative measure of EFL proficiency: a substitute for essays on college entrance examinations. *Language Learning* 41(3), 1991:313-336.
- FRIES, C.C. *Teaching and learning English as a foreign language*. Ann Arbor: University of Michigan Press, 1945.
- GAIRNS, R. & S. REDMAN. *Working with words*. Cambridge: Cambridge University Press, 1986.
- GALVES, C. Reflexões sobre o lugar da gramática no ensino de leitura em língua estrangeira. *Trabalhos em Lingüística Aplicada* 1. Campinas: Universidade Estadual de Campinas, 1983:69-86.
- _____. & J. BUSNARDO. Leitura em língua estrangeira e compreensão e produção de textos em língua materna. *Redação e Leitura*. São Paulo: Pontifícia Universidade Católica, 1983:305-311.
- GERLOFF, P. Identifying the unit of analysis in translation: some uses of think-aloud protocol data. In C. Faerch & G. Kasper, eds., *Introspection in second language research*. Clevedon, 1987.
- GIDDENS, A. *New rules of the sociological method: a positive critique of interpretative sociologies*. London: Hutchinson, 1976.
- GIFI, A. *Nonlinear multivariate analysis*. The Netherlands: Leiden, University of Leiden, Department of Data Theory, 1981.
- GLEASON, H.A. *An introduction to descriptive linguistics*. New York: Holt, Rinehart, 1961.
- GOODMAN, K.S. Psycholinguistic universals in the reading process. In P. Pimslei & T. Quinn, eds., *The psychology of second language learning*. Cambridge: Cambridge University Press, 1971:135-142.
- _____. Reading: a psycholinguistic guessing game. In H. Singer and R. Ruddell, eds., *Theoretical models and processes of reading*, 2nd ed., Newark: International Reading Association, 1973a (Originalmente publicado em 1967).
- _____. Psycholinguistics universals of the reading process. In F. Smith, ed., *Psycholinguistics and reading*. New York: Holt, Rinehart & Winston, 1973:21-29.
- _____. Behind the eye: what happens in reading. In H. Singer & R. Ruddell, eds., *Theoretical models and processes of reading*. Newark, Delaware: International Reading Association, 1976.

- _____. The reading process. In P. Carrell, J. Devine and D. Eskey, eds., *Interactive approaches to second language reading*. Cambridge: Cambridge University Press, 1988:11-21.
- GOODWIN, C. & A. DURANTI. Rethinking context: an introduction. Em C. Goodwin and A. Duranti, eds., *Rethinking context: language as an interactive phenomenon*. Cambridge: Cambridge University Press, 1992:1-42.
- GOUGH, P. One second of reading. In H. Singer & R. Ruddell, eds., *Theoretical models and processes of reading*. Newark, Delaware: International Reading Association, 1976.
- GRABE, W. Reassessing the term 'interactive'. In P. Carrell, J. Devine and D. Eskey, eds., *Interactive approaches to second language reading*. Cambridge: Cambridge University Press, 1988:56-70.
- GREENACRE, M.J. *Theory and applications of correspondence analysis*. New York: Academic Press, 1984.
- GRICE, H.P. Utterer's meaning, sentence meaning, and word-meaning. *Foundation of Language* 4, 1968:225-42.
- _____. *The thread of discourse*. The Hague: Mouton, 1975.
- GROTJAHN, R. On the methodological basis of introspective methods. In C. Faerch & G. Kasper, eds., *Introspection in second language research*. Clevedon, 1987:54-55.
- GUTTMAN, L. The quantification of a class of attributes: a theory and method of scale construction. In P. Horst, ed. *The prediction of personal adjustment*. Social Science Research Council. New York, 1941:319-348.
- HAARSTRUP, K. Using thinking aloud and retrospection to uncover learners' lexical inferencing procedures. In C. Faerch & G. Kasper, eds., *Introspection in second language research*. Clevedon, 1987:197-212.
- _____. *Lexical inferencing procedures or talking about words*. Copenhagen: Copenhagen Business School. Tese de doutorado, 1989a.
- _____. The effect of context on learners' intake of new words. Comunicação apresentada na AILA Conference, Greece, 1989b.
- HALLIDAY, M.A. *Spoken and written language*. Oxford: Oxford University Press, 1989.
- _____. & R. HASAN. *Cohesion in English*. London: Longman, 1976.
- HAMMADOU, J. Interrelationships among prior knowledge, inference, and language proficiency in foreign language reading. *The Modern Language Journal* 75(1), 1991:27-37.
- HAMP-LYONS, L. Two approaches to teaching reading: a classroom based-study. *Reading in a Foreign Language* 3, 1985:363-73.
- HARRISON, C. *Readability in the classroom*. Cambridge: Cambridge University Press, 1980.
- HARSTE, J.C. & C.L. BURKE. A new hypothesis for reading teacher research: both *teaching* and

