



O padrão lexical dos demonstrativos (in)variáveis no *corpus* CORUEPA: a realidade linguística como ferramenta de ensino-aprendizagem

Dóris Figueiredo de Souza¹
Maria Sílvia Jucá Acácio²

Resumo

Este estudo tem por objetivo analisar como os candidatos do PRISE e do RICEX usaram os demonstrativos (in)variáveis nos vestibulares da Universidade do Estado do Pará (UEPA); a fim de incentivar o docente a valer-se das variedades da língua, estimular o envolvimento dos alunos nas aulas de língua portuguesa e questionar se a gramática ensinada na escola valoriza a construção identitária dos sujeitos nas respectivas comunidades linguísticas. O estudo objetiva listar a frequência dos demonstrativos nas redações do CORUEPA, categorizar a configuração de uso e identificar o padrão de colocação dos demonstrativos no PRISE e RICEX. O estudo fundamenta-se nos aportes de Sardinha (2004), Bagno (2001; 2012) e Castilho (2014). Na metodologia aplica-se o corpus nas ferramentas de análise do programa WordSmith Tools 5.0 (Scott, 2008). O resultado das amostras revela o uso prevaletente da colocação anteposta e o desuso do pronome aquele e do sistema ternário, diante do predomínio do pronome esse/essa/isso e do sistema binário este/esse. Espera-se com este estudo, contribuir para o processo de ensino-aprendizagem da língua portuguesa, ao considerar o padrão de uso dos candidatos da UEPA e propor que os docentes apliquem como ferramenta de ensino a realidade linguística dos alunos, construída dentro/fora da escola.

Palavras-chave:

Linguística de de Corpus; Língua Portuguesa; Descrição Linguística

Sobre os autores:

¹ Graduada em Letras Língua Portuguesa pela Univerisdade do Estado do Pará.

E-mail: dorisfigueiredo1999@gmail.com | **ORCID:** 0000-0001-7903-6927

² Doutora em Língua Portuguesa e Filologia pela Universidade do Estado do Pará.

E-mail: marajuca@uepa.br | **ORCID:** 0000-0002-2591-4533

CONSIDERAÇÕES PRELIMINARES

Os estudos em Linguística Computacional favorecem o ensino de línguas e possibilitam o avanço na compilação de uma ampla quantidade de dados (*corpus*), com a precisão de resultados. Neste artigo, objetiva-se analisar o uso de demonstrativos (in)variáveis no *corpus* de estudo da UEPA, doravante CORUEPA¹, para observar como os candidatos usaram esses pronomes nos processos seletivos da Universidade do Estado do Pará (UEPA).

A pesquisa fundamenta-se na escrita dos candidatos tanto da 3ª etapa do Programa de Ingresso Seriado (PRISE)² de 2016, quanto do Processo Seletivo Especial dos cursos de Relações Internacionais e de Tecnologia em Comércio Exterior (RICEX)³ de 2017, ao colocar em pauta, se os candidatos usaram os demonstrativos conforme a gramática prescritiva, marcada pelo sistema ternário este/esse/aquele, ou se usaram os pronomes conforme o binário esse/aquele incluso na língua falada.

Desse modo, os objetivos específicos para este estudo incluem listar a frequência dos demonstrativos nas produções textuais do CORUEPA, com auxílio da ferramenta *WordList*, do programa *WordSmith Tools*⁴; categorizar a configuração de uso dos demonstrativos nas redações; e identificar o padrão de colocação dos demonstrativos (in)variáveis tanto no PRISE como no RICEX, já que se trata de públicos-alvo distintos.

A pesquisa baseia-se teoricamente na discussão acerca do ensino de língua padrão e linguagem coloquial; do uso de demonstrativos (in)variáveis no Português Brasileiro; do desuso do pronome demonstrativo; da Linguística de *Corpus* no ensino de línguas, a partir de autores como Sardinha (2004), Porfirio; Bidarra (2004), para compor o estudo da Linguística de *Corpus*; bem como os linguistas Bagno (2001; 2012), Castilho (2014), Haug (2015), Cunha e Cintra (2017), para auxiliar na discussão sobre os demonstrativos.

¹ O CORUEPA é o *Corpus* de Estudo da UEPA, origina-se no projeto de pesquisa do Programa Institucional de Bolsas de Iniciação Científica (PIBIC), sob o título “*Compilação de corpus composto por redações de processos seletivos da UEPA, com vistas à criação de banco de dados, para análise linguística e ensino de língua materna*”, de autoria do Prof. Dr. Ednalvo Apóstolo Campos e colaboração dos integrantes Dóris Figueiredo de Souza e João Pedro Saraiva. O projeto contempla 400 redações da 3ª Etapa do PRISE e 33 do RICEX. O CORUEPA efetivou-se após a etapa de transcrição manual e limpeza do *corpus*.

² O PRISE era o Programa de Ingresso Seriado da UEPA, implantado em 1997. O *corpus* desta pesquisa refere-se as redações da 3ª etapa do PRISE de 2016. Este programa era composto por três etapas, correspondente aos três anos do Ensino Médio.

³ RICEX é o Processo Seletivo Especial para os cursos de Relações Internacionais e de Tecnologia em Comércio Exterior (RICEX). Na UEPA, fez parte de um projeto estadual, que visava à capacitação de profissionais e sociedade. Os dados utilizados para este estudo são os do processo seletivo do ano de 2017.

⁴ Software do Windows destinado a encontrar padrões de palavras. Publicado pela Lexical Analysis Software e Oxford University Press, desde 1996. Disponível em: <https://www.lexically.net/wordsmith/>. Acesso em 23.03.2022.

Esta pesquisa visa contribuir para o processo de ensino-aprendizagem, a partir da análise do padrão de uso dos demonstrativos. A motivação para este estudo parte do envolvimento com o projeto do Programa Institucional de Bolsas de Iniciação Científica (PIBIC), intitulado “*Compilação de corpus composto por redações de processos seletivos da UEPA, com vistas à criação de banco de dados, para análise linguística e ensino de língua materna*”, posto que a Linguística Computacional permite que o *corpus* CORUEPA sirva para pesquisas na área da descrição linguística e para o ensino de língua portuguesa.

