

METODOLOGIA

A atividade de compreensão textual foi investigada a partir das questões de múltipla escolha presentes nas cinco provas de nivelamento em leitura a fim de analisarmos: o tipo de questão de múltipla escolha, a relação entre o nível de dificuldade de cada questão com a inteligibilidade dos textos, e a habilidade/competência de compreensão.

Sabemos que já existem vários estudos sobre a natureza das provas do PISA no Brasil e do SAEB, a destacar Bonamino et al. (2002), e o trabalho de Lopes-Rossi e Orlando de Paula (2012) que traça um estudo comparativo entre as habilidades testadas no PISA e no ENEM. Concernente às provas de nivelamento em leitura em português da PUC-Rio, contudo, ainda não há uma investigação acadêmica, razão suficiente para a escolha do corpus desta pesquisa ser constituído apenas por estas provas, atribuindo ao presente trabalho um ineditismo pertinente, esperando que sua contribuição acadêmica possa ser grande para as pesquisas sobre letramento na universidade.

A pesquisa foi realizada, grosso modo, em duas etapas, nas quais foram levantados dados estatísticos e dados qualitativos das provas realizadas pelos leitores desta pesquisa.

Neste capítulo são apresentadas as etapas de tratamento de dados que alicerçaram as análises quantitativas e qualitativas. Para isso, contei com o auxílio da ferramenta Coh-Metrix-Port, a qual forneceu o índice Flesch de inteligibilidade dos textos acadêmicos das PNLs, de um lado; e as análises manuais das estruturas sintáticas que têm sido apontadas pela literatura como de maior custo de processamento cognitivo, de outro.

A descrição da macroestrutura de cada texto acadêmico é o tópico do próximo item.

8.1

Layout das provas de nivelamento em leitura – a proposta pedagógica da universidade

A PNL de 2009.2⁶⁰ contém três textos. O primeiro foi extraído e adaptado do livro *Do que é feito o pensamento* do autor Steven Pinker, 2008. O segundo foi adaptado do artigo de Bastien Bosa, publicado na revista eletrônica de Estudos de Antropologia Social MANA, em 2009. O terceiro trabalha com a modalidade infográfica, na qual uma tabela é apresentada com dados do IBOPE.

A PNL 2010.1 tem 03 textos. O primeiro também foi extraído do livro *Uma introdução à história do design* do autor Rafael Cardoso, 2008. O gráfico no texto 2 disponibiliza os índices de vendas no comércio de rua e em shopping centers. O texto 3 - *A parte que nos cabe: consumo sustentável?* – é um fragmento do capítulo de mesmo nome do autor Fábio Feldmann, publicado no livro *Meio Ambiente no século 21*, 2003.

Em ambas as provas acima mencionadas, os textos foram intercalados com questões de múltipla escolha. No entanto, para a PNL 2010.2, a banca examinadora inovou. Os textos 1 e 2 são apresentados de forma sequenciada, ou seja, as questões de múltipla escolha aparecem ao final do segundo texto. A prova finaliza com o texto 3 e suas respectivas questões.

Na PNL 2010.2, o texto 1 é um extrato do livro de Rosiska Darcy de Oliveira - *Elogio da diferença: o feminino emergente* -, publicado em 1999. O segundo é um artigo publicado na Revista CULT, em 2010, no qual Franklin Leopoldo e Silva argumenta sobre Ética e situações-limite. Após as questões de múltipla escolha, temos o texto 3 que introduz uma tabela extraída de uma pesquisa realizada pelo sociólogo Alberto Carlos Almeida e posteriormente publicada em seu livro *“A Cabeça do Brasileiro”*, em 2007. A tabela traz os resultados da Pesquisa Social Brasileira (PESB) cujo objetivo era investigar o apoio social do brasileiro ao chamado “jeitinho” brasileiro. Após este terceiro texto, outras questões de múltipla escolha se sucedem.

⁶⁰ 1 refere-se ao primeiro semestre do ano mencionado; .2, ao segundo semestre.

A PNL 2011.1 igualmente traz três textos, sendo o texto 1 adaptado do capítulo - Com fé e orgulho - de Marilena Chauí no livro Brasil: Mito fundador e sociedade autoritária, 2000. O segundo texto também é uma adaptação do capítulo O motivo edênico no imaginário social brasileiro de José Murilo de Carvalho, publicado no livro CIDADANIA, justiça e violência, lançado em 1999. Este segundo texto explora duas tabelas que tratam de informação verbal e numérica.

Outra inovação se observa na PNL 2011.1, pois as questões são agora apresentadas ao final do terceiro texto - Jovens têm orgulho de ser brasileiros e acreditam que país é importante no mundo. O texto foi adaptado do original publicado no site da UOL.

Por fim, a PNL 2011.2 apresenta a mesma organização de apresentação das questões de múltipla escolha ao final do terceiro texto, porém nenhuma outra modalidade foi explorada além da discursiva. O primeiro texto é uma adaptação do capítulo - O conhecimento científico - publicado em 1886 no livro Filosofando: uma introdução à filosofia, das autoras Maria Helena Martins e Maria Lúcia de Souza Aranha. O texto 2 é uma definição de dicionário de filosofia. A adaptação do verbete retirado do Dicionário Básico de Filosofia de 1991: Senso comum dos autores Danilo Marcondes e Hilton Japiassu. Finalmente, o terceiro texto - Leonardo da Vinci, neurocientista – é uma Adaptação de texto da revista Mente & Cérebro do autor Jonathan Pevsner, em 2006.

Pela descrição dos textos das PNLs, podemos notar a proposta pedagógica da banca examinadora não somente em avaliar a compreensão leitora do leitor-universitário, mas igualmente em lhe prover condições de se conscientizar da imperiosa necessidade de dominar os parâmetros discursivos necessários à comunicação oral e escrita naquela comunidade acadêmica.

Na tabela a seguir, podemos observar que os itens⁶¹ das provas são elaborados a partir de uma matriz de especificação de habilidades e competências cognitivas a serem avaliadas, correlacionando-as a um determinado nível de proficiência. Toma-se por pressuposto que a probabilidade de resposta a um item é reflexo da interação entre a proficiência leitora do graduando, do tipo de

⁶¹ Nesta pesquisa, os termos item e questão são utilizados com o mesmo sentido.

habilidade testado no item, tanto quanto o nível de dificuldade que o tipo de item impõe.

Ao usarmos a classificação das questões, conforme as operações cognitivas e estratégias metacognitivas, é importante lembrar que uma questão de múltipla escolha dificilmente se encaixa em apenas uma das classificações. As classificações aqui propostas têm por base teórica os níveis de processamento inventariados por Perfetti: nível lexical, nível sintático, nível textual; e a outra fonte de categorização das questões foi a utilizada pelo Pisa (recuperação de informação explícita, interpretação e integração, reflexão e avaliação). Há ainda uma terceira fonte de categorização gerada a partir da própria especificidade das PNLs. Neste sentido, foram criadas as categorias identificação ou inferência de significado de vocabulário, e processamento sintático-semântico subcategorizado em três domínios, tal como podemos ver na tabela abaixo.

Tabela 7: Habilidades e competências testadas em cada item das PNLs

PNL	Identificação ou Inferência de significado de vocabulário	Identificação e recuperação de informações explícitas	Processamento sintático-semântico			Interpretação e integração (síntese) do assunto principal (sentido global)	Identificação e interpretação e/ou integração de ideias no nível local	Reflexão e avaliação do texto
			Domínio					
			Identificação de relação lógico-semântica interfrástica (coesão sequencial)	Paráfrase	Estabelecimento de relação correferencial			
2009.2	Q3	Q2, Q5, Q7, Q8	Q4			Q1	Q6, Q9	Q10
2010.1		Q1	Q6, Q7		Q3		Q4, Q5	Q2, Q8, Q9, Q10
2010.2	Q3	Q2				Q4, Q6	Q1, Q7, Q8	Q5, Q9, Q10
2011.1		Q2, Q6		Q7	Q5	Q1, Q4,	Q3, Q8	Q9, Q10
2011.2	Q9	Q7			Q4		Q1, Q2, Q6	Q3, Q5, Q8, Q10

H1 – Identificação e inferência de vocabulário - reconhecer ou inferir significado de vocábulo não muito usual na linguagem coloquial, através de seu emprego na frase.