- learning of reading are theoretically based. In P.D. Person, ed, *Reading: theory, research and practice*. 26th Yearbook of the National Reading Conference. Clemson: National Reading Conference, 1977:32-40.
- HAYNES, M. Patterns and perils of guessing in second language reading. In J. Handscombe et alii, eds., *On TESOL '83*. Washington, D.C: TESOL, 1983:163-176.
- HEILENMAN, L.K. The use of a cloze procedure in foreign language placement. *The Modern Language Journal* 67(2), 1983:121-125.
- HENNING, G.H. Remembering foreign language vocabulary: acoustic and semantic parameters. *Language Learning* 23, 1973:185-96.
- HENRIQUES, L. Conscientizando os alunos sobre os efeitos da inferência lexical na compreensão textual. Artigo apresentado no IV Seminário de ESP para ETFs, São Paulo, 1990.
- HILL, M. O. Correspondence analysis: a neglected multivariate method. *Applied Statistics* 23(3), 1974:340-354.
- HINOFOTIS, F.B. Cloze as an alternative method of ESL placement and proficiency testing. In J. W. Oller & K. Perkins, eds., *Language tests at school: a pragmatic approach*. London: Longman, 1979.
- HIRSHFELD, H.O. A connection between correlation and contingency. *Cambridge Philosophical Society Proceedings* 31: 1935:540-544.
- HOCK, T.S. The role of prior knowledge and language proficiency as predictors of reading comprehension among undergraduates. Comunicação apresentada na AILA Conference, 1987.
- HOEY, M. Signalling in discourse. Discourse Analysis Monograph no.6, *English Language Annals* 9, 1976:117-129.
- HOLMES, J. Snarks, quarks and cognates. An elusive fundamental particle in second language reading processes. São Paulo: Pontifícia Universidade Católica, mimeo, 1986.
- HOSENFELD, C.A. Preliminary investigation of the reading strategies of successful and non-successful second language learners. *System* 5, 1977:110-123.
- _____. Case study of ninth grade readers. In A.H. Urquhart & J.C. Alderson, eds., *Reading in a Foreign Language*. New York: Longman, 1984:231-249.
- HORST, P. Measuring complex attitudes. *Journal of Social Psychology* 6, 1935:369-374.
- HUDSON, T. The effects of induced schemata on the "short circuit" in L2 reading: non-decoding factors in L2 reading performance. In P. Carrell, J. Devine and D. Eskey, eds., *Interactive approaches to second language reading*. Cambridge: Cambridge University Press, 1988:183-205.
- HUTCHINSON, T. & A. WATERS. Performance and competence in ESP. *Applied Linguistics* 2, 1981:39-49.
- HYMES, D.H. On communicative competence. In J.B. Pride & J. Holmes, orgs., *Sociolinguistics*.

Harmondworth: Penguin, 1972.

JACOB, E. Qualitative research: a defense of traditions. *Review of Educational Research* 59(2), 1989:229-235.

JENKINS, J.R., D. PANY & J. SCHRECK. Vocabulary and reading comprehension: instructional effects. *Technical Report no. 100*. Urbana, IL: Center for the Study of Reading. University of Illinois (ERIC Document Reproduction Service N. ED 100 999), 1978.

_____, N.L. STEIN & K. WYSOCKI, Learning vocabulary through reading. *America Educational Research Journal* 21(4), 1984: 767-87.

JOHNS, T.F. The text and its message: an approach to the teaching of reading strategies for students of development administration. In H. Faber and A. Maley, eds., *Leseverstehen im Fremdensprachen-Unterricht*. Munich: Goethe Institut, 1978: 147-170.

JOHNSON, P. Effects of reading comprehension on building background knowledge. *TESOL Quarterly* 16, 1982:503-16.

KAMEENUI, E.J., D.W. CARNINE & R. FRESCHI. Effects of text construction instructional procedures for teaching word meanings on comprehension and recall. *Reading Research Quarterly* 17(3), 1982:367-388.