O ENSINO DO PORTUGUÊS BRASILEIRO: A LÍNGUA PADRÃO E A LINGUAGEM COLOQUIAL

Bagno (1999, p. 34) afirma que o português não padrão ou coloquial é transmitido naturalmente pela tradição oral, enquanto o português padrão é adquirido artificialmente no âmbito escolar a partir da tradição escrita. Recomenda-se que o docente de língua portuguesa propicie o contato dos alunos com o português padrão e o coloquial, para não desenvolverem uma baixa “autoestima linguística” (idem, 2001) e para saberem se adequar às situações comunicativas e ambientais.

Para exemplificar, observa-se que a depreciação do uso não padrão ocorre desde o século XVII - XVIII, quando a língua geral esteve constante na fala da população paulista, até que deixasse de ter notoriedade, devido à expansão da língua portuguesa, que limitava o uso do tupi apenas ao âmbito doméstico, pois os falantes deveriam aparentar aceitação ao prestígio do português, de acordo com Holanda (2014). Assim, do mesmo modo que havia falantes da língua portuguesa contrários ao uso da língua geral, na atualidade, há contextos em que a linguagem coloquial ainda é desvalorizada diante da língua padrão, uma vez que,

As gramáticas normativas brasileiras permanecem apegadas ao mito da “língua única” e se apresentam como “descrição” de uma variedade linguística supostamente empregada pelas “pessoas cultas” do país, isto é, pelas classes dominantes, apresentando-a sempre como o “padrão” a ser imitado (BAGNO, 2001 p. 26)

Desse modo, o enaltecimento da língua padrão nas gramáticas normativas é visto como parâmetro de uso ou variante de prestígio, e isto dificulta a aceitação da linguagem coloquial. Para Faraco (2008, p. 74) “*seu prestígio não decorre de suas propriedades gramaticais, mas de processos sócio-históricos que agregam valores a ela*”. Em consonância, Bagno (2009) discute que os gramáticos, ao invés de elaborarem as suas regras do uso, a partir da escrita de autores brasileiros, por exemplo, preferem extrair amostras preconcebidas de obras literárias, isto é, “*opções linguísticas que eles, gramáticos, já de antemão consideram as boas, as*

bonitas, as corretas” (*Ibid.*, p.157).

Faraco (2008, p. 76) defende que “*as gramáticas e os dicionários adquiriram, então, certa força coercitiva*”, por interferirem no ensino de língua portuguesa e na liberdade linguística dos indivíduos, já que ao optar por um uso coloquial na redação, o emprego torna-se inadequado ao contexto, e pode custar uma boa pontuação e o ingresso na universidade. Nesse sentido, qual método se deve adotar no ensino da língua portuguesa nas escolas? Possenti (2000) propõe que a prática docente encoraje a gramática internalizada do aluno, para poder aproximá-lo da gramática normativa, isto é, possibilitar que construa o conhecimento, ao invés de impor um padrão de língua a ser seguido.

Segundo Bagno (2001) a escola deve aproximar o aluno das manifestações linguísticas, “*concretizadas no maior número possível de gêneros textuais e variedades de língua: rurais, urbanas, orais, escritas, formais, informais, cultas, não cultas*” (p. 157). Logo, o ensino de língua padrão e da linguagem coloquial nas escolas é importante para que o aluno saiba adaptar-se aos diferentes contextos conversacionais.

Outrossim, Possenti (2000) afirma que desconhecer a variedade padrão da língua não impossibilita que o indivíduo utilize suas competências para se comunicar, visto que “*é perfeitamente possível aprender uma língua sem conhecer os termos técnicos com os quais ela é analisada*” (p. 53), isto é, o domínio competente da língua não se restringe ao uso da variedade padrão.

Entretanto, há contextos que incentivam esta restrição a língua padrão, como os vestibulares que cobram a adequação dos candidatos à gramática normativa, conforme demonstram os critérios das redações do PRISE e RICEX, tais como: a adequação ao tema; adequação a coletânea; coesão e coerência; adequação ao tipo de texto (dissertativo/ narrativo); e o uso adequado da norma padrão. Em suma, as redações dos processos seletivos evidenciam a realidade linguística dos sujeitos, pertencentes aos diversos segmentos sociais, uma vez que, o uso do português não padrão nas redações, ocorre independente da exigência de adequação a língua padrão.

METODOLOGIA O USO DE PRONOMES DEMONSTRATIVOS (IN)VARIÁVEIS NO PORTUGUÊS BRASILEIRO

Os pronomes demonstrativos são definidos, em termos gerais, como aqueles que “*indicam a posição dos seres em relação às três pessoas do discurso*” (BECHARA, 2009 p. 140). Estes pronomes apresentam formas tanto variáveis (variação em gênero e número) quanto invariáveis, conforme demonstra o quadro (1)

Quadro 1 – Formas dos pronomes demonstrativos

PESSOA	VARIÁVEIS				INVARIÁVEIS
	MASCULINO		FEMININO		
1 ^a	este	estes	esta	estas	isto
2 ^a	esse	esses	essa	essas	isso
3 ^a	aquele	aqueles	aquela	aquelas	aquilo

Fonte: Adaptado pelas autoras, a partir de Cunha e Cintra (2017, p. 343)

No quadro (1) notam-se os pronomes demonstrativos situados a uma pessoa no discurso. Os pronomes variáveis são descritos na 1^a pessoa como *este/estes* e *esta/estas*; na 2^a pessoa, como *esse/esses* e *essa/essas*; e na 3^a, como *aquele/aqueles* e *aquela/aquelas*. Dentre as formas variáveis descritas, temos os masculinos: *este(s)*, *esse(s)*, *aquele(s)*; e os femininos, *esta(s)*, *essa(s)*, *aquela(s)*. Já nas formas invariáveis, temos *isto* na 1^a pessoa; *isso* na 2^a pessoa e *aquilo*, na 3^a.

A título de registro, observa-se que além dos invariáveis mencionados no quadro (1), Rocha Lima (2011, p. 160) elenca vocábulos como: *mesmo(s)*, *mesma(s)*, *próprio(s)*, *própria(s)*, *tal*, *tais*, *o(s)*, *a(s)*, que funcionam como invariáveis, assim como os advérbios, *aí*, *ali*, *aqui*, *lá*, que acompanham os demonstrativos, a exemplo:

(I) Este *aqui* terminou a prova.

