

UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO
CENTRO DE EDUCAÇÃO E HUMANIDADES
INSTITUTO DE LETRAS
COORDENAÇÃO DE PÓS-GRADUAÇÃO EM LETRAS

AS FORMAÇÕES *X-INHO* NAS MODALIDADES ORAL E ESCRITA: UM ESTUDO
CONTRASTIVO BASEADO NA LINGÜÍSTICA DE CORPUS

MARCOS ANTÔNIO GONÇALVES

Rio de Janeiro
2006

MARCOS ANTÔNIO GONÇALVES

AS FORMAÇÕES *X-INHO* NAS MODALIDADES ORAL E ESCRITA: UM ESTUDO
CONTRASTIVO BASEADO NA LINGÜÍSTICA DE CORPUS

Dissertação apresentada à Coordenação Geral Stricto
Sensu de Pós-Graduação em Letras da Universidade
do Estado do Rio de Janeiro, como exigência parcial
para a obtenção do título de Mestre em Lingüística.
Orientador: Professora Doutora Tânia Maria Granja
Shepherd.

Rio de Janeiro

2006

CATALOGAÇÃO NA FONTE
UERJ/REDE SIRIUS/NPROTEC

G635 Gonçalves, Marcos Antônio.
As formações x-inho nas modalidades oral e escrita : um estudo contrastivo baseado na lingüística de corpus / por Marcos Antônio Gonçalves. – 2006.
92f.

Orientadora : Tânia Maria Granja Shepherd
Dissertação (Mestrado) – Universidade do Estado do Rio de Janeiro, Instituto de Letras

1.Língua portuguesa – Sufixos e prefixos – Teses. 2. Análise lingüística (Linguística) – Teses. 3. Lingüística – Processamento de dados – Teses. I. Shepherd, Tânia Maria Granja. II. Universidade do Estado do Rio de Janeiro. Instituto de Letras. III. Título.

CDU 806.90-541.112.12

EXAME DE DISSERTAÇÃO

GONÇALVES, A. Marcos. *As formações x-inho nas modalidades oral e escrita: uma análise contrastiva baseada na Lingüística de Corpus*. Dissertação de Mestrado em Lingüística. Rio de Janeiro, Coordenação de Pós-graduação em Letras, UERJ, 2006. 100 p.

Orientador: Professora Doutora Tânia Maria Granja Shepherd (UERJ)

Professora Doutora Sandra Pereira Bernardo (PUC-Rio)

Professor Doutor Helênio Fonseca de Oliveira (UERJ)

Defendida a Dissertação

Conceito:

Em: ____/____/____

“Tudo se transforma.

No caso das palavras, próprias da natureza humana, tudo se transforma realmente, de acordo com a latitude, a hora, o momento, o ambiente, a circunstância, a inflexão, o gesto, o tom, mil e um fatores, mil e uma nuances.”

Pedro Bloch

AGRADECIMENTOS

A Deus por me permitir, com saúde e força, concluir esta etapa almejada.

A meus pais pelo exemplo de vida e fé, pelo amor e pela responsabilidade na minha educação, formação acadêmica e, principalmente, humana.

À minha filha pela compreensão e paciência.

Aos amigos que fiz durante esses anos na UERJ, meu agradecimento.

À professora Doutora Tânia Shepherd, pela competência, disponibilidade, atenção e segurança tão importantes na orientação deste trabalho.

À professora Doutora Sandra Pereira Bernardo pelo apoio deste os tempos de meu curso de Especialização.

A todos que, de algum modo, viabilizaram esta pesquisa, o meu agradecimento.

SINOPSE

Estudo contrastivo do uso subjetivo das formações *x-inho* nas modalidades oral e escrita, com base em suas funções semântico-pragmáticas. Uma abordagem baseada na área de investigação denominada Linguística de Corpus.