KAPLAN, R.B. Process vs product: problem or strawman. *Lenguas Modernas* 15, 1988:35-44.

KATO, M. Reconhecimento instantâneo e processamento em leitura. *Série Estudos* 8. Uberaba, Minas Gerais, 1982.

_____. Estratégias em interpretação de sentenças e compreensão de textos. *Cadernos da PUC-SP* 16. São Paulo: EDUC/Cortez, Pontificia Universidade Católica, 1983:9-33.

_____. Estratégias gramaticais e lexicais na leitura em língua estrangeira. *Cadernos da PUC-SP* 17. São Paulo: EDUC, Pontificia Universidade Católica, 1984:132-141.

_____. *O aprendizado da leitura*. São Paulo: Martins Fontes, 1985.

KAVALE, K. & R.L. SCHREINER. The reading processes of above average and average readers: a comparison of the use of reasoning strategies to standardized comprehension measures. *Reading Research Quarterly* 15(1), 1979:102-28.

KELLOG, G.S. & M.J.A. HOWE. Using words and pictures in foreign language learning. *Alberta Journal of Educational Research* 17, 1971:89-94.

KERN, R.G. Second language reading strategy instruction: its effects on comprehension and word inference ability. *The Modern Language Journal* 73(2), 1989:135-143.

KLEIMAN, A.B. Estratégias de inferência lexical na leitura de segunda língua. *Ilha do Desterro*. Florianópolis: Universidade Federal de Santa Catarina, 1985:67:82.

_____. *Leitura: ensino e pesquisa*. São Paulo: Pontes, 1989a.

_____. *Texto e leitor - aspectos cognitivos da leitura*. Campinas: Pontes, 1989b.

- _____ & S.B. TERZI. A self-correcting approach to reading in a foreign language. *Ensaio de Lingüística — Cadernos de Lingüística e Teoria da Literatura* 5. Belo Horizonte: Faculdade de Letras, Universidade Federal de Minas Gerais, 1981:9-24.
- KOCH, I.G.V. e L.C. TRAVAGLIA. *Texto e coerência*. São Paulo: Cortez, 1989.
- KRASHEN, S. *Writing: research, theory and applications*. New York: Prentice-Hall, 1984.
- _____ *The input-hypothesis: issues and applications*. New York: Longman, 1985.
- _____ Do we learn to read by reading? The relationship between free reading and reading ability. In D. Tannen, ed., *Linguistics in context: connecting observation and understanding*. Norwood, NJ: Ablex, 1988:269-98.
- _____ We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis. *The Modern Language Journal* 73(4), 1989:440-464.
- KRINGS, H. The use of introspective data in translation. In C. Faerch & G. Kasper, eds., *Introspection in second language research*. Clevedon, 1987.
- KUCERA, H. & W.N. FRANCIS. *A computational analysis of present-day American English*. Providence, R.I.: Brown University Press, 1967.
- LABERGE, D & S.J. SAMUELS. Toward a theory of automatic information processing in reading. In H. Singer & R.B. Ruddell, eds., *Theoretical models and processes of reading*. Newark, Delaware: International Reading Association, 1976.
- LAJOLO, M. *Do mundo da leitura para a leitura do mundo*. São Paulo: Ática, 1994.
- _____ & R. ZILBERMAN. *A leitura rarefeita*. São Paulo: Brasiliense, 1991.
- LARSEN-FREEMAN, D. & M. LONG. Second language acquisition research methodology. In *An introduction to second language acquisition research*. London: Longman, 1991:10-51.
- LAUFER, B. Possible changes in attitudes towards vocabulary acquisition research. *IRAL* 24(1), 1986.
- _____ What percentage of text-lexis is essential for comprehension? Comunicação apresentada no LSP Symposium, Vaasa, Israel, 1987.
- _____ "Sequence" and "order" in the development of L2 lexis: some evidence from lexical confusions. *Applied Linguistics* 11(3), 1990.
- _____ & D.D. SIM. Taking the easy way out, non-use and misuse of clues in EFL reading. *English Teaching Forum* 23(2), 1985:7-10.
- LEBART, L. & J.P. FENELON. *Statistique et informatique appliquées*. Paris: Dunod, 1973.
- LEE, J.F. Background knowledge & L2 reading. *The Modern Language Journal* 70(4), 1986: 350-354.
- LEECH, G. N. *Explorations in semantics and pragmatics*. Amsterdam: John Benjamins, 1980.