Vale ressaltar, que os vocábulos latinos *ipse, ipsa* > *mesmo*; *ipsum* > *próprio*, já funcionavam como demonstrativos – ver, Almeida (2000, p. 162). O autor acrescenta que os pronomes pessoais, *is* > *ele (este)*, *ea* > *ela (esta)*, *id* > *o (isto, isso, aquilo)* também exerciam essa função.

Como vimos, a definição dos demonstrativos permeia outras classes de palavras, e este fator incentiva Castilho (1993) a nomear essa função como *mostrativos*, à medida que “*se referem ao conjunto de palavras que ocorrem dispersas em outras classes*” (BAGNO, 2012, p. 773). Logo, as classes como artigo definido, pronome demonstrativo, pronome pessoal, advérbio, integram os mostrativos. Em contrapartida, nas gramáticas consultadas, nota-se que os autores não adotam a nomenclatura dos mostrativos e se referem às classes mencionadas, como todas pertencentes aos demonstrativos. Neste artigo, a análise considera o conceito dos mostrativos, mas prioriza os pronomes demonstrativos (in)variáveis.

Segundo Cunha e Cintra (2017, p. 343) e Haug (2015, p. 692) os demonstrativos na forma variável exercem a função de pronomes adjetivos ou pronomes substantivos, a depender da posição que ocupam no sintagma, enquanto os invariáveis funcionam apenas como pronomes substantivos. No caso dos demonstrativos variáveis, quando exercem a função de pronome adjetivo,

podem vir antepostos ou pospostos aos substantivos, para indicar o quanto a fala/escrita foi monitorada, já que “na língua falada, o demonstrativo posposto exprime um alto grau de monitoramento linguístico” (BAGNO, 2012 p. 495).

Quadro 2 – Funções dos demonstrativos na forma variável feminina

PRONOME	PESSOA	DÊITICA	
		ESPAÇO	TEMPO
Esta	1 ^a	situação próxima	presente
Essa	2 ^a	situação intermediária ou distante	passado ou futuro pouco distante
Aquela	3 ^a	situação longínqua	passado vago ou remoto

Fonte: Adaptado pelas autoras, a partir de Cunha e Cintra (2017, p. 345)

No quadro (2) nota-se a função dêitica, descrita por Cunha e Cintra (2017) como temporal e espacial. Na 1^a pessoa, a dêixis espacial é usada quando o que se quer indicar estiver próximo ao falante, no tempo presente; o de 2^a pessoa, quando o objeto estiver próximo ao ouvinte, no tempo passado ou futuro; e o de 3^a pessoa, ocorre ao estar afastado tanto do falante como do ouvinte, no tempo remoto. Isto ocorre desde o latim, visto que *hic > isto*, refere-se a um “objeto presente ou próximo”, e *ille > aquele*, a um “objeto ausente ou afastado” (ALMEIDA, 2000, p. 162).

Por vezes, o uso dos demonstrativos está mais associado ao espaço da função dêitica, do que ao tempo, já que este uso “escapa à correspondência estrita com as (não) pessoas do discurso e depende, bem mais, das relações afetivas de proximidade ou distância”, de acordo com Bagno (2012, p. 792).

Assim, a definição de que os demonstrativos indicam a posição dos objetos em relação à 1^a, 2^a e 3^a pessoa no discurso, é suficiente para que o falante consiga usar e distinguir esses demonstrativos na escrita? Cunha e Cintra (2017, p. 345) defendem que as distinções desses pronomes não são “rigorosamente obedecidas na prática”, uma vez que, os falantes diferenciam os usos conforme a “predominância dos seus valores estilísticos sobre seus valores gramaticais” (ROCHA LIMA, 2011, p. 400).

O DESUSO DO PRONOME DEMONSTRATIVO

Bagno (2012, p. 123) discute sobre a substituição do sistema ternário *este/esse/aquele* pelo sistema binário *esse/aquele* na língua portuguesa falada. E acrescenta que, essa mudança “plenamente instalada” já é visível na língua escrita, com uma “ampla repercussão nos gêneros escritos monitorados”.

Para o autor (*op. cit.*), a predominância do sistema binário ocorre devido ao processo de assimilação, quando um segmento integra o outro com característica igual ou semelhante. Nos demonstrativos, este processo efetua-se quando o “grupo -st presente em *este/esta/isto* tende a sofrer assimilação” (*Ibid.*, p. 793), isto é, quando o -s induz a assimilação à consoante -st. Além disso, defende que os demonstrativos *este/esta/isto* desapareceram do vocabulário dos brasileiros, mas ressurgiram nas escolas mediante o ensino que prioriza o sistema ternário, conforme afirma em

Comprovam o desaparecimento dos demonstrativos com -st no vernáculo geral brasileiro, a ponto de podermos dizer que eles não pertencem a nossa língua materna, constituindo um desses muitos conjuntos de formas gramaticais aos quais só temos acesso por meio do letramento escolar (BAGNO, 2012, p. 795)

Portanto, o sistema ternário *este/esse/aquele* não predomina no Português Brasileiro (PB), uma vez que, este sistema é substituído pelo sistema binário tanto na língua falada quanto na língua escrita. Nesta, o contexto de produção textual no vestibular, exige que o candidato se adapte à norma, recorra ao uso do sistema ternário vinculado à pragmática, e conserve esse sistema, ainda que não seja condizente com a língua falada.

Castilho (2014, p. 498) afirma que “*decretar que o sistema ternário dessa subclasse tenha desaparecido é um pouco precipitado, pois na língua escrita, quando se configuram algumas necessidades dêiticas, esse sistema reaparece*”. Neste artigo, ao considerar que os processos seletivos exigem a adequação ao padrão formal, se o uso do sistema ternário, *este/esse/aquele* não ocorrer com frequência, pode-se ter indicadores de um sistema binário nas redações, se *esse/essa* se sobrepõe a *este/esta*, de acordo com Nascentes (1965, p. 3-5).