H2 - Identificação e recuperação de informações - localizar informação relevante e explícita no texto.

H3 - Processamento sintático-semântico - reconhecer ideias apresentadas por diferentes expressões e recursos linguísticos (paráfrase); estabelecer relações entre partes de um texto, identificando repetições ou substituições que contribuem para a continuidade de um texto (questões de retomada anafórica pronominal); estabelecer relação de causa e consequência entre partes e elementos do texto; estabelecer relações lógico-discursivas presentes no texto, marcadas por conjunções, advérbios .

H4 - Interpretação e integração (síntese) do assunto principal (sentido global) - interpretar informações localizadas em diferentes partes do texto, e as integrar em uma síntese que dê sentido global ao texto (de parágrafo para parágrafo, de parágrafo para texto); diferenciar as partes principais das secundárias de um texto.

H5 - Identificação e interpretação e/ou integração de ideias no nível local (no parágrafo e entre parágrafos)

H6 - Reflexão e avaliação do texto (interagindo com conhecimento enciclopédico) - o leitor faz o uso de informações extratextuais, ou seja, seu conhecimento enciclopédico e suas experiências pessoais para avaliar o texto ou completar-lhe lacunas; reconhecer diferentes formas de tratar uma informação na comparação de textos que tratam do mesmo tema, em função das condições em que ele foi produzido e daquelas em que será recebido (questões de intertextualidade); utilizar conhecimentos prévios para imaginar um exemplo que corresponde a uma categoria descrita em um texto; diferenciar tipologia e/ou gênero textual, tanto quanto identificar os recursos linguísticos e suas funções em dado texto.

É assim que se justifica a utilização das PNLs para a avaliação das competências leitoras dos universitários na PUC-Rio, em um recorte que põe em dinâmica interativa as habilidades e competências leitoras para a resolução de itens cujas propriedades, no entanto, demandam a utilização mais expressiva de uma dada habilidade específica. Já vimos que mais de uma habilidade concorre para a resolução de itens de múltipla escolha, todavia elege-se uma habilidade específica para cada item, a qual passa a ser a habilidade âncora.

A escala de proficiência é fundamental para que o universitário tenha clareza das habilidades leitoras que já desenvolveu. No entanto, para a PUC-Rio vale ressaltar também a(s) habilidade(s) que o aluno ainda precisa solidificar ou desenvolver, donde o feedback dado ao leitor-universitário ser sempre pautado em cada item que ele errou. Para melhor esclarecer, um leitor que tenha logrado o nível 3 de proficiência, por exemplo, uma nota 6,5 na PNL. Este leitor será considerado proficiente para a comunidade PUC-Rio, no entanto ele tomará ciência de suas habilidades que ainda carecem de aprimoramento. O quadro abaixo é um exemplo do retorno que a PUC-Rio dá ao graduando.

MENSAGEM PARA OS ALUNOS CUJA PONTUAÇÃO NA PROVA CORRESPONDE AO NÍVEL III (6 A 6,9)
<p>Na prova de COMPREENSÃO DE TEXTO, você obteve nota 6,5 (Nível III). Está, portanto, dispensado de realizar novo Teste de Nivelamento em Leitura. Com base na análise de suas respostas, percebe-se, no entanto, que você ainda precisa desenvolver as seguintes habilidades:</p> <p>Questão 4. Habilidade testada: Identificar relação lógico-semântica entre idéias expressas em orações interligadas.</p> <p>Questão 9. Habilidade testada: Identificar informações em gráficos e relacioná-las com conhecimento prévio sobre notações e termos matemático-estatísticos.</p>

8.2

Dos tipos de questões de múltipla escolha nas PNLs

Como asseverado anteriormente, na descrição da resolução de cada item o enfoque reside na descrição dos atos cognitivos que o leitor-universitário desenvolve enquanto resolve e comunica suas soluções. Desta forma, foi possível fazer uma tipologia dos itens das PNLs, tendo por base o Guia de Elaboração e Revisão de Questões e Itens de Múltipla Escolha⁶², mas também considerando as especificidades das questões das PNLs.

Evitou-se excessivo uso de categorização dada a dificuldade em se trabalhar, dentro do escopo de uma dissertação de mestrado, com inúmeras categorizações e subcategorizações. O caráter sucinto das categorizações nos traz, de um lado, maior objetividade às análises; de outro, ele se faz acompanhar de uma carência de uma análise mais minuciosa dos dados. Entendemos, contudo, que, ainda em caráter preliminar, as análises aqui partilhadas podem contribuir para fomentar e subsidiar o debate sobre letramento na universidade e inteligibilidade textual. A tipologia das questões das PNLs abrange 4 categorias e 2 subcategorias, a saber:

Questão de resposta única: enuncia o problema ou a situação problema na forma de pergunta e apresenta as alternativas de resposta.

- Nível I – opções com uma única palavra ou com poucas sentenças;
- Nível II – opções com proposições mais longas (3 ou mais orações)

Questão de afirmação incompleta: apresenta o enunciado do problema ou situação problema como uma afirmação a ser completada por uma das alternativas.

Questão de resposta múltipla: apresenta uma situação contextualizada com afirmativas pertinentes a ela. A seguir, enuncia o problema ou situação problema na forma de pergunta ou afirmação incompleta e apresenta uma chave de resposta.

Questão de foco negativo: apresenta várias respostas corretas e apenas uma incorreta, que é a solicitada.

⁶² Ver, para tanto, <http://cenfopciencias.files.wordpress.com/2011/10/guia-de-elaboracao-de-itens.pdf>

A distribuição dos tipos de itens ao longo das PNLs pode ser observada na tabela abaixo.

Tabela 8: Tipologia dos itens

Item / PNL	1	2	3	4	5	6	7	8	9	10
2009.2	Resposta Única - nível II	Resposta Única - nível II	Resposta Única - nível I	Resposta Única - nível I	Resposta múltipla	Foco negativo	Afirmação incompleta	Afirmação incompleta	Foco negativo	Resposta Única - nível II
2010.1	Resposta Única - nível II	Resposta Única - nível II	Resposta Única - nível I	Resposta Única - nível I	Resposta Única - nível I	Resposta Única - nível I	Resposta Única - nível I	Resposta Única - nível I	Resposta Única - nível I	Resposta Única - nível II
2010.2	Resposta múltipla	Resposta Única - nível I	Resposta Única - nível I	Resposta múltipla	Resposta Única - nível I	Afirmação incompleta	Afirmação incompleta	Resposta múltipla	Resposta Múltipla	Foco negativo
211.1	Resposta Única - nível I	Resposta múltipla	Resposta múltipla	Foco negativo	Resposta Única - nível I	Resposta Única - nível I	Resposta Única - nível II	Afirmação incompleta	Afirmação incompleta	Foco negativo
2011.2	Resposta múltipla	Resposta Única - nível I	Foco negativo	Resposta Única - nível I	Resposta Única - nível I	Foco negativo	Resposta Única - nível I	Resposta Única - nível I	Resposta Única - nível II	Resposta Única - nível I

A tipologia foi importante para cotejar o tipo de item que mais demandou cognitivamente dos graduandos e a respectiva competência leitora avaliada na questão, o que será discutido mais adiante. Antes, contudo, vamos entender a análise computacional do Coh-Metrix feita para cada texto das provas, assunto do próximo item.