SUMÁRIO

1. INTRODUÇÃO.....	12
2. REFERENCIAL TEÓRICO.....	17
2.1 – Visão geral das formações x-<i>inho</i>	17
2.1.1 – Aspectos morfológicos do sufixo <i>-inho</i>	23
2.1.2 – Lexicalização ou idiomatização	27
2.1.3 – Aspectos estilísticos	28
2.1.4 – Aspectos lexicográficos	29
2.1.5 – Estrutura sintática	30
2.1.6 – Os traços pejorativo e afetivo	30
2.2 – O sufixo <i>-inho</i> visto como grau	31
2.2.1 – Algumas diferenças entre flexão e derivação.....	32
2.3 – Os Modos de Organização do Discurso	36
2.3.1 – Modo de Organização Descritivo.....	38
2.3.2 – Modo de Organização Narrativo.....	39
2.3.3 – O Narrativo/Descritivo	40
2.4 – Oposição língua falada/língua escrita	41
2.5 – A noção de contexto	48
2.5.1 – O Contexto de Cultura e o Contexto de Situação.....	50
3. PROCEDIMENTOS METODOLÓGICOS.....	54
3.1- A Lingüística de Corpus	54

3.1.1 – Breve histórico.....	54
3.1.2 – Conceitos centrais.....	56
3.1.3 – Metodologia ou abordagem.....	60
3.1.4 – Corpus-driven e corpus-based.....	61
3.1.4.1 – Abordagem conduzida por corpus (corpus-driven).....	61
3.1.4.2 – Abordagem baseada em corpus (corpus-based).....	62
3.2- Tamanho e representatividade de um <i>corpus</i>	63
3.3- Seleção e descrição dos <i>corpora</i>	64
3.4- O programa WordSmith Tools	67
3.5- Passo a passo da extração e análise	70
4. ANÁLISE E INTERPRETAÇÃO.....	72
4.1- Contornos estruturais para identificação de <i>x-inho</i>	72
4.1.1 – A função diminutiva.....	73
4.1.2 – A função afetiva positiva.....	74
4.1.3 – A função afetiva negativa.....	76
4.1.4 – A função intensificadora.....	77
4.2- Interpretação dos dados	81
5. CONSIDERAÇÕES FINAIS.....	92
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	95
ANEXO 1 – Lista de concordâncias do <i>corpus</i> escrito.....	
ANEXO 2 – Lista de concordâncias do <i>corpus</i> oral narrativo.....	
ANEXO 3 – Lista de concordâncias do <i>corpus</i> oral descritivo.....	

LISTA DE QUADROS

QUADRO 1	Grau de intensificação de bases negativas e positivas.....	21
QUADRO 2	Grau de dimensão em bases substantivas.....	22
QUADRO 3	Eixos constitutivos da linguagem	26
QUADRO 4	Funções desempenhadas pelas três ferramentas do programa WordSmith Tools.....	69

LISTA DE TABELAS

TABELA 1	Classificação relativa do tamanho de <i>corpus</i>	63
TABELA 2	Quantidade de ocorrências e formas lexicais nos diferentes subcorpora.....	67
TABELA 3	Total de itens e frequências das formações <i>x-inho</i> nos <i>corpora</i>	82
TABELA 4	Ocorrência de <i>x -inho</i> nos <i>corpus</i> oral por modo.....	82
TABELA 5	Ocorrências da noção dimensiva no modo narrativo.....	84
TABELA 6	Frequência total do <i>corpus</i> escrito.....	88
TABELA 7	Ocorrência de <i>x-inho</i> no <i>corpus</i> escrito.....	89

LISTA DE GRÁFICOS

GRÁFICO 1	Noções expressas por x-inho no <i>corpus</i> oral narrativo	83
GRÁFICO 2	Noções expressas por x-inho no <i>corpus</i> oral descritivo	86
GRÁFICO 3	Noções expressas por x-inho no <i>corpus</i> escrito	89

LISTA DE FIGURAS

FIGURA 1	Tela do programa WordSmith Tools mostrando linhas de concordâncias de palavras terminadas em <i>-inho</i>	68
----------	---	----

CAPÍTULO 1 - INTRODUÇÃO

A proposta deste trabalho é desenvolver um estudo sobre as formações *x-inho*, em que *x* representa qualquer base compatível com elementos mórficos *-inho* e *-zinho*, considerados, neste trabalho, formalmente equivalentes.

A motivação inicial desta pesquisa adveio da percepção de que essas formações não assinalam apenas o grau diminutivo sintético dos substantivos e nem somente emotividade, quer positiva ou negativa. O sufixo *-inho* tem a função de indicar uma dimensão menor, melhor ou pior daquilo que é considerado como padrão normal, e desta forma apresenta também uma atitude avaliativa do enunciador em relação ao objeto por ele dimensionado. Conforme Lapa (1998:93, *apud* Emílio, 2003:26) “É nos sufixos que a descarga das paixões se dá com maior energia”. Assim, consideramos que o uso das formações *x-inho* não se restringe ao significado-base, ou ao acréscimo da noção de tamanho reduzido ou daquilo que o objeto focado possua de melhor ou pior, mas pode apresentar as percepções do enunciador dentro de um contexto específico.