A LINGUÍSTICA DE CORPUS NO ENSINO DE LÍNGUAS

Sardinha (2004, p. 3) afirma que no século XX, o estudo do *corpus* era desenvolvido por linguistas, que coletavam os corpora ou “conjunto de dados linguísticos”, para descrever manualmente a língua e suas variações, bem como contribuir para o ensino de línguas. Atualmente, a descrição conta com sistemas computacionais que possibilitam a construção/coleta do *corpus* eletrônico, para a aplicação na fonologia, semântica, sintaxe, pragmática, nas áreas como “*ensino de língua estrangeira, análise de gênero, metáfora, tradução e linguística forense*” (KADER; RICHTER, 2013, p. 21).

Entretanto, para que o pesquisador utilize a Linguística Computacional, deve se adequar aos procedimentos necessários para a compilação do *corpus*

eletrônico, a partir da: limpeza do *corpus*⁵, isto é, a remoção de cabeçalhos, acentuações e outros elementos; nomeação e armazenamento em formatos de texto (.txt), dentre outros mecanismos que contribuem para a anotação automática e semiautomática, por meio da aplicação dos corpora nas ferramentas computacionais. Assim, considera-se que

Anotação linguística pode ser qualquer nível que se queira, isto é, nos níveis morfossintático, sintático, semântico, discursivo, etc, sendo inserida de três formas: manualmente (por linguistas), automaticamente (por ferramentas de Processamento de Língua Natural - PLN) ou semi-automaticamente (correção manual da saída de outras ferramentas) (ALUÍSIO; ALMEIDA, 2006, p. 161).

Sinclair (2005) afirma que, a princípio, as evidências do estudo de *corpora* eram desconsideradas, mas com a precisão de resultados oriundos da Estatística nos *softwares*, isto possibilitou que a Linguística de *Corpus* alcançasse uma notoriedade por seu caráter científico. Além disso, o estudo dos corpora dedica-se à compilação e análise de uma vasta quantidade de dados na língua escrita e/ou falada, a fim de propor “*representativos da totalidade do uso linguístico ou de algum de seus âmbitos*” (SANCHEZ, 1995, p. 8-9).

Portanto, neste estudo, com o propósito de descrever o padrão de uso dos demonstrativos, utiliza-se a colocação, descrita por Sardinha (2004, p. 40) como a “*associação entre itens lexicais ou entre o léxico e campos semânticos*”. Em consonância, Sheperd (2009, p. 154) define colocação como a “*associação sintagmática de itens lexicais com outros itens lexicais*”, seja à esquerda ou à direita do sintagma.

No entanto, o que é padronização lexical? E como a Linguística Computacional pode favorecer este estudo? De acordo com Porfirio; Bidarra (2004) a padronização ocorre quando um conjunto de unidades lexicais coexiste com outros itens lexicais e demonstram uma regularidade de uso. Oliveira (2004) ressalta que a frequência é determinante para identificar o padrão lexical nas estruturas sintagmáticas. E como veremos, o *WordSmith Tools* (Scott, 2008), *software* da Linguística de *Corpus*, dispõe deste recurso.

Desse modo, descrever o padrão de uso das categorias lexicais, com apoio na linguística computacional, é interessante para o ensino de língua portuguesa, à medida que disponibiliza amostras concretas para o sistema escolar, referente à realidade linguística dos alunos e contribui para o processo de ensino-aprendizagem.

⁵ A limpeza do *corpus* favorece o processamento computacional do texto escrito ou falado, nos *softwares* como *WordSmithTools* e *AntConc*.

METODOLOGIA

O CORPUS DA PESQUISA

Neste estudo, elenca-se 58 produções textuais para compor o *corpus* desta pesquisa, 25 da 3ª Etapa do PRISE, de 2016 e 33 do RICEX, de 2017. Este *corpus* é um recorte das 433 redações, que constituem o *Corpus* de Estudo da UEPA (CORUEPA), está no formato de texto (*txt*) para aplicação nas ferramentas computacionais do *WordSmith Tools* 5.0 (Scott, 2008).

DOS SUJEITOS DA PESQUISA

Os candidatos do Programa de Ingresso Seriado (PRISE) são, exclusivamente, concluintes do 3º ano do Ensino Médio, enquanto os do Processo Seletivo Especial dos cursos de Relações Internacionais e de Tecnologia em Comércio Exterior (PSE RICEX) são servidores públicos efetivos do Estado do Pará (54% das vagas) e os que concluíram o Ensino Médio (46% das vagas ofertadas).

Destaca-se que outras variáveis são relevantes para a análise, como: faixa etária, sexo, escolaridade, a fim de entender os fatores que motivaram o uso dos demonstrativos, entretanto, em razão do sigilo que o certame impõe por se tratar de processos seletivos, não foi possível ter acesso às variáveis dos candidatos, para a elaboração deste artigo.

Neste estudo, considera-se que os participantes do PRISE e RICEX, pertencentes a públicos-alvo distintos, trazem na produção escrita, um reflexo da realidade linguística que se constitui nos diversos âmbitos (educacional, familiar, social) em que estão inseridos. No sistema escolar público e privado, o processo de ensino-aprendizagem pode estar pautado na norma padrão dos livros didáticos e/ou nos usos da língua escrita/falada dos alunos.

DA COLETA E SISTEMATIZAÇÃO DOS DADOS

As ocorrências dos demonstrativos foram coletadas pelo *software WordSmith Tools* 5.0, que dispõe de ferramentas como *WordList* (lista de palavras) e *Concord* (concordanciador). A partir da coleta desses pronomes, listou-se a frequência, categorizou-se ou a configuração de uso e identificou-se ou o padrão de colocação (anteposição/ posposição), como se pode observar no quadro 3.