- _____. *Semantics*. Harmondsworth: Penguin, 1974.
- _____. *Semantics*. 2nd edition. Harmondsworth: Penguin, 1981.
- _____. *Principles of pragmatics*. London: Longman, 1983.
- LEFFA, V.J. *The role of comprehension monitoring skills and syntactic competence on reading comprehension in a foreign language*. Austin: Graduate School of the University of Texas. Tese de doutorado, 1984.
- LEVENSTON, E.A. Second language acquisition: issues and problems. *Interlanguage Studies Bulletin* 4, 1979:147-160.
- LEVINSON, S. *Pragmatics*. Cambridge: Cambridge University Press, 1983.
- LIU, N. & I.S.P. NATION. Factors affecting guessing vocabulary in context. *RELC Journal* 16(1), 1985:33-42.
- LORD, R. Learning vocabulary. *IRAL* 12(3), 1974:239-47.
- LORGE, I. & J. CHALL. Estimating the size of vocabularies of children and adults: an analysis of methodological issues. *Journal of Experimental Education* 32(2), 1963:147-157.
- LYONS, J. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press, 1968.
- _____. *Semantics*. 2 vols. London & New York: CUP, 1977.
- _____. *Language and linguistics: an introduction*. Cambridge: Cambridge University Press, 1981.
- MACIEL, A.M.B. Novos rumos para um antiga questão: proficiência em língua ou proficiência em leitura. *Trabalhos em Lingüística Aplicada* 13. Campinas: Universidade Estadual de Campinas, 1989:255-262.
- MACKEY, W.F. *Language Teaching Analysis*. Longman: London, 1965.
- MACNAMARA, J. Comparative studies of reading and problem solving in two languages. McGill University: Language Research Group, 1967.
- _____. The linguistic independence of bilinguals. *Journal of Verbal Learning and Verbal Behaviour* 6, 1967:729-36.
- MALACA CASTELEIRO, J. , A. MEIRA & J. PASCHOAL. *Nível Limiar*. Conselho da Europa/ICALP/Ministério da Educação, Lisboa, 1988.
- MAQUERA SOSA, B.R. *Estudo de tabelas de contingência através da análise de correspondência*. Campinas: Universidade Estadual de Campinas, IMECC. Dissertação de Mestrado, 1989.
- MARTON, W. Foreign vocabulary learning as a problem no. 1 of language teaching at the advanced level. *Interlanguage Studies Bulletin* 2, 1977:33-57.
- MCCARTHY, M.J. A new look at vocabulary in EFL. *Applied Linguistics* 5(1), 1984:12-22.

- McCLELLAND, J. & D. RUMELHART. An interactive activation model of the effect of context in perception. *Psychological Review* 88, 1981:375-407.
- MEARA, P.M. Vocabulary acquisition: a neglected aspect of language learning. *Language Teaching & Linguistics Abstracts* 13, 1980: 221-246.
- _____. Word associations in a foreign language: a report on the Birkbeck vocabulary project. *Nottingham Linguistic Circular* 11(2), 1982:29-38.
- _____. *Vocabulary in a second language* specialised bibliography no. 3). London: Centre for Information on Language Teaching and Research, 1983.
- _____. The study of lexis. In A. Davies, C. Criper & A. Howatt, eds., *Interlanguage*. Edinburgh: Edinburgh University Press, 1984:225-235.
- _____. *New approaches to testing vocabulary knowledge* (no prelo).
- _____. & G. JONES. Vocabulary size as a placement indicator. In P. Grunwell, ed., *Applied linguistics in society*. London: CILT, 1988.
- _____. & G. JONES. *The Eurocentres vocabulary size tests 10KA*. Zurich: Eurocentres, 1990.
- MELKA TEICHROEW, F.J. Receptive versus productive vocabulary: a survey. *Interlanguage Studies Bulletin* 6. Utrecht, 1982:5-33.
- MEY, J.L. *Whose language: a study in linguistics pragmatics*. Amsterdam: John Benjamins, 1985.
- MINSKY, M. A framework for representing knowledge. In P. Evinston, ed., *The psychology of computer vision*. New York: Winston, 1975.
- MODIANO, N. *Reading comprehension in the national language: a comparative study of bilinguals and all Spanish approaches to reading instruction in selected Indian schools in the highlands of Chiapas, Mexico*. New York University. Tese de doutorado, 1966.
- MOITA LOPES, L.P. *Discourse analysis and syllabus design: an approach to the teaching of reading*. London: University of London, Department of English to Speakers of Other Languages. Tese de doutorado, 1986.
- _____. Applied Linguistics in Brazil: Perspective. *The British Association of Applied Linguistics Newsletter* 39, 1991:27-30.
- _____. *Linguística Aplicada no Brasil: um olhar retrospectivo e prospectivo*. *Boletim Informativo Anpoll* 21. Brasília, Goiânia, 1994.
- MONDRIA, J-A. & WIT-DE BOER, M. The effects of contextual richness on the guessability and the retention of words in a foreign language. *Applied Linguistics* 12(3), 1991:249-267.
- MORAES, M.G. de D. *O saber e o poder do professor de línguas: algumas implicações para uma formação crítica*. Campinas: Universidade Estadual de Campinas. Dissertação de mestrado, 1990.
- MOURA, E.V. *Estratégias de aprendizagem de língua estrangeira entre alunos de diferentes*