8.3

Da ferramenta computacional Coh-Metrix-Port

A ferramenta computacional Coh-Metrix-Port (Almeida & Aluísio, 2009) foi desenvolvida para o projeto PorSimples⁶³ (Simplificação Textual do Português para Inclusão e Acessibilidade Digital), todavia ela é disponibilizada gratuitamente para os mais variados fins. O Coh-Metrix-Port foi adaptado para o português brasileiro (PB) por pesquisadores da USP de São Carlos, os quais tomaram por base as métricas disponibilizadas pela ferramenta Coh-Metrix da

⁶³ <http://caravelas.icmc.usp.br:3000/>

Universidade de Memphis. Importante frisar que, até o momento, mais de 500 métricas estão disponíveis em uma versão restrita do Coh-Metrix. Dessas 500 métricas, o Departamento de Psicologia da Universidade de Memphis disponibiliza apenas 60 na versão gratuita on-line no site do projeto⁶⁴. A ferramenta Coh-Metrix foi desenvolvida para a análise linguístico-computacional de textos em língua inglesa. Ressalte-se ainda que até agora apenas 35 das 60 métricas foram adaptadas para o português brasileiro, em três níveis, como se observa abaixo.

Tabela 9: Índices do Coh-Metrix-Port

ÍNDICES LEXICAIS	índice Flesch; número de palavras
ÍNDICES SINTÁTICOS	palavras antes de verbos principais; tipo/token; sintagmas nominais; conectivos (causais, aditivos, temporais, lógicos); operadores lógicos
ÍNDICES SEMÂNTICOS	anáforas; sobreposição de argumentos adjacentes; sobreposição de argumentos

Ainda na primeira etapa desta pesquisa, os índices Flesch dos textos das PNLs provieram da ferramenta computacional Coh-Metrix-Port com o fim de se investigar a inteligibilidade dos mesmos. Após se submeter o texto para análise, o Coh-Metrix-Port disponibiliza métricas de diversos aspectos que vão desde o número de palavras até às análises morfossintáticas dos aspectos que garantem a inteligibilidade do texto, como, por exemplo, índice de verbos, índice de conectivos, dentre outros.

O Coh-Metrix-Port extrai automaticamente diversos atributos dos textos baseando-se na anotação sintática produzida pelo analisador sintático (parser) de Charniak (Scarton & Aluisio, 2010, p.4). A análise sintática não é total, mas somente parcial (identificação de sintagmas). Este parser identifica uma sentença pelo ponto final (.) e o parágrafo pela quebra de linhas; a partir daí o analisador obtém o número total de sentenças no texto.

Quando há quebra de linha, o parser conta o parágrafo, donde ser possível o número de sentenças por parágrafos. Este é um aspecto de suma importância, ou seja, ao se copiar o texto e colá-lo no espaço indicado pelo Coh-Metrix-Port, deve-se ter a preocupação em organizar seus parágrafos tal como os mesmos se apresentam no texto original.

⁶⁴ Atualmente a versão é 3.0. A adaptação da ferramenta para o português foi feita a partir da versão 2.0

Note-se, no entanto, que o sistema Coh-Metrix-Port não aprecia a presença de voz passiva, também não associa funções sociais ou confere pesos diferenciados entre tipologia textual e gênero textual. Em que pese a diferença entre tipo textual e gênero textual, o Coh-Metrix-Port atem-se à natureza linguística da composição do texto, ou seja, à sua tipologia: aspectos lexicais, sintáticos, tempos verbais, relações lógicas, número de palavras por sentença, palavras frequentes, etc. Isto equivale a dizer que os aspectos sociohistóricos que dão existência e manutenção aos gêneros textuais não são apreciados em suas características sócio-comunicativas nem em suas propriedades funcionais. Scarton & Aluísio (2010, p.46), desta forma, conferem relatividade aos dados da ferramenta web. As autoras sustentam, citando McNamara, que

As fórmulas de inteligibilidade superficiais não conseguem capturar a coesão e dificuldade de um texto nem avaliar mais profundamente as razões e correlações de fatores que tornam um texto difícil de ser entendido.

As análises computacionais não medem outros fatores relacionados à inteligibilidade, tais como dificuldade conceitual, excesso de ideias novas, organização do material (em colunas, estilo livro, multimodal), conteúdo ou características do usuário, etc. As estratégias de compreensão, de construção de sentidos locais e globais, enfim, as formas de interação com o texto são dinâmicas e flexíveis e, portanto, não repousam somente no conhecimento linguístico do leitor, mas também em outras plataformas de conhecimento e de fatores. Bohn (1990, p.2) amplia a relatividade dos dados obtidos por fórmulas computacionais.

The concept that text complexity is a physical feature or set of features which can objectively be taken from the text and plotted into an equation stems from the structural point of view of language description. This is a naïve interpretation of textual structure and a naïve way to look at the psychological and neurological baggage of the reader and the reading process itself. Under this perspective formulas contradict current knowledge about the reading process and they don't include in the measures important textual and personal variables that influence the understanding of a text⁶⁵.

⁶⁵ A ideia de que a complexidade do texto é um aspecto físico ou um conjunto de aspectos que podem ser objetivamente retirados do texto e lançados em uma equação decorre do ponto de vista estrutural da descrição de linguagem. Este é uma interpretação ingênua da estrutura textual e uma forma ingênua de olhar para a bagagem psicológica e neurológica do leitor e o próprio processo de leitura. Sob esta perspectiva, fórmulas contradizem o conhecimento atual sobre o processo leitura e elas não incluem nas medidas importantes variáveis textuais e pessoais que influenciam a compreensão de um texto. (Tradução da Autora)

De fato, os parâmetros disponibilizados pelo Coh-Metrix-Port não consideram os conhecimentos sócio-culturais que o leitor tem em sua memória, ou seja, seu conhecimento enciclopédico ou conhecimento de mundo, tanto quanto não computa características psicológicas do leitor que igualmente influenciam na compreensão do texto, como por exemplo, o seu nível da representação dos conhecimentos e o seu maior ou menor grau de facilidade para acessar informações em sua memória de longo prazo. Por outro lado, a análise linguístico-computacional tem atraído não somente críticas como as de McNamara e Bohn acima sinalizadas, mas igualmente elogios. Os elogios dirigem-se para o objetivo central dessas ferramentas, qual seja, o de identificar que fatores tornam um texto de fácil compreensão, possibilitando a re-escritura de textos complexos em um nível de inteligibilidade que alcance o leitor médio e o público infantil.

Assim, a despeito da existência de alguns pontos negativos acima considerados, nesta pesquisa fizemos uso da métrica índice Flesch para avaliar o nível de dificuldade apresentado por cada texto das PNLs em conjugação com outros fatores considerados a partir da análise sintática manual desses textos. Deste modo, a análise manual quantitativa das estruturas sintáticas tidas na literatura como complexas também foi indispensável ao aprofundamento da discussão acerca dos fatores que podem ter afetado o desempenho dos graduandos nas PNLs.

Para os textos em português, o índice Flesch é obtido a partir da fórmula: $206.835 - (1.015 \times \text{ASL}) - (84.6 \times \text{ASW})$. Onde ASL é o tamanho médio de sentenças, ou seja, o número de palavras dividido pelo número de sentenças; e ASW é número médio de sílabas por palavra, isto é, o número de sílabas dividido pelo número de palavras.

Os pesquisadores adaptadores da ferramenta web estabeleceram quatro faixas de classificação das dificuldades de leitura para a língua portuguesa:

- **Textos muito fáceis** (índice Flesch entre 75 - 100), que seriam adequados para leitores com nível de escolaridade até a quarta série do ensino fundamental;
- **Textos fáceis** (índice Flesch entre 50 - 75), que seriam adequados a alunos com escolaridade até a oitava série do ensino fundamental;
- **Textos difíceis** (índice Flesch entre 25 - 50), que seriam adequados para alunos cursando o ensino médio ou universitário e;
- **Textos muito difíceis** (índice Flesch entre 0 - 25), que em geral seriam adequados apenas para áreas acadêmicas específicas. (Scarton & Aluísio, 2010, p. 3)

Como visto acima, o índice Flesch exprime o grau de inteligibilidade de um texto escrito, estabelecendo uma correlação entre tamanhos médios de palavras e extensão de sentenças, e a facilidade de leitura. Segundo Liberato & Fulgêncio (op. cit., p. 19), ``o cérebro requer tempo para tomar suas decisões e interpretar o que foi visto``, donde sentenças longas podem representar um fator de acréscimo de dificuldade no processamento por, por exemplo, exigir que o leitor decida onde e como proceder aos fatiamentos dos constituintes.