Quadro 3 – Da coleta de dados

WordList																																							
File	Edit	View	Compute	Settings	Windows	Help																																	
Word	Freq.	%																																					
O	2	8.55	<table border="1"> <thead> <tr> <th colspan="3">Search</th> </tr> <tr> <td>Search column</td> <td>Word</td> <td>OK</td> </tr> <tr> <td></td> <td>Freq.c</td> <td>Help</td> </tr> <tr> <td></td> <td>%</td> <td></td> </tr> <tr> <td>Search for</td> <td>*ISSO</td> <td>Cancel</td> </tr> <tr> <td colspan="3">ISSO</td> </tr> </thead> <tbody> <tr> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td></td> <td></td> </tr> </tbody> </table>				Search			Search column	Word	OK		Freq.c	Help		%		Search for	*ISSO	Cancel	ISSO																	
Search																																							
Search column	Word	OK																																					
	Freq.c	Help																																					
	%																																						
Search for	*ISSO	Cancel																																					
ISSO																																							
A	8	3.42																																					
SE	5	2.14																																					
PARA	4	1.71																																					
FAZ	3	1.28																																					
NOSSO	3	1.28																																					
ISSO	3	1.28																																					
NADA	2	0.85																																					

Fonte: Elaborado pelas autoras, com base no *corpus* retirado do CORUEPA e aplicado no *WordList* do programa *WordSmith Tools 5.0*.

No quadro (3) a ferramenta *WordList* apresenta três colunas centrais: *word*, *freq.* e *%*, de modo que: a coluna *word* lista os vocábulos que constituem a redação; a coluna *freq.* demonstra a frequência destes itens e a coluna *%* traz este percentual. O processo de anotação linguística dispõe da caixa de diálogo do *Search*, que serve para restringir os resultados da busca por esses vocábulos, mecanismo que se utilizou para a coleta de dados.

Em seguida, utiliza-se a ferramenta *Concord*, para extrair os excertos que contêm estes pronomes, categorizar a configuração de uso e identificar o padrão de colocação. Para exemplificar, há três sentenças de três candidatos do PRISE, em que o *Concord* identifica o pronome variável *essa*

(I) Entretanto, o estrago maior foi no mercado brasileiro, *essa* informação chegou sem a cautela que deveria na mídia.

(II) Ademais, as exportações *dessa* caíram 90%.

(III) A recente operação da Polícia Federal – chamada “Carne Fraca” (...) Segundo ele, casos *dessa* investigação, trazidos a público pela Polícia Federal e pelas mídias.

Nos exemplos (II), (III) e (IV), a posição dos demonstrativos adjetivos no sintagma nominal, SN + demonstrativo / demonstrativo + SN aponta para: (1) anteposição; (2) posposição e (3) anteposição. Esses exemplos mostram indicativos de um baixo grau de monitoramento, por conta do predomínio do demonstrativo anteposto.

ANÁLISE DE DADOS

A análise da frequência dos demonstrativos registra 204 ocorrências nas 58 redações, com 160 formas variáveis e 44 invariáveis, destaca-se que o uso desses pronomes não ocorre em 04 redações. A coleta da frequência dos demonstrativos realiza-se mediante a aplicação da ferramenta *WordList*, do programa *WordSmith Tools 5.0*.

Tabela 1 – Ocorrência dos demonstrativos nas redações

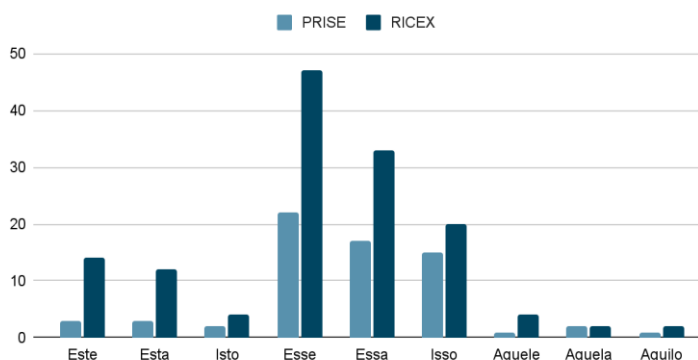
<i>Word</i>	<i>Freq.</i>	%
Esse(s)	69	33.82
Essa(s)	50	24.50
Isso	35	17.15
Este(s)	17	8.33
Esta(s)	15	7.35
Isto	6	2.94
Aquele(s)	5	2.45
Aquela(s)	4	1.96
Aquilo	3	1.47
Total	204	100%

Fonte: Elaborado pelas autoras, a partir da ferramenta *WordList*

Na tabela (1) observa-se que os pronomes *esse/essa/isso* da 2ª pessoa do discurso, apresentam uma frequência elevada em relação aos demais demonstrativos, enquanto *aquele/aquela/aquilo* da 3ª pessoa, demonstram uma frequência inferior.

Bagno (2012) discute que o sistema binário *esse/aquele* é predominante na língua falada, por outro lado, na escrita das redações do PRISE e RICEX, o sistema binário que prevalece é *esse/este* e o sistema ternário só não desaparece, por compor uma das demandas dos processos seletivos. A seguir, o gráfico (1) indica a frequência dos demonstrativos em cada processo seletivo.

Gráfico 1 – Frequência dos demonstrativos no PRISE e RICEX



Fonte: Elaborado pelas autoras, a partir da ferramenta *WordList*.

No gráfico (1) apresenta-se 08 ocorrências no PRISE referente às formas na 1ª pessoa do discurso, *este* possui uma frequência equivalente a *esta*, mas inferior a *isto*. No RICEX, temos 30 ocorrências, *este* tem frequência maior que *esta* e *isto*.

Na 2ª pessoa do discurso observa-se 54 ocorrências no PRISE e 100 ocorrências no RICEX. Em ambos os processos seletivos, *esse* indica maior frequência em comparação a *essa* e a *isso*. Na 3ª pessoa do discurso apresenta-se 04 ocorrências no PRISE, *aquela* aponta uma frequência maior que *aquele* e *aquilo*, enquanto no RICEX tem-se 08 ocorrências, e *aquele* se mostra superior à frequência de *aquela* e *aquilo*.

Segundo o gráfico (1), nota-se que os candidatos do RICEX foram os que mais usaram os demonstrativos nas três pessoas do discurso, com o quantitativo de 138 ocorrências em relação às 66 do PRISE. Entende-se que o sistema ternário *este/esse/aquele* ainda consta na realidade linguística escolar e nas produções textuais dos candidatos de 2016 – 2017. No entanto, é o sistema binário *esse/este*, que predomina na língua escrita dos participantes, com ênfase na forma da 2ª pessoa *essa/esse/isso*.