níveis de rendimento e de proficiência. Assis: Universidade Estadual Paulista. Dissertação de mestrado, 1992.

MUNBY, J. *Communicative syllabus design*. Cambridge: Cambridge University Press, 1978.

NAGY, W.E. & R.C. ANDERSON. How many words are there in printed school English. *Reading Research Quarterly* 19(3), 1984:304-330.

_____ & P.A. HERMAN. Breadth and depth of vocabulary knowledge: implications for acquisition and instruction. In M.G. McKewon and M.E. Curtis, eds., *The nature of vocabulary acquisition*. Hillsdale: Lawrence Erlbaum, 1987.

NATION, I.S.P. Teaching and testing vocabulary. *Guidelines* 5(1), 1983:12-25.

_____ *Vocabulary list: words, affixes and stem*. New Zealand: English Language Institute, Wellington: Victoria University of Wellington, 1984.

_____ *Word Lists* (revised edition). Wellington: Victoria University English Language Centre, 1986.

_____ & J. COADY. Vocabulary and reading. In R. Carter & M. McCarthy. *Vocabulary and language teaching*. Longman: London, 1988:97-110.

NATTINGER, J.R. A lexical phrase grammar for ESL. *TESOL Quarterly* 14, 1980: 337-44.

NISHISATO, S. *Analysis of categorical data: dual scaling and its applications*. Toronto: University of Toronto Press, 1980.

OLIVEIRA, P.S.X. Alemão instrumental para estudantes de filosofia - avaliação da fase piloto - I e II semestres de 1989. Relatório de pesquisa inédito. Campinas: Universidade Estadual de Campinas.

_____ Cursos básicos de leitura em língua estrangeira: elucidando o papel da gramática, com o alemão por preferência. *Trabalhos em Lingüística Aplicada* 20. Campinas: Universidade Estadual de Campinas, 1992:69-89.

OLLER, J.W. The cloze test and ESL proficiency. *Language Learning* 21, 1971:183-196.

_____ *Language tests at school*. New York: Longman, 1979.

_____ & C. CONRAD. Cloze tests of second language proficiency and what they measure. *Language Learning* 23, 1973: 105-19.

_____ & K. PERKINS. *Research in language testing*. Rowley: Newbury House Publishers, 1969.

OLSHAVSKY, J.E. Reading as problem solving: an investigation of strategies. In *Reading Research Quarterly* 12, 1976-1977: 654-74.

ORLANDI, E.P. *Sobre a estruturação do discurso*. Campinas: Universidade Estadual de Campinas, 1981.

_____ *A linguagem e seu funcionamento: as formas do discurso*. São Paulo: Brasiliense,

1983.