No próximo item, alguns dados quantitativos serão apresentados. Eles se relacionam com as ocorrências de estruturas sintáticas complexas nos textos das PNLs. Para se ter uma base de análise, o número de períodos existentes em cada texto é disponibilizado. O Coh-Matrix_Port não conta orações, mas períodos. É a contagem automática do Coh-Matrix-Port que está sendo considerada aqui. Com isso, pode-se ter a razão entre o número de estruturas complexas e número de períodos em cada texto.

8.4

Dos dados quantitativos a partir de análises manuais

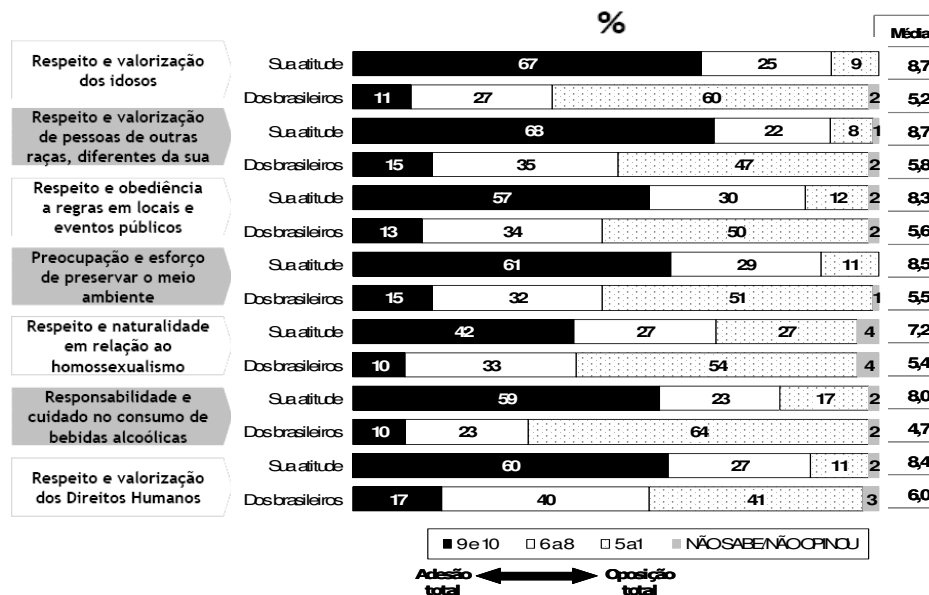
Ainda na primeira etapa de tratamento dos dados, uma análise manual nos trouxe dados quantitativos das PNLs no que tange aos textos e seus respectivos níveis de complexidade sintático-semântica. As estruturas sintáticas consideradas de maior complexidade linguística foram contabilizadas em números absolutos na tabela a seguir.

Tabela 10: Estruturas linguísticas complexas - Ocorrências em números absolutos

PNL	Texto	Índice Flesch	Passivas Analíticas	Estruturas Intercaladas (encaixe central)	Relativas de Objeto Direto	Relações Correferenciais pronominais
2009.2	1	57.54 (26 períodos)	8	6	-	5
	2	17.05 (22 períodos)	14	4	-	11
	3	tabela	-	-	-	-
	1	27.02 (23 períodos)	13	2	-	3
2010.1	2	gráfico	-	-	-	-
	3	34.54 (18 períodos)	5	4	-	4
	1	55.61 (24 períodos)	10	2	1	16
2010.2	2	33.99 (42 períodos)	20	14	6	18
	3	tabela	-	-	-	-
	1	39.18 (51 períodos)	17	11	2	17
2011.1	2	52.91 (53 períodos)	15	5	1	6
	3	57.43 (12 períodos)	5	-	1	1
	1	42.40 (44 períodos)	17	1	2	15
2011.2	2	23.44 (3 períodos)	3	-	1	-
	3	35.69 (29 períodos)	13	3	2	12

Destaque-se ainda que o Coh-Matrix-Port não trabalha com estatísticas multifatoriais, como, por exemplo, a tabela trazida no texto 3 da PNL 2009.2. Vê-se na figura abaixo que o texto é apresentado no formato planilha cuja terceira coluna revela o percentual atribuído à atitude pessoal e a dos brasileiros em geral em relação a certos valores. Neste caso, o leitor encaixou sua resposta em uma das 4 possibilidades, variando de adesão total até a oposição total.

Atitude pessoal e dos brasileiros



Essa análise computacional não é possível de ser obtida pelo Coh-Metrix-Port, razão pela qual não foi obtido o índice Flesch do texto 3, porém foram criados critérios de categorização de inteligibilidade para as tabelas e gráficos presentes nas PNLs. O nível de dificuldade das tabelas e gráficos foi analisado a partir do número de entradas (variáveis), tendo-se atribuído as seguintes categorias:

- Gráfico ou tabela com duas entradas – fácil/muito fácil
- Gráfico ou tabela com três ou quatro entradas – difícil
- Gráfico ou tabela – acima de quatro entradas – muito difícil

Tabela 11: Análise de gráficos e tabelas das PNLs

Análise de gráfico / tabela	2009.2 - tabela (texto 3)	2010.1 - gráfico (texto 2)	2010.2 - tabela (texto 3)
	Muito difícil	Fácil	Difícil

Como visto anteriormente, em todas as PNLs há um total de 10 questões de múltipla escolha que testam as habilidades de compreensão leitora. As notas obtidas pelos graduandos são lançadas em planilhas Excel pela PUC-Rio.

A partir das planilhas da PUC, foi feita uma planilha com percentuais de desempenho dos universitários em cada questão. A meta era obter o percentual de

acerto em cada item de forma a se registrarem aquelas que apresentaram um maior nível de dificuldade para os alunos, considerando-se as de maior grau de dificuldade as questões com porcentagem abaixo de 50% de acerto.

Registradas as questões mais difíceis para os graduandos, o objetivo agora era analisá-las correlacionando inteligibilidade do texto, tipologia do item, habilidade / competência avaliada no item, e estruturas sintáticas complexas presentes nos textos. Com isso, punha-se em relação dinâmica a inteligibilidade dos textos e o nível de compreensão leitora dos graduandos através dessas questões. Mencione-se ainda que há itens que podem ser classificados como híbridos porque envolvem dois ou três textos da prova. Em caso de itens híbridos, as questões foram definidas pelo peso maior que continham em um dos textos. Por exemplo, se a questão de intertextualidade trabalhava em cima de um texto de índice Flesch fácil e outro muito difícil, a questão foi categorizada como muito difícil. Da mesma forma, se a questão trabalhava com dados de uma tabela categorizada como fácil e um texto de índice Flesch difícil, o item foi tido como difícil.

No item 8.5 a seguir, algumas considerações sobre as questões mais difíceis são partilhadas com você, leitor(a). Nenhuma análise aqui tem caráter conclusivo nem de súmula vinculante que o(a) obrigue a aceitá-las para si. Postas as minhas análises e interpretações, as mesmas estão abertas à discussão, objetivo primordial desta pesquisa.