Na tabela (2) subdivide-se o uso dos demonstrativos em três categorias, na primeira (A), o sistema binário *este/esse*; na segunda (B) o sistema *esse/aquele*, que se aproxima da língua falada; na terceira (C), o sistema ternário *este/esse/aquele* e o binário *este/aquele*, associado a gramática prescritiva; na quarta (D), os candidatos que optaram por usos peculiares, *essa, aquela* ou *esta*.

Tabela 2 - Configuração de uso dos demonstrativos nos processos seletivos

PRISE		Nº	RICEX	Nº
A	Isso/ esta	1	Esse/ esta	3
	Isto/ esse	1	Essa/ este	1
	Este/ essa/ isso	1	Esta/ essa	1
	Este/ esta/ isto/ isso	1	Este/esse/essa/isso/essa/esta/este	1
B	Aquela/ esse	3	Aquela/ essa/ esse/ isso/ aquele	1
	Aquilo/ esse	1	Essa/ isso/ esse/ aquela	1
			Isso/ aquele/ esse	1
			Esse/ essa/ aquilo	1
C	Este/ esse/ aquele	1	Este/ isto/ aquele	2
D	Aquela	1	Esta/ este/ isto	5
	Essa/ esse/ isso	1	Essa/ esse/ isso	1
		4		3
Nº: o número de candidatos, que usaram os respectivos demonstrativos no PRISE e RICEX.				

Fonte: Elaborado pelas autoras, a partir da ferramenta *Concord*.

Observa-se que a categoria (D) prevalece no PRISE e RICEX, visto que 27 dos candidatos preferem usar apenas os pronomes *essa/esse/isso* em suas redações, enquanto 10 participantes utilizam somente os pronomes *essa/esta*, 08 optam por *aquela/essa*, 05 por *este/esta/isto*, 02 por *este/aquela*, 01 por *este/esse/aquela* e 01 por *aquela*.

Segundo as amostras, se pode observar que o sistema ternário está desaparecendo da língua escrita dos indivíduos, até mesmo no contexto do vestibular, que demanda o uso deste sistema. Nas redações da UEPA, nota-se que a baixa frequência do pronome *aquela* e do sistema ternário, ocasiona na predominância dos pronomes *esse/essa/isso* – 14 no PRISE e 13 no RICEX; e do sistema binário *essa/esta*, com 04 no PRISE e 06 no RICEX.

Elenca-se a categoria (C), associada ao uso da gramática normativa, com o sistema ternário *este/esse/aquela*, para exemplificar se o uso dos demonstrativos realmente segue à norma padrão, critério de correção da UEPA, ou se predomina a linguagem coloquial do participante. Ilustra-se o uso ternário *este/esse/aquela* pertencente ao informante A do PRISE

(V) Não obstante, tem-se conhecimento que *estes* quatro elementos, terra,

fogo, ar e água são imprescindíveis em diversos aspectos (...)

(VI) Entretanto, na desenfreada disputa das potências mundiais, *esses* elementos tornam-se vilões (...)

(VII) E mais recente, tem-se o caso das Usinas em Minas Gerais, cujo uso inadequado da superfície terrestre *naquele* local (...)

Em (V) e (VI) observa-se que a dêixis temporal dos pronomes “este” e “esse” corresponde à norma por indicar tempo presente, enquanto na dêixis espacial, ambos manifestam proximidade em relação ao leitor. Em (VII) a dêixis espacial do demonstrativo “naquele” refere-se a um distanciamento entre o escritor e o leitor, já a dêixis temporal indica passado vago ou remoto. Nota-se adequação do candidato à prescrição gramatical, mediante o uso dos verbos com a partícula ‘se’ e dos demonstrativos conforme a dêixis.

Conclui-se que a categoria (C) pertencente ao sistema ternário *esse/este/aquele* retrata o uso da língua padrão. Por outro lado, o predomínio da categoria (D) uso exclusivo do pronome *essa/esse/isso*, seguido da categoria (A) sistema binário *esse/este*, revela uma realidade linguística divergente do que preconiza as gramáticas tradicionais e os livros didáticos, diante do desuso da categoria (C) – conforme a tabela (2).

Enfatiza-se que 3,98% dos pronomes *essa/esse/isso* corresponde à língua padrão, enquanto 2,96% não segue a norma, o mesmo ocorre no sistema binário *esse/este*, visto que 3,9% refletem o ensino gramatical nas escolas e 0,6% referem-se ao português não padrão, incluso na tradição oral. Logo, apesar de 01 candidato utilizar o sistema ternário, a análise das categorias (A), (C) e (D) infere que a maioria dos candidatos do PRISE e do RICEX segue o critério normativo estabelecido pela UEPA, porém há um percentual de uso não padrão, que requer atenção do ensino para ser desenvolvido nas aulas de língua portuguesa.

A seguir, discute-se sobre o padrão de colocação dos demonstrativos adjetivos nos processos seletivos da UEPA, para entender se houve uma preocupação dos candidatos em pospor esses pronomes aos substantivos, e assim, monitorar seus usos linguísticos.

Tabela 3 – Ocorrência dos demonstrativos adjetivos no PRISE

	Pronome	Freq.	Anteposição	Posposição
PRISE	Esse	22	19	1
	Essa	17	17	-
	Este	3	3	-
	Esta	3	2	1
	Aquele	1	1	-
	Aquela	2	2	-
	Total	48	44	2

Fonte: Elaborado pelas autoras, a partir da ferramenta *Concord*

Na tabela (3) observa-se 48 demonstrativos variáveis, 46 exercem a função de adjetivos e 02 de substantivos. No PRISE dentre os 46 demonstrativos adjetivos, registram-se 44 antepostos, com destaque para *esse/essa*, que obteve maior frequência e 02 pospostos somente com *esse/esta*.

Tabela 4 – Dos demonstrativos adjetivos no RICEX

	Pronome	Freq.	Anteposição	Posposição
RICEX	Esse	47	41	1
	Essa	33	27	4
	Este	14	12	-
	Esta	12	11	-
	Aquele	4	0	-
	Aquela	2	2	-
	Total	112	93	5

Fonte: Elaborado pelas autoras, a partir da ferramenta *Concord*.