- OXFORD, R. & D. CROOKALL. Research on language strategies: methods, findings, and instructional issues. *The Modern Language Journal* 73(4), 1989:405-420.
- PAES E BARROS, A. & R. ROJO. Convergência e divergência em leitura: reflexões sobre uma análise de resumos. *Trabalhos em Lingüística Aplicada* 4. Campinas: Universidade Estadual de Campinas, 1984:47-66
- PAIN, S. A capital cure for the Soviet Union's ecological ills. *New Scientist*, 1/September/1990.
- PANY, L. & J.R. JENKINS. Learning word meanings: a comparison of instructional procedures. *Learning Disability Quarterly* 1, 1978:21-32.
- _____, J.R. JENKINS & J.B. SCHRECK. Vocabulary instruction: effects on word knowledge and reading comprehension. *Learning Disability Quarterly* 5, 1982:202-215.
- PARRY, K.J. Too many words: learning the vocabulary of an academic subject. Comunicação apresentada na 22nd Annual Conference of TESOL Convention, Chicago, 1988.
- PATROCÍNIO, E.M.F. *Repensando o conceito de competência comunicativa no "aquecimento" da aula de português-língua estrangeira: uma perspectiva estratégica*. Campinas: Universidade Estadual de Campinas. Dissertação de mestrado, 1993.
- PAZ, O. *Los hijos del limo*. Barcelona: Seix Barral, 1974.
- PERFETTI, C.A. *Reading ability*. New York: Oxford University Press, 1985.
- _____. Cognitive and linguistic components of reading ability. In B. Foosman & A. Siegel, eds., *Acquisition of reading skill*. Hillsdale: Lawrence Erlbaum, 1986a:11-40.
- _____. Reading acquisition and beyond: decoding includes cognition. Em N. Stein, ed., *Literacy in American schools*. Chicago: University of Chicago Press, 1986b:41-61.
- PICKERING, M. Context-free and context-dependent vocabulary learning: an experiment. *System* 10(1), 1982:79-83.
- POLITZER, R.L. Errors of English speakers of German as perceived and evaluated by German natives. *The Modern Language Journal* 62, 1978:253-261.
- _____. & F. Politzer. *Teaching English: teaching English as a second language*. Lexington, MA, 1972.
- PRESSLEY, M. Children's use of the key word method to learn simple Spanish vocabulary words. *Journal of Educational Psychology*, 69 (5), 1977:465-72.
- RAMOS, R.G. *Estratégias usadas por falsos principiantes na leitura de textos acadêmicos em inglês*. São Paulo: Pontificia Universidade Católica. Dissertação de mestrado, 1988.
- RAUGH, M.R. & R.C. ATKINSON. A mnemonic method for the acquisition of a second-language vocabulary. *Psychology and Education Series Technical Reports*, 1974:224.

- READ, J. Towards a deeper assessment of vocabulary knowledge. Comunicação apresentada no 8th World Congress of Applied Linguistics, Sydney, Australia, 1987.
- REICHARDT, C. & T. COOK. Qualitative and quantitative methods in evaluation research. In T. Cook & C. Reichardt, eds., *Qualitative and quantitative methods in evaluation research*. Beverly Hills: Sage Publications, 1979:7-32.
- RHODES, L. Comprehension and predictability: an analysis of beginning reading materials. In J.C. Harste & R.F. Crey, eds., *New perspectives on comprehension*. Bloomington: Indiana University School of Education, 1979.
- RICHARDS, J. The role of vocabulary teaching. *TESOL Quarterly* 10(1), 1976.
- RICHARDSON, M. & G.F. KUDER. Making a rating scale that measures. *Personnel Journal* 12, 1933:36-40.
- RIVERS, W.M. *Communicating naturally in a second language*. New York: CUP, 1983.
- ROBERTS, J.T. Recent Developments in ELT. *Language Teaching* 15(2), 1982:94-110.
- ROBINSON, P.J. *ESP (English for specific purposes)*. Oxford: Pergamon Press, 1980.
- _____. Procedural knowledge and language learning. *Journal of Pragmatics* 13, 1989:523-546.
- RUDEL, R.B. Psycholinguistic implications for a system of communication model. In K. Goodman & J. Flemming, eds., *Psycholinguistics and the teaching of reading*. Newark, Delaware: International Reading Association, 1976.
- RUMELHART, D.E. Toward an interactive model of reading. In S. Dorniv, ed., *Attention and performance VI*. Hillsdale: Lawrence Erlbaum, 1977:573-603.
- _____. & A. ORTONY. The representation of knowledge in memory. In R.C. Anderson, R.J. Spiro & W.E. Montague, eds., *Schooling and the acquisition of knowledge*. Hillsdale: Lawrence Erlbaum, 1977:99-135.
- RUSSEL, D.H. & I.Q. SAADEH. Qualitative levels in children's vocabularies. *Journal of Educational Psychology* 53(4), 1962:170-174.
- SAMUELS, S.J. & M.L. KAMIL. Models of the reading process. In P. Carrell, J. Devine and D. Eskey, eds., *Interactive approaches to second language reading*. Cambridge: Cambridge University Press, 1988:22-36.
- SAVILLE-TROIKE, M. Reading and the audio-lingual method. *TESOL Quarterly* 7(4), 1973:395-405.
- SCARAMUCCI, M.V.R. O resumo e a compreensão em leitura em língua estrangeira. *Trabalhos em Lingüística Aplicada* 15. Campinas: Universidade Estadual de Campinas, 1990:65-86.
- _____. A construção do sentido na leitura em inglês como língua estrangeira: algumas implicações para o ensino, *Intercâmbio*. São Paulo: Pontifícia Universidade Católica, 1992:227-238.