8.5

Considerações sobre as questões mais difíceis nas PNLs

Pullin et.al. (2009) consideram legítimo pressupor que os modos de ler produções técnicas, especializadas ou científicas passam a ser desenvolvidos ao longo da vivência acadêmica no curso de origem do graduando. Ora, do ponto de vista da vivência dos graduandos da PUC-Rio, resta-nos fazer uma pergunta crucial a Pullin, diante de suas considerações a respeito da competência leitora: por que há graduandos e graduados em todo o país (ENADE, INAF) que ainda

são avaliados como leitores pouco proficientes? Há resultados nesta pesquisa (e em outras aqui citadas) que nos concitam a uma postura cautelosa para afirmar que o desenvolvimento de novos modos de ler adequados ao tratamento dispensado às produções textuais acadêmicas são decorrentes do próprio ofício dos alunos no Ensino Superior, o que nos mantém sob a ótica investigativa acerca dos fatores que influenciam negativamente as habilidades leitoras de universitários no Brasil.

Da população investigada na presente pesquisa, por exemplo, há universitários que realizaram a prova de nivelamento por dois e até três anos consecutivos; alguns lograram sucesso, todavia até o final de 2011 existiam casos de leitores que não atingiram o nível III (6,0 a 6,9) exigido pela PUC-Rio. Após 2-3 anos desenvolvendo o ofício de estudante do Ensino Superior, os graduandos que chegaram a 2011.2 com 3 e / ou 4 reprovações nas PNLs acentuaram em mim a necessidade de buscar as razões para seus baixos desempenhos nessas provas. Se a prática de leitura de textos acadêmicos, técnicos, científicos ou especializados não os ajudou a desenvolver estratégias eficazes de leitura quando o contexto era a PNL, que fatores, então, precisam ser considerados para explicar o baixo desempenho desses graduandos?

Parece imprescindível problematizar as interinfluências entre os fatores envolvidos na compreensão leitora avaliada nas PNLs. Estarei me ocupando dessas possíveis correlações no próximo item.

A tabela a seguir nos traz os parâmetros em correlação.

Tabela 12: Questões mais difíceis – interinfluências de fatores⁶⁶

Ind Flesch do texto/ análise do gráfico	2009.2		2010.1		2010.2		
	Fácil	Muito difícil	Difícil	Difícil	Fácil	Difícil	Difícil
Habilidade / Competência	Identificação ou Inferência de significado de vocabulário	Identificação e recuperação de informações explícitas	Identificação e recuperação de informações explícitas	Reflexão avaliação	Interpretação e integração (síntese) do assunto principal (sentido global)	Reflexão avaliação	Reflexão avaliação
Tipo de item	Resposta única nível I	Resposta Múltipla	Resposta única nível II	Resposta única nível II	Afirmação incompleta	Resposta Múltipla	Foco negativo
Acertos	424	408	53	103	284	250	279
% acertos	44,9	43,2	9	17,6	47,2	41,5	46,3

Ind Flesch do texto/ análise do gráfico	2011.1			2011.2		
	Difícil	Fácil	Fácil	Fácil	Fácil	Fácil
Habilidade / Competência	Interpretação e integração (síntese) do assunto principal (sentido global)	Processamento sintático-semântico	Identificação e interpretação e/ou integração de ideias no nível local	Reflexão avaliação	Processamento sintático-semântico	Identificação ou Inferência de significado de vocabulário
Tipo de item	Foco negativo	Resposta única nível II	Afirmação incompleta	Foco negativo	Resposta única nível I	Resposta única nível II
Acertos	83	81	18	164	224	275
% acertos	47,2	46	10,2	28,4	38,8	47,6

Com base nos dados da tabela acima, podemos observar que os graduandos tiveram dificuldades com a identificação ou inferência de significado de vocabulário (2 itens dos 3 propostos) e a identificação e recuperação de informações explícitas (2 itens dos 9 propostos); a reflexão e avaliação do texto mostrou-se especialmente difícil (4 itens dos 14 propostos); a interpretação e integração do assunto principal (sentido global proposto em 5 itens, 2 difíceis) e a identificação e interpretação e/ou integração de ideias no nível local (1 item dos 12 propostos) também sinalizaram certa dificuldade dos graduandos em integrar ideias relevantes dos textos. Por fim, os dados também mostram uma dificuldade com paráfrase (único item proposto) e retomada anafórica pronominal na competência de processamento sintático-semântico (2 itens dos 3 propostos).

⁶⁶ Dadas as regras de formatação de margens para dissertações apresentadas à PUC-Rio, a tabela foi dividida em duas.

Ainda pelo panorama que a referida tabela nos dá, constata-se que a dificuldade maior relaciona-se com a competência em interagir o conhecimento enciclopédico com as novas informações trazidas pelo texto. A maior dificuldade dos graduandos da PUC-Rio na competência reflexão e avaliação do texto pode nos ajudar a trazer a velha discussão acerca da prática de leitura por parte dos leitores brasileiros. Sem cunho reducionista nem de caráter condenatório, o que as pesquisas sobre letramento nas IESs têm nos apontando é a fragilidade na prática de leitura por parte dos estudantes brasileiros. Neste sentido, Cunha & Santos (2004, p.7) constata que

Apesar do grande número de erros gramaticais e de construção lingüística encontrados normalmente nos textos produzidos por vestibulandos e universitários ingressantes, **o problema real não pode ser debitado ao ensino da gramática, mas à falta de leitura e conseqüente falta de informação e de referenciais culturais, por parte dos estudantes.** Vale destacar que os universitários, ao contrário dos leitores iniciantes, têm habilidades e estratégias que os ajudaram a entrar na universidade. **Então, na prática, esses universitários necessitam de estratégias que os ajudem no desenvolvimento das suas capacidades lingüísticas.** (grifos nossos)

Da mesma forma, Beluce et al. (p. 4) asseveram que

Pesquisas têm revelado que os universitários brasileiros apresentam dificuldades na compreensão em leitura, **destacando problemas referentes ao conhecimento prévios dos alunos e a habilidade em realizar a integração e associação de novas informações ao contexto em que se encontra inserido** (Cabral & Tavares, 2005; Cunha & Santos, 2006; Santos et al, 2004; Silva & Witter, 2008). (grifos nossos)

Há uma convergência entre essas constatações e os dados obtidos a partir da análise dos desempenhos de graduandos que realizaram a PNL por mais de duas vezes. Uma outra planilha foi necessária para computar os graduandos que se submeteram à PNL por mais de duas vezes, em semestres consecutivos ou em espaço de tempo superior a 6 meses. Mediante essa listagem, o desempenho individual de 06 leitores imaturos foi analisado almejando-se relacionar seus acertos com seu respectivo nível de letramento.

No anexo 3, vemos a trajetória de realizações de PNLs por parte de graduandos com perfil de leitor imaturo e seus acertos nos itens dessas provas. Todos os seis graduandos aqui analisados lograram acertar questões consideradas difíceis (em negrito) no computo geral dos desempenhos dos leitores em cada uma das PNLs. Como não temos ferramenta para nos certificarmos se o leitor imaturo

acertou na base do ``chute``, o que os dados nos sugerem é que as competências avaliadas foram construídas nestes seis leitores-universitários, porém não se desenvolveram até patamares considerados proficientes para a comunidade discursiva PUC-Rio.

Parece que podemos supor que a frágil prática de leitura crítica dos graduandos os torna leitores de baixa qualidade lexical (Perfetti, 2001) e de baixa experiência em integrar informações localizadas em diferentes partes do texto, ou de as integrar com seus conhecimentos prévios (Cunha & Santos, 2004). Em contrapartida, a menor dificuldade em identificar e interpretar e/ou integrar informações no nível local (1 item difícil dentre 12 propostos) pode, a princípio, parecer uma contradição que abala a assertiva de que nossos graduandos apresentam dificuldades de sintetizar informações. Podemos, assim, inferir que a dificuldade está no nível da macroestrutura do texto, desacentuando quando o leitor faz microanálise sem a preocupação ou mesmo o controle metacognitivo de integrar todas suas microanálises e, dessa forma, sem preocupação em manter o texto coeso em sentido global.