Nesse sentido, entendemos (cf. Halliday, 1985) que é na situação concreta de comunicação que esse enunciador se identifica e se alinha (ou não) ao ouvinte, não só através da significação social e da adequação de registros: mediante a intenção, o falante escolhe a estrutura lingüística de acordo com seu ouvinte e a situação, e também se identifica pela linguagem, algo que Halliday (*op.cit.*) já descreveu como as macrofunções ideacional e interpessoal da linguagem. Na realidade, o uso de *-inho* não é de caráter somente expressivo, nem somente dimensivo, mas significativamente mais expressivo. O falante, ao fazer sua escolha por *-inho*, na maioria das vezes, opta pelo sufixo por ser um recurso matizado que permite convergir sentidos sobre um significado base.

Sabemos, portanto, que a amplitude de uso e produtividade¹ dessas formações determinam uma dificuldade também na sua análise teórica. Porém, essas formações atendem às necessidades comunicativas, dentro de uma determinada função discursiva. A partir dessa perspectiva, esta dissertação buscará levantar a produtividade das formações *x-inho* através da investigação de amostras de dois *corpora* eletrônicos, um de natureza oral e outro de

¹ Definimos neste trabalho o termo produtividade como sendo a habilidade de um falante produzir novas formas, sendo equivalente à frequência, ou frequência relativa de uma forma.

natureza escrita². O *corpus* oral foi subdividido em dois *subcorpora*, englobando textos preponderantemente narrativos e descritivos, respectivamente, segundo a classificação de Charaudeau (1992). A finalidade do trabalho, portanto, é detectar, entre as possibilidades de uso dessas formações, qual a função mais freqüente e produtiva entre os *corpora* escrito e oral e, nesse último, qual o modo de discurso (narrativo ou descritivo) apresenta mais opções pela forma *x-inho*.

Tendo em vista que até o presente não há trabalhos sobre a formação *x-inho* baseados em *corpora* eletrônicos, a pesquisa relatada aqui enfatiza que são os exemplos advindos do *corpus*, sua distribuição e tipologia que formam a base da investigação. Para tanto, foram adotadas algumas noções centrais à área da Lingüística, conhecida como Lingüística de Corpus³ (Biber et al., 1998; Sinclair, 1991 e Berber Sardinha, 2004).

A Lingüística de Corpus é uma área que “se ocupa da coleta e exploração de *corpora*, ou conjuntos de dados lingüísticos textuais, em formato legível por computador, que foram coletados criteriosamente com propósito de servirem para a pesquisa de uma língua ou variedade lingüística” (Berber Sardinha, 2004: 3). Tendo por base a conceituação Hallidayana (Halliday, 1985) da língua como um sistema probabilístico, a Lingüística de Corpus entende que as ocorrências lexicais se dão dentro de uma certa freqüência e distribuição dentro de qualquer língua, combinando-se e associando-se a outras ocorrências que definirão suas funções e seu uso. Enquanto sistema probabilístico, portanto, não é necessário analisarem-se todas as instâncias de determinado padrão lingüístico para se poder calcular a probabilidade de esse padrão ocorrer numa língua como um todo. Uma amostra representativa da variedade que se quer estudar pode servir para dimensionar como o léxico se comporta e como se distribui dentro da língua.

Desse modo, não incluímos no trabalho as ocorrências das formações *x-inha*, *x-inhas*, *x-inhos*, por considerá-las semanticamente equivalentes a *x-inho* e, probabilisticamente passíveis de ocorrerem com freqüência e distribuição semelhantes. Além da flexão de gênero, portanto, deixamos também de lado a flexão de número dessa formação, à exceção de itens lexicalizados em expressões polilexicais, que veremos no capítulo 2.

² A composição dos *corpora* será explicitada no capítulo 3 deste trabalho.

³ O capítulo 3 desta dissertação apresenta uma explicação mais abrangente do que se entende por Lingüística de Corpus.

Neste trabalho pretendemos analisar, portanto, os aspectos relevantes das formações *x-inho* na língua portuguesa partindo de uma abordagem funcional⁴, mais abrangente do que as explicitadas nas gramáticas tradicionais que enfocaremos mais adiante. O trabalho ressalta, a importância da Lingüística de Corpus na descrição dos fenômenos lingüísticos, através da análise dessas formações a partir de *corpora* eletrônicos de onde se extraem as regularidades manifestas.