- SCHANK, R.C. & R.P. ABELSON. *Scripts, plans, goals, and understanding: an inquiring into human knowledge structures*. Hillsdale: Lawrence Erlbaum, 1977.
- SHIFFRIN, R.M. & W. SCHNEIDER. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review* 84(2), 1977:127-190.
- SCHMITZ, J.R. O termo "estratégia": um conceito útil para a Linguística Aplicada? *Alfa* (no prelo).
- SCOTT, M. Vocabulary inference effects in reading English as a foreign language. Florianópolis: Universidade Federal de Santa Catarina, mimeo, 1984.
- _____. *Demystifying the jabberwocky: a research narrative*. Lancaster: University of Lancaster. Tese de doutorado, 1990.
- _____. What does inferring from context mean? In M.S.Z. de Paschoal and M.A.A. Celani, orgs., *Linguística aplicada: da aplicação da linguística à linguística transdisciplinar*. São Paulo: EDUC, Pontificia Universidade Católica, 1992:195-222.
- SEARLE, J.R. *Speech acts: an essay in the philosophy of language*. Cambridge: Cambridge University Press, 1969.
- SEIBERT, L.C. An experiment on the relative efficiency of studying French vocabulary in associated pairs versus studying French vocabulary in context. *Journal of Educational Psychology* 21, 1930:297-314.
- SHOHAMY, E. Predicting speaking proficiency from cloze tests: theoretical and practical considerations for tests substitution. *Applied Linguistics* 3(2), 1982:161-169.
- SILVA, E.T. *Elementos de pedagogia da leitura*. São Paulo: Martins Fontes, 1988.
- SINCLAIR, J.M., R. Daley & S. Jones. English lexical studies, Report no. 5060.
- _____. & A. RENOUF. A lexical syllabus for language learning. In R. Carter & M. McCarthy, eds., *Vocabulary and language teaching*. London: Longman, 1988:140-60.
- SINGER, M.H. & J. CROUSE. The relationship of context-use skills to reading: a case for an alternative logic. *Child Development* 52, 1981:1326-29.
- SOUZA, A.M.R. de. *Análise de correspondência*. São Paulo: Universidade de São Paulo, Instituto de Matemática e Estatística. Dissertação de Mestrado, 1982.
- SOUZA, N.M. de. *Análise de correspondência*. Porto Alegre: Universidade Federal do Rio Grande do Sul, Bacharelado em Estatística. Monografia, 1990.
- SOUZA E SILVA, M.C.P. A intertextualidade e os tempos verbais: fatores determinantes para a caracterização tipológica dos textos acadêmicos. *Trabalhos em Linguística Aplicada* 15. Campinas: Universidade Estadual de Campinas, 1990:87-104.
- SPADA, N. Relationships between instructional differences and learning outcomes: a process-product study of communicative language teaching. *Applied Linguistics* 8(2), 1987:137-155.