Em tempos de tecnologia da informação, com a consequente democratização da informação e do acesso a ela, talvez seja reducionismo teórico afirmar que nossos graduandos não lêem.

(...) é clara a progressiva valorização da leitura, à medida que avança a escolarização dos entrevistados: em todos os suportes (livro, revista, jornal e Internet), **o ensino superior define um índice maior de leitura**: os entrevistados com esse nível de ensino lêem muito mais que a média livros técnicos (35%), obras sobre História, Política e Ciências Sociais (37%), Ensaios e Humanidades (15%), Biografias (30%), e usam muito mais a Internet (31%)⁶⁷. (grifos nossos)

Todavia, os que os dados nos apontam é, de um modo geral, para uma ``**leitura de jornal**``, na qual cada informação é tratada de forma local, sem ligação com as demais. Em outras palavras, as informações são lidas como se fossem manchetes de jornal, as quais o graduando lê para se informar, mas não para produzir novos conhecimentos.

Após a análise qualitativa desses dados, tive a grata oportunidade de ler o trabalho de Paviani (2010), a qual fez um estudo com universitários,

⁶⁷ Ver, para tanto, Retratos da leitura no Brasil: olhares e reflexões, acessível em <http://www.leituracritica.com.br/pesquisa11/pdf/cadernoderefleoes-retratosep175.pdf>

considerando, dentre outras variáveis, as áreas de conhecimento dos estudantes (graduação), ingressantes e concluintes no Ensino Superior. Sua pesquisa aqui nos interessa no ponto em que ela fez um levantamento da preferência de leitura por parte desses universitários (o que mais gosta de ler?). Seus dados mostraram a tendência para leitura de jornais por alunos iniciantes, das diferentes áreas do conhecimento, pois é a leitura na primeira posição. A mesma tendência foi evidenciada entre os concluintes, no conjunto das áreas, ou seja, a preferência pela leitura de jornais manteve a 1ª posição (op. cit., p. 8).

Relativamente ao desempenho em leitura dos alunos iniciantes e concluintes, por área de conhecimento, Paviani pontua que seus

dados são indicativos de que, de um modo geral, o desempenho em leitura de alunos universitários, iniciantes e concluintes, de todas as áreas de conhecimento está a desejar, ou melhor, não corresponde às expectativas do ensino superior. (op.cit, p.8)

A prática de leituras rápidas e pouco desafiantes, uma vez levada para a vida acadêmica, esbarra com um volume considerável de leitura que toda graduação supõe, sem, contudo, se fazer acompanhar, muitas das vezes, da criativa construção de novos e substanciais saberes a partir da reorganização dos conhecimentos prévios em integração com as novas informações que os textos trazem. Ressalte-se ainda que os textos acadêmicos não necessariamente se organizam em narrativas que se desenvolvem linearmente no tempo, tal como os textos jornalísticos. Os textos acadêmicos muitas das vezes se organizam em torno de ideias, o que nos faz concordar com Paviani: ``Enquanto os alunos forem fazendo, de modo mais frequente e preferencialmente, leituras breves e, às vezes, superficiais, de textos jornalísticos e similares, poucos avanços se lhes ocorrem em termos de habilidades de leitura`` (idem, p.12).

A seguir vamos nos aproximar dessas questões difíceis e trazer uma discussão sobre as interinfluências entre fatores psicolinguísticos que possam ter respondido pelo baixo desempenho dos graduandos nessas questões tidas como difíceis por eles.

8.5.1

Questões difíceis da PNL 2009.2

No que tange à PNL 2009.2, ela oportunizou que os universitários lidassem com um texto fácil e um muito difícil do ponto de vista do índice Flesch; a seu turno, a tabela (texto 3) foi categorizada como muito difícil. Neste cenário, duas questões de múltipla escolha resultaram difíceis para os universitários.

Observa-se que a questão 03, relacionada ao texto 1 (fácil), foi a de maior grau de dificuldade na PNL 2009.2. Trata-se de uma questão que explora a qualidade lexical do leitor-universitário, ou seja, seu conhecimento de forma e de sentidos de palavras em um texto de índice Flesch fácil. O item de resposta única apresentava um único item lexical em cada alternativa, ou seja, uma tipologia fácil do ponto de vista do processamento cognitivo.

Sob a perspectiva interacional, vemos que o nível fácil de inteligibilidade do texto associado a uma tipologia de item fácil não garantiu um percentual de acerto mais substancial no item 03 (44,9%). Este dado talvez nos autorize a afirmar, como observado acima, que a fraca prática de leitura de textos mais desafiantes para o leitor está associada, de forma estreita, a uma baixa qualidade lexical.

Atente-se ainda que o próprio enunciado deste item pode ter influenciado no desempenho dos graduandos, visto que há duas ocorrências de participio – conhecido e desencadeado -, os quais nem sempre favorecem uma leitura fácil quanto a serem nominalizações adjetivas ou formas verbais. Concorrem para a manutenção da ambiguidade as cinco alternativas: explorado, deflagrado, destacado, revelado, liberado. Parece razoável afirmar que os três últimos são mais facilmente interpretados como adjetivos, dada a frequência com que são usados enquanto nominalizações; enquanto explorado e deflagrado, como formas verbais. O item pede um sinônimo para o participio desencadeado em: “Eu gostaria de explorar um debate menos conhecido desencadeado pelo 11 de setembro”. No caso, o participio é um exemplo de estrutura complexa, qual seja, uma passiva analítica reduzida. Se o leitor não logrou interpretar e compreender essa passiva dentro do texto, tampouco a compreenderia na resolução da questão, o que o pode ter levado a assinalar o que ele interpretou como adjetivo.

A competência de acessar e recuperar informações explícitas foi a prevista para o item 5, cujo baixo percentual de acerto (43,2%) o situa como difícil. No entanto, os itens 02, 07 e 08 também avaliaram esta competência, tendo sido seus percentuais de acerto bastante expressivos (86,8%, 75,1%, 86,2%, respectivamente). Isto nos remete à consideração de que uma competência pode ser testada por diferentes tipos de itens, respondendo esta tipologia também pela maior ou menor dificuldade da questão. No caso, o item 5 trouxe a habilidade de acessar e recuperar informações explícitas dentro do tipo de item de resposta múltipla, na qual o leitor precisa considerar todas as afirmativas corretas para marcar uma chave de resposta, o que requer várias etapas de processamento de resolução da questão. Leitores imaturos que se encontram em nível básico de letramento (INAF, 2011) apresentam habilidades leitoras instáveis, donde o aumento no número de etapas de processamento da informação ser especialmente difícil para esse grupo de leitores. Neste caso, o item 5 buscou avaliar a compreensão local de diferentes partes do texto 2, cujo índice Flesch o situa no nível muito difícil. Parece que podemos supor que o casamento entre item de resposta múltipla (múltiplos processamentos) e texto difícil tenha respondido pelo baixo desempenho de 56,8% dos leitores avaliados.

8.5.2

Questões difíceis da PNL 2010.1

O item 1 da PNL 2010.1 explorou a estratégia de retornar ao texto para checar e controlar o grau de compreensão de uma informação explícita. O leitor identifica a informação e recupera, de forma crítica, sua compreensão da passagem. O checar e controlar de forma consciente sua própria compreensão leitora, como visto no capítulo 6, é definido na literatura como metacognição. O item explorou as estratégias de metacognição no texto 1, de índice Flesch difícil.

Ressalte-se que o texto apresenta apenas três parágrafos e, pela análise computacional, estes parágrafos estão estruturados em um total de 23 períodos. Disso resulta que os parágrafos são longos e cheios de informação sobre

mudanças nos padrões de consumo desde a chamada Revolução Industrial. Opera, nesse caso, a complexidade informativa (Kleiman) tanto quanto a densidade informacional dos parágrafos, tendo ambos os aspectos possivelmente influenciado a própria formulação do item.