Na tabela (4), apresenta-se 112 demonstrativos variáveis, 98 exercem a função de adjetivos e 14 de substantivos. No RICEX, dentre os 98 demonstrativos adjetivos, registra-se 93 antepostos, ênfase para *esse/essa* e 05 pospostos, somente com *esse/essa*.

A descrição do padrão de colocação dos pronomes nos processos seletivos revela que os demonstrativos adjetivos são antepostos com maior frequência tanto no PRISE como no RICEX, logo a produção escrita de ambos não corresponde a usos monitorados. Por outro lado, identifica-se que o RICEX contém uma posposição maior que o PRISE, contudo, os usos não são suficientes

para configurar uma escrita monitorada diante do padrão da anteposição.

Em suma, evidencia-se que a frequência dos demonstrativos é alta nas formas *essa* e *esse/este* dos processos seletivos. A configuração de uso comprova o elevado índice de ocorrências para o emprego exclusivo do pronome *essa/esse/isso* com 58,3% nas redações do PRISE e 43,3% do RICEX, assim como do sistema binário *esse/este* com 16,6% do PRISE e 20% do RICEX. Para somar a esses indicadores da realidade linguística, acrescentam-se como os pronomes foram colocados na sentença e conclui-se que o uso à esquerda do sintagma é predominante.

RESULTADOS ALCANÇADOS

Neste estudo pretendeu-se mostrar a realidade linguística dos participantes dos processos seletivos de 2016 e 2017 da UEPA, por apresentar o uso dos demonstrativos: se expõe uma configuração ternária vinculada à gramática normativa, ou se denota usos binários. Dentre os resultados observados, aponta-se: (i) o uso apoiado na língua falada *esse/aquele*; (ii) a variante *esse/este* apoiada na gramática normativa; (iii) os que não usaram o sistema binário nem o terciário, apenas o pronome *esse/essa/isso* – a maioria; e (iv) os que não aplicaram nenhum demonstrativo (in)variável nas redações.

Em (i), observa-se que os candidatos possuem uma dificuldade em diferenciar *esse* de *este* ou uma preferência de uso pela 1ª e 3ª pessoa do discurso, em ambos os casos, a escolha deve estar vinculada ao modo como usam esses pronomes na língua falada. Em (ii), os candidatos mostram um domínio para diferenciar as formas *esse/este*, visto que o uso deste sistema binário remete a uma inserção no padrão formal, os participantes não usam o pronome *aquele*, por associar a uma linguagem informal e/ou desconhecer como utilizá-lo.

No (iii), têm-se os participantes que preferem usar *esse/essa/isso*, as formas na 2ª pessoa a fim de evitar equívocos na escrita das redações e/ou os que desconhecem, esquecem como usar os pronomes *este/aquele*. No (iv), os participantes demonstram um desconhecimento referente ao uso dos demonstrativos (in)variáveis e/ou preferem não usar esses pronomes nas redações.

Destaca-se, que apenas um candidato usa o sistema ternário *este/esse/aquele*, portanto, além deste sistema estar em desuso na língua escrita dos participantes do PRISE e do RICEX, também não é encontrado na língua falada do português brasileiro, de acordo com linguistas; ressalta-se que o pronome *aquele* também está desaparecendo da escrita.

Na literatura, os linguistas defendem que a forma *esta* é substituída por *essa*, pois nas situações conversacionais, o sujeito não diferencia *este* de *esse*, apenas usa *esse* para se referir a algo próximo e *aquele* para algo distante. Neste artigo observa-se que os candidatos se preocupam em distinguir *este* de *esse*, pois a

maioria se adéqua a norma. Entretanto, é importante que o ensino se preocupe com o percentual que não alcança o critério da UEPA de adequação à norma, pois revela a necessidade de o ensino relacionar o português padrão com o não padrão e reconhecer a realidade linguística dos alunos nas escolas.

Na colocação dos demonstrativos (variáveis) adjetivos, o pronome pode vir à direita ou à esquerda do sintagma, quando anteposto é pouco monitorado e quando posposto possui um alto grau de monitoramento. O padrão de colocação aponta para a anteposição, em virtude de a posposição ser um recurso estilístico, pouco usado pelos candidatos, seja em razão de desconhecimento ou preferência linguística.

O estudo da configuração de uso dos pronomes propõe auxiliar na atuação dos profissionais da educação, pois permite descrever como o aluno se relaciona com o tema, como utiliza os elementos gramaticais. Assim, para romper com a prática padronizadora, propõe-se que as escolas apliquem como ferramenta de ensino a realidade linguística dos alunos, construída dentro/fora do âmbito escolar, para visibilizar os saberes dos sujeitos, suas linguagens em prol da educação intercultural.

Desse modo, recomenda-se que o docente use “amostras concretas” para facilitar o entendimento dos alunos, acerca do tema sobre os demonstrativos, ao dialogar como esses pronomes aparecem na escrita; se é possível identificar outros sistemas binários, haja vista que, não existe apenas o sistema *esse/aquele*, apoiado na língua falada; se usam demonstrativos pospostos ou antepostos; e se o uso apresenta alguma semelhança com o sistema ternário.

CONSIDERAÇÕES FINAIS

A realidade linguística existente nas redações atua como ferramenta para o processo de ensino-aprendizagem do português brasileiro, por incentivar a prática docente a reconhecer a interculturalidade dos alunos e como estão utilizando os demonstrativos nas produções textuais. É interessante repensar a aplicação de exercícios estruturais no âmbito escolar, pois a preeminência da norma padrão intimida as manifestações linguísticas presente na sala de aula.

Como vimos, a literatura defende que o sistema binário *esse/aquele* é preeminente na língua falada em detrimento do pronome *este*. Por outro lado, na língua escrita conclui-se que o pronome *aquele* está desaparecendo da língua escrita e observa-se o desuso do sistema ternário *este/esse/aquele* nas redações, diante do predomínio do pronome *esse/essa/isso* e do sistema binário *este/esse*, descrito a partir das ferramentas computacionais. O resultado das amostras revela uma preocupação dos participantes em usar os pronomes nas redações, prova disso, é o elevado percentual em conformidade a gramática normativa, ainda que o padrão lexical aponte o uso prevalecente da colocação anteposta, indicativo de baixo monitoramento.