- SPEARITT, J. Identification of subskills of reading comprehension by maximum likelihood factor analysis. *Reading Research Quarterly* 8, 1972:92-111.
- SPERBER, D. & WILSON, D. Irony and the use-mention distinction. In P. Cole, org., *Syntax and Semantics 9: Pragmatics*. Academic Press, 1981.
- STAHL, S. & FAIRBANKS, M.M. The effects of vocabulary instruction: a model-based meta-analysis. *Review of Educational Research* 56(1), 1986:72-110.
- STANOVICH, K.E. Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly* 16, 1980:32-71.
- STENBERG, R.J. & J.S. POWELL. Comprehending verbal comprehension. *American Psychologist* 38, 1983:878-893.
- STEVIK, E.W. Research on what? Some terminology. *The Modern Language Journal* 74(2), 1990:143-152.
- STROTHER, J. & J. ULLIN. Does syntatic rewriting affect English for science and technology text comprehension? In J. Devine, P. Carrell and D. Eskey, eds., *Research in reading as a foreign language*, 1987:89-100.
- TARVIN, W.L. & A.Y. AL-ARISHI. Rethinking communicative language teaching: reflection and the EFL classroom. *TESOL Quarterly* 25(1), 1991: 9-25.
- TAVARES, K. *O uso da introspecção: da técnica de pesquisa para o ensino de leitura*. Rio de Janeiro: Universidade Federal do Rio de Janeiro. Dissertação de mestrado, 1993.
- TAYLOR, W.L. Cloze procedure: a new tool for measuring readability. *Journalism Quarterly* 33, 1953:42-8.
- TAYLOR, I. & M. TAYLOR. *The psychology of reading*. New York: Academic Press, 1983.
- TERZI, S.B. A extensão do texto como determinante de mudanças de postura frente à leitura: uma experiência. *Trabalhos em Lingüística Aplicada* 3. Campinas: Universidade Estadual de Campinas, 1984:69-79.
- THORNDIKE, E.L. & I. LORGE. *The teacher's word book of 30,000 words*. Columbia University, Teachers College, 1944.
- TIERNEY, R.J. & J. LAZANSKY. The rights and responsibilities of readers and writers: a contractual agreement. *Reading Education Report* 15. Illinois: University of Illinois at Urbana-Campaign, Center for the Study of Reading, 1980.
- TUINMAN, J.J. & M.E. BRADY. How does vocabulary account for variance on reading comprehension tests: a preliminary instructional analysis. *Twenty-third yearbook of the national reading conference*. Clemson, SC: National Reading Conference, 1974.
- TWADELL, F. Vocabulary expansion in the TESOL classroom. In K. Croft, ed., *Readings in English as a second language*. 2nd edition. Cambridge, MA.: Winthrop, 1980.
- TWOMEY, A *bibliography of research carried out in the field of vocabulary acquisition in a second language*. London: Birbeck College. Dissertação de mestrado, 1979.

- ULIJN, J.M. Conceptualisation in second language reading. Comunicação apresentada no 5th International Congress of Applied Linguistics, Montreal, 1978.
- _____ & G.A.M. KEMPEN. The role of the first language in second reading comprehension: some experimental evidence. In G. Nickel, ed., *Proceedings of the 4th International Congress of Applied Linguistics 1*, Stuttgart: Hochschulverlag, 1976.
- VAN DIJK, T. & W. KINTSH. *Strategies of discourse comprehension*. New York: Academic, 1983.
- VAN EK, J.A. *The threshold level*. London: Longman, 1976.
- VAN PARREREN, C.F. & M. SCHOUTEN-VAN PARREREN. Contextual guessing: a trainable reader strategy. *System* 9(3), 1981:235-241.
- VILLALOBOS-AGUAYO, M.T. *Análise de correspondência e modelos log-lineares: um enfoque para a análise exploratória de dados categóricos*. Campinas: Universidade Estadual de Campinas, IMECC. Dissertação de Mestrado, 1993.
- VYGOTSKY, L.S. *Mind in society*. Cambridge: M.I.T University Press, 1978.
- XUE G. & I.S.P. NATION. A university word list. *Language Learning and Communication* 3(2), 1984:215-229.
- YAMAMOTO, M.K. *Português Instrumental: o processo de compreensão da metáfora no texto didático*. São Paulo: Pontifícia Universidade Católica. Dissertação de Mestrado, 1991.
- WALLACE, M. *Teaching vocabulary*. London: Heinemann, 1982.
- WEST, M. *A general service list of English words*. London: Longman, Green & Company, 1953.
- WIDDOWSON, H.G. The process and purpose of reading. In H. Widdowson, ed., *Explorations in Applied Linguistics*. New York: Cambridge University Press, 1979:171-183.
- _____ *O ensino de línguas para a comunicação*. Pontes: Campinas, 1991.
- _____ *Learning purpose and language use*. Oxford: Oxford University Press, 1983.
- _____ Knowledge of language and ability for use. *Applied Linguistics* 10(2), 1989:128-137.
- WILKINS, D.A. *Linguistics and language teaching*. London: Edward Arnold, 1972.
- ZILBERMAN, R. (org.) *Leitura em crise na escola: as alternativas do professor*. Porto Alegre: Mercado Aberto, 1982.
- ZIMMERMAN, J., P. BRODER, J. SHAUGHNESSY & B. UNDERWOOD. A recognition test of vocabulary using signal-detection measures, and some correlates of word and non-word recognition. *Intelligence* 1, 1977: 5-31.