O tipo de item é o de resposta única, nível II, ou seja, enuncia a situação problema e apresenta as alternativas constituídas de proposições longas (3 ou mais orações). Parece que neste caso a dificuldade do item foi suscitada pela própria complexidade da operação de analisar as proposições apresentadas nas alternativas, compará-las com as informações do texto, e, assim, julgar se as proposições estavam ou não corretas. Tal processamento inflaciona o tempo de resolução do item ainda que a habilidade de leitora seja básica, ou seja, acessar e recuperar informações explícitas. O fato é que 91% dos graduandos não lograram pontuar essa questão.

Já na questão 2, a banca previa que o leitor fizesse o uso de informações extratextuais, ou seja, seu conhecimento enciclopédico, de tipologia e gênero textuais, suas experiências com os vários discursos em que o tema do texto pode ser explorado, para refletir e avaliar o mesmo. Convém mais uma vez analisar o item em interação com o todo da prova para que se obtenha melhor compreensão de que fator atuou negativamente no desempenho dos leitores-universitários.

O que se observa é que a competência de refletir e avaliar apareceu em outros itens, quais sejam, 08, 09 e 10, contudo talvez seja de novo o caso de se analisar a competência em interação com a tipologia de cada item aqui. E não só: a própria inteligibilidade e o gênero dos textos em que tais questões se basearam podem responder pela diferença no desempenho dos graduando nas 04 questões que avaliaram a competência de refletir e avaliar. Neste sentido, os textos das 04 questões foram categorizados como difíceis, todavia as questões 08, 09 e 10 foram em cima de um gráfico. Já a questão 4 pautou-se em um texto contínuo, argumentativo, e ainda de complexidade informativa (Kleiman), dado que o tema consumo é tratado desde o século XVIII. A distância temporal entre o assunto do texto e o leitor representa uma complexidade informativa.

Por um lado, o texto contínuo é geralmente constituído de sentenças que, a seu turno, se organizam em parágrafos sem imagens ou gráficos. O texto contínuo e longo pressupõe o leitor como uma pessoa que consiga se concentrar e, portanto, focar sua atenção por um tempo relativamente extenso no material escrito sem o

apoio de ilustrações, cores e outros sinais para facilitar sua compreensão textual. Pelo exposto, talvez os dados nos estejam também sinalizando que o gênero gráfico é mais inteligível que um texto contínuo. Assim, embora a habilidade descrita fosse idêntica nos itens 02, 08, 09 e 10, o grau de dificuldade da questão 04 foi maior para o texto contínuo, o que se refletiu no desempenho de 82,4% de leitores-graduandos.

8.5.3

Questões difíceis da PNL 2010.2

Nas questões 6, 9 e 10 da PNL 2010.2, previa-se o uso de operações cognitivas que demandam maior nível de abstração. No item 6, o leitor precisava interpretar informações localizadas em diferentes partes do texto, e as integrar em uma síntese que desse sentido global ao texto (de parágrafo para parágrafo, de parágrafo para texto). Por isso, tem-se na interpretação uma competência de maior nível de abstração, ou seja, as informações processadas precisam passar do plano de dados dos sentidos (sistema perceptivo) ao plano da abstração de suas localizações no texto ou mesmo de sua ordenação proposta pelo autor do texto. Tudo isso precisa ocorrer sob o domínio consciente e deliberado do leitor, o que sem dúvida invoca a competência de criar um distanciamento da situação concreta em que os dados foram automaticamente lidos e inconscientemente processados para ceder lugar a um maior esforço consciente de reorganização mental: o resultado é o texto interpretado. Dos 602 universitários avaliados, somente 47,2% demonstram competência para lidar com um item de maior complexidade e abstração.

As questões 9 e 10 também demandaram que o leitor fizesse uso de um nível maior de abstração, condicionando a reconstrução da textualidade pela reflexão e avaliação das informações e ideias contidas no texto original. Para tanto, o leitor precisava fazer interagir seu conhecimento enciclopédico e suas experiências pessoais com o texto, emprestando-lhe, por conseguinte, sentidos. Contudo, para refletir sobre e avaliar as representações mentais construídas a partir do material escrito, o leitor precisa pensar de forma bem abstrata, logrando

construir inferências e hipóteses por si próprio, e mantendo na memória processual todas as representações importantes que construiu ao ler as passagens do texto. A representação mental e organizada do texto demanda igualmente um certo distanciamento do material escrito, o que ocorrerá em função, dentre outros fatores, da maturidade do leitor em confiar em suas interpretações ao invés de somente dar elevado peso à informação provida pelo autor do texto.

Se a competência de interpretar e integrar já anunciava uma dificuldade do leitor avaliado na PNL 2010.2, o uso de recursos extratextuais para se ter um maior alcance de compreensão do texto foi ainda mais desafiador para esses leitores. Atente-se ainda que o item 09 foi do tipo de resposta múltipla, o que capitalizou maior dificuldade para 58,5% dos leitores com baixo desempenho na questão. No mesmo diapasão, a tipologia do item 10 é fonte de dificuldade que se acumula com a competência de avaliar e refletir. O item é do tipo de foco negativo, o qual é tida como difícil. Costa (1998, p, 21) argumenta que ``as perguntas que contêm uma negação na raiz costumam suscitar dificuldades na compreensão da instrução, se implicam converter à forma negativa o que no texto lido é apresentado em forma afirmativa``. Isto pressupõe estabelecer uma relação negativa, o que implica uma operação lógica de maior nível de abstração. De fato, Graesser et al. (2008, p. 12) ressaltam que ``negações são comparativamente difíceis para os humanos processarem (...)``⁶⁸, o que podemos constatar com 53,7% dos graduandos que não obtiveram sucesso nesta questão.

8.5.4 **Questões difíceis da PNL 2011.1**

A questão 4 exigiu do leitor a competência de interpretar e integrar de forma global o texto, donde o controle consciente de suas análises locais estar necessariamente envolvido. No caso da PNL em comento, este controle também foi requerido na questão 8, a qual traz de novo o leitor à consciência de sua competência para integrar informações no nível local, estabelecendo relações entre informações do texto dentro de um mesmo parágrafo ou em parágrafos não

⁶⁸ Tradução do original ``negations are comparatively difficult for humans to process (...)``

contínuos. Como vimos, a precária prática de leitura de textos mais desafiantes torna a integração e interpretação uma competência pouco desenvolvida entre graduandos.

Finalmente, a questão 7 apostou tanto no conhecimento linguístico quanto gramatical do leitor, quando então o leitor precisaria analisar as paráfrases de uma proposição dada no enunciado da questão. As cinco alternativas eram re-escrituras desta proposição dentro de um item de resposta única, nível II, ou seja, alternativas com proposições mais longas (3 ou mais orações).

Neste item, ao menos três operações cognitivas eram demandas: 1) construir o sentido de cada uma das proposições apresentadas nas 5 alternativas; 2) comparar as proposições com o trecho original do texto 2 de índice Flesch fácil; 3) comparar as proposições entre si de forma a escolher a que melhor parafraseou o trecho trazido do enunciado. Pesquisas têm apontado que há uma razão inversamente proporcional entre o número de etapas de processamento cognitivo e competência leitora entre leitores imaturos. Como vimos, os dados do INAF 2011 mostraram que 34% dos universitários brasileiros (graduandos e formados) se encontram no nível básico (alfabetismo funcional), pois, dentre outras razões, eles lêem e compreendem textos de média extensão, mas revelam limitações quando as operações requeridas envolvem maior número de elementos, etapas ou relações.

8.5.5

Questões difíceis da PNL 2011.2

Os dados apontam para a PNL de 2011.2 como a prova mais difícil. Os índices Flesch de seus três textos os situam na categoria de texto fácil, muito difícil e difícil, respectivamente, sinalizando os dois últimos como textos adequados para alunos cursando o ensino médio ou universitário.