Espera-se com este estudo, contribuir para o processo de ensino-aprendizagem da língua portuguesa, ao considerar o padrão de uso dos candidatos da UEPA e propor que os docentes imersos na sala de aula apliquem como ferramenta de ensino a realidade linguística dos alunos, construída dentro/fora da escola.

Admite-se que o uso natural da língua predomina fora do âmbito escolar, pois o indivíduo se relaciona com diversas comunidades linguísticas, logo expor as realidades linguísticas na sala de aula a partir de gêneros textuais, da produção escrita do aluno, visa contextualizar os usos da língua materna. Reforça-se a necessidade de ampliar pesquisas nesta área, para descrever se o uso efetuado por aprendizes nas redações se ampara ou não, na norma-padrão contemplada por gramáticas normativas e requisitada por vestibulares.

REFERÊNCIAS

- ALMEIDA, Napoleão. *Gramática Latina: curso único e completo*. SP: Saraiva, 2000.
- ALUÍSIO, S; ALMEIDA, G. O que é e como se constrói um *Corpus*? Lições aprendidas na compilação de vários *corpora* para pesquisa linguística. *Rev. Calidoscópico*, RS, 2006.
- BAGNO, M. *A Língua de Eulália: novela sociolinguística*. 4ª ed. SP: Contexto, 1999.
- BAGNO, M. *Dramática da Língua Portuguesa: Tradição gramatical, mídia & exclusão social*. SP: Edições Loyola, 2001.
- BAGNO, M. *A Norma Oculta: Língua & poder na sociedade brasileira*. SP: Parábola, 2009.
- BAGNO, M. *Gramática Pedagógica do Português Brasileiro*. SP: Parábola, 2012.
- BECHARA, Evanildo. *Moderna Gramática Portuguesa*. RJ: Nova Fronteira, 2009.
- CASTILHO, A. T. de. Os mostrativos no português falado. In: CASTILHO, A. T. de. (org.). *Gramática do Português Falado*, vol. III. Campinas: Unicamp, 1993.
- CASTILHO, A. T. de. *Nova Gramática do Português Brasileiro*. SP: Contexto, 2014.
- CUNHA, Celso & CINTRA, L. F. L. *Nova Gramática do Português Contemporâneo*. 2. ed. Rio de Janeiro: Nova Fronteira, 2017.
- FARACO, C. *Norma Culta Brasileira: desatando alguns nós*. SP: Parábola Editorial, 2008.
- HAUY, Amini Boainain. *Gramática da Língua Portuguesa Padrão*. 1 ed. SP: Edusp, 2015.

- HOLANDA, S.B. de. *Raízes do Brasil*. 27ª ed. SP: Companhia das Letras, 2014.
- KADER, C; RICHTER, M. Linguística de corpus: possibilidades e avanços. *Revista de Estudo e Pesquisa em Educação*. Juiz de Fora, v.15, jan./jun. 2013.
- NASCENTES, Clóvis Monteiro. “Êste”, “êsse”. *Miscelânea*. Rio de Janeiro: Editora do Professor, 1965.
- OLIVEIRA, Aparecida A. de. *Análise de textos da prova de inglês do vestibular da FUVEST: Um estudo baseado em Corpus*. 2004. Dissertação de Mestrado, PUC/SP, Ribeirão Preto, 2004.
- PORFIRIO, Lucielen; BIDARRA, Jorge. *Padronização Lexical: Um estudo descritivo inicial sobre colocação, coligação e prosódia semântica aplicados à extração automática de informação de textos*. In: Congresso Nacional de Linguística e Filologia, n. 1, série VIII, 2004, RJ, Cadernos do CNFL, RJ: Instituto de Letras. UERJ, 2004.
- POSSENTI, Sírio. *Por que (não) ensinar gramática na escola*. SP: Mercado de Letras, Associação de Leitura do Brasil, 2000.
- ROCHA LIMA, C. *Gramática Normativa da Língua Portuguesa*. RJ: J. Olympio, 2011.
- SANCHEZ, A. Definicion e Historia de los Corpus. In: A. SANCHEZ et al (org.). *CUMBRE – Corpus Linguístico de Espanol Contemporâneo*. Madrid: SGEL, 1995.
- SARDINHA, T.B. *Linguística de Corpus*. SP: Editora Manole, 2004.
- SCOTT, Mike. *WordSmith Tools. Versão 5.0*. Reino Unido: Lexical Analysis Software Ltd, 2008. Disponível em: <<https://www.lexically.net/wordsmith/>>. Acesso em: 01 jun. 2022.
- SHEPERD, Tania M.G. O Estatuto da Linguística de *Corpus*: metodologia ou área da linguística? *Revista Matraca*, RJ, v.16, n.24, jan./jun. 2009.
- SINCLAIR, J. *Corpus and Text - Basic Principles. Developing Linguistic Corpora: a Guide to Good Practice*, ed. M. Wynne. Oxford: Oxbow Books. 2005.



The Lexical Standard of (in)variable demonstratives in the corpus CORUEPA: the linguistic reality as a teaching-learning tool

Dóris Figueiredo de Souza¹
Maria Sílvia Jucá Acácio²

Abstract

This study aims to analyze how PRIZE and RICEX candidates used (in)variable statements in the entrance exams at the State University of Pará (UEPA); in order to encourage the teacher to make use of language varieties, encourage student involvement in Portuguese language classes and question whether the grammar taught at school values the identity construction of subjects in their respective language communities. The study aims to list the frequency of statements in CORUEPA's textual productions, categorize the configuration of use in essays and identify the placement pattern of variable statements in PRIZE and RICEX. The study is based on Sardinha (2004), Bagno (2001; 2012) and Castilho (2014). The methodology applies the corpus in the analysis tools of the WordSmith Tools 5.0 program (Scott, 2008). The result of the samples reveals the predominance of the use of the prefixed placement and the disuse of the ternary system, given the predominance of the pronoun those/these/that and the binary system this/these. This study is expected to contribute to the teaching-learning process of the Portuguese language, by considering the pattern of use of UEPA candidates and proposing that teachers apply the linguistic reality of students as a teaching tool, built inside/outside the school.

Keywords:

Corpus Linguistics; Portuguese Language; Linguistic Description