Na questão 3, temos de novo a competência de refletir e avaliar sendo testada. A propósito, esta é uma competência bem avaliada pela banca ao longo das 5 provas (14 itens). Na prova em tela, esta competência foi a eleita em 03

questões, sendo 02 exploradas no tipo de item de resposta única, nível I, ou seja, as alternativas têm uma única palavra ou um período curto. Em contrapartida, a questão 3 foi explorada no tipo de questão conhecido como foco negativo, isto é, o leitor precisa inibir as alternativas corretas para focar na alternativa errada ou inadequada e marcá-la.

Uma base empírica, por assim dizer, para a escolha do pronome demonstrativo anafórico como elemento que tem seu peso no custo de processamento de anáforas e que, por conseguinte, de certa forma abala a inteligibilidade textual pode ser buscada na questão 4. Esta questão requeria que o graduando assinalasse a alternativa correta em que o pronome anafórico e o termo antecedente sublinhado mantivessem relação correferencial. Os elementos anafóricos nem sempre são facilitadores da compreensão leitora. De fato, para um leitor proficiente as anáforas podem tornar o texto mais fácil porque mais coeso, mas a mesma facilidade não é sentida pelo leitor imaturo, visto que, em algumas situações, um elemento pronominal pode, em termos potenciais, ser associado a mais de um antecedente.

O baixo percentual (38,8%) de acerto na questão 4 é revelador da dificuldade suscitada pelo fenômeno da referenciação. Não obstante a coesão do texto, para o leitor pouco proficiente parece que, quanto maior a distância entre o termo antecedente e a forma remissiva, maior a dificuldade de processar o fenômeno da referenciação. No entanto, na questão 4, as referências anafóricas são adjacentes, ou seja, elas ocorrem entre sentenças adjacentes. Busquei, assim, analisar as alternativas assinaladas pelos leitores de forma a construir uma base de entendimento acerca do grau de aceitabilidade da opção marcada.

Cumprime primeiramente esclarecer que, por questões técnico-administrativas, das 578 provas feitas tive acesso a 400. Esclareça-se que somente foram computadas as respostas marcadas de forma inequívoca pelos leitores. Questão em branco ou assinalada de forma a deixar qualquer elemento de dúvida foi desconsiderada. Desta forma, foram registradas as respostas de 357 leitores, sendo que 143 gabaritaram a questão, ou seja, 40% marcaram a letra D). Os 60% restantes respondem pelas marcações que os leitores fizeram nas letras: A) 58 ; B) 18; C) 52; E) 86.

Chama a atenção a quantidade de graduandos que assinalaram a letra (e) abaixo transcrita:

E) Em comparação com a ciência, o conhecimento espontâneo é fragmentário, por não estabelecer conexões em situações em que essas poderiam ser verificadas. (l.39 e 40) – **situações**

Possivelmente o que respondeu aqui pela dificuldade dos graduandos em compreender a relação correferencial é o fato de que dois termos femininos e no plural aparentemente se perfilam a candidato a antecedente: conexões e situações. Atente-se também que, pragmaticamente falando, tanto situações quanto conexões podem ser verificadas, o que pode ter conduzido o leitor a uma leitura mais ancorada em suas experiências linguístico-discursivas do que nas pistas textuais e mesmo nas relações estabelecidas entre os termos do período em que esses itens lexicais aparecem.

No que tange à semântica, leitores pouco proficientes parecem menos flexíveis para atribuírem diferentes sentidos a uma mesma palavra, dependendo do contexto em que ela é usada no texto. A próxima questão que vou analisar prevê a habilidade leitora de identificação ou inferência de significado de vocabulário, quando, então, a mesma frase é reescrita tendo apenas um de seus termos sido substituído por outro. O leitor deveria marcar a opção na qual a substituição foi feita pelo sinônimo correto.

Assentando minhas análises sobre o rápido acesso e eficaz retomada de um item lexical na memória do leitor, podemos questionar o que caracteriza uma alta qualidade lexical. Perfetti aponta dois fatores que respondem por essa alta qualidade:

A representação lexical precisa ter tanto a informação semântica quanto a fonológica fortes o suficiente para recuperar sua localização na memória. E, de outro lado, essa qualidade deve ser duradoura o suficiente para que os processos subsequentes possam desenvolver suas funções` (Tradução do original. Perfetti, 2001, 2).

Perfetti finaliza definindo a qualidade da representação lexical como o conhecimento detalhado sobre formas e sentidos da palavra (idem, 10).

Formas e sentidos são os conhecimentos linguísticos envolvidos na questão 9, a qual foi tida com difícil pelos leitores (47,6% de acertos). Contudo, o que observei foi que houve 111 leitores que obtiveram nota igual ou superior a 6,0, mas erraram a questão em comento. Estes dados nos fazem questionar o que

separa os leitores em dois grupos: proficientes e imaturos. Que fatores, então, precisam ser considerados para entendermos os resultados da questão 9?

Para analisarmos a eficácia de uma determinada habilidade de leitura, precisamos fazê-la dialogar com outros fatores. A relativização dos resultados na PNL 2011.2 precisa considerar a relação entre diversos fatores, tal como tivemos a oportunidade de observar na análise da questão 4. Neste sentido, podemos concordar com Dubay (2004, 5) que alerta para o caráter desmotivador de um texto considerado muito difícil: ``When texts exceed the reading ability of readers, they usually stop reading``⁶⁹. A competência envolvida na questão 9 é relativamente simples: inferência de vocabulário. O que aumenta a dificuldade da questão é sua tipologia, qual seja, um item de foco negativo porque apresenta várias respostas corretas e apenas uma incorreta, que é a solicitada. Entende-se que esta tipologia é difícil porque a solicitação da resposta incorreta orienta o leitor para a busca do erro, e não do acerto. Ademais, três etapas de processamento cognitivo estão envolvidas na resolução deste item: 1) O leitor precisava primeiramente localizar no contexto de origem a informação retomada na alternativa para se certificar de sua compreensão acerca do sentido do termo sublinhado e substituído. Como a questão trabalha com os três textos desta PNL, previa-se que o leitor seria capaz de lidar com a intertextualidade em textos de índice Flesch fácil; 2) O leitor deveria construir nova representação contemplando, agora, o sinônimo fornecido na alternativa; 3) O leitor deveria comparar as duas representações – trecho original e trecho com sinônimo - e checar a relação de sinonímia entre elas.

A questão 9 parece ser um exemplo para se considerarem fatores em interação quando se busca compreender uma dada competência leitora instável. Neste caso, a questão fundou-se em textos de índice Flesch fácil e a habilidade em si requerida situa-se, primeiramente, no eixo da lexicalidade, ou seja, no conhecimento linguístico do leitor. No entanto, ambos aspectos não se traduziram em bom desempenho acadêmico no item 9 (47,6 % de acertos), o que nos faz trazer de volta a pergunta: o que separa os leitores em dois grupos: proficientes e imaturos?

⁶⁹ Quando o texto excede a habilidade leitora dos leitores, eles normalmente param de ler. (T.A.)

Identificação e inferência de vocabulário, interpretação de trechos e comparação de sentidos fazem desta questão um item híbrido, pois que a identificação do sentido de um item lexical em um trecho necessariamente abrange o domínio sintático-semântico para entender as relações entre termos da proposição. Além disso, a competência avaliada neste item propunha o reconhecimento da linguagem verbal como objeto de reflexão sobre sua função e uso social, quando, então, diferentes formas de se expressar uma mesma ideia estava em jogo. Este fato gera para o leitor o compromisso de usar seus conhecimentos esquemáticos acerca da representação da cultura e das regras sociais de interpretação dentro de sua comunidade discursiva com o fim de resolver a questão.

Em suma, a competência básica de identificação ou inferência de sentido não se resumiu à habilidade de acessar a palavra no léxico mental do leitor. Mais que isso: a questão requereu um maior alcance e complexidade da análise, o que pode ter interferido no desempenho dos graduandos.