Dentre os trabalhos já realizados sobre as formações *x-inho*, podemos citar Ezarani (1989) que teve como objetivo examinar os contextos lingüísticos e extralingüísticos de ocorrências dessas formações na fala espontânea do português carioca. Seus objetivos principais eram tentar identificar as funções semântico-pragmáticas de *x-inho* e avaliar quantitativa e probabilisticamente no *corpus* pesquisado a correlação dessas formações a fatores lingüísticos, sociais e situacionais. A autora utilizou o *corpus* do Projeto de Estudo da Norma Urbana Culta (NURC) e o *corpus* do Projeto Censo da Variação Lingüística, ambos orais. Outro trabalho é o de Cláudio Moreno (1977) que estuda os diminutivos em *inho* e *zinho* e a delimitação do vocábulo nominal em português, fundamentando-se nas colocações de Matoso Câmara e Yonne Leite, as quais teve a pretensão de corroborar ou ampliar. Podemos citar também o trabalho de Emílio (2003) que estuda o fenômeno estilístico - uso do diminutivo x grau normal - à luz da análise variacionista. Há, ainda, os trabalhos de Piza (2001) e Rosa (1982) que abordam em suas pesquisas algumas questões do grau e, em contrapartida, do sufixo *-inho*.

Embora haja uma série de publicações e de trabalhos que se baseiam em *corpus*, não se encontrou nenhuma pesquisa que utilizasse a metodologia da Lingüística de *Corpus* para investigar as formações *x-inho*, por isso, torna-se relevante nossa pesquisa pelas contribuições e inovações que possam vir a trazer para o estudo dessas formações na língua portuguesa.

Nossa hipótese inicial é de que fatores como as influências lingüísticas de um grupo, os fatores extralingüísticos (sexo, faixa etária, nível intelectual etc.) e o tipo de modalidade (oral e escrita) podem *incorporar* determinadas noções a algumas palavras do léxico da língua, dando, às vezes, a um determinado paradigma um outro sentido fora daquele que é de uso esperado ou que esteja previsto na gramática. Pode-se incluir nessa hipótese o alto índice de

⁴ Definimos 'funcional' como "uma abordagem lingüística direcionada para o uso da língua, cujo interesse está centrado na língua como instrumento de interação social e não apenas como um sistema isolado." (Neves, 1997: 56).

produtividade das formações *x-inho* fora da sua noção tradicional (representar tamanho pequeno), as quais não são avaliadas de maneira mais explícita em alguns compêndios gramaticais. Observa-se, portanto, que as abordagens tradicionais prendem-se com mais ênfase aos valores de grau⁵ (tamanho pequeno) dessa formação, não enfocando de maneira mais exaustiva a subjetividade do mesmo, que tem um nível de produtividade alto e pode concorrer com o diminutivo.

Baseando-se na hipótese inicial acima, a presente dissertação faz as seguintes perguntas de pesquisa:

- a) Ao investigarmos quantitativamente um *corpus* eletrônico de natureza oral e outro de natureza escrita qual a frequência do sufixo *-inho* nesses *corpora*?
 - Que tipo de funções o sufixo desempenha e como se distribuem?
 - Como se comporta o sufixo *-inho* no corpus oral, se o dividirmos em dois Modos de Organização do Discurso distintos (narrativo e descritivo)?
- b) Como se comporta o sufixo *-inho* contrastivamente em termos de frequência e distribuição de funções nos dois *corpora*?

O tema proposto e as perguntas de pesquisa abrangem questões morfológicas e semânticas e requerem abordagem complexa do ponto de vista teórico, metodológico e empírico. Para tal, organizamos nossa discussão nesta dissertação em cinco capítulos. O capítulo de Introdução comenta as motivações da pesquisa. No capítulo 2, denominado Referencial Teórico, faz-se um levantamento das questões teórico-gramaticais relacionadas ao sufixo *-inho*, como as que dizem respeito ao grau e aos aspectos morfológico e semântico. Também nesse capítulo fundamentamos aspectos referentes a oposição língua falada x língua escrita, e a classificação do discurso em modos de organização narrativo e descritivo, segundo Charaudeau (2002), além da noção de contexto, conceitos essenciais para que possamos entender tanto o estado da arte sobre os estudos da formação em *x-inho* como a natureza do trabalho com *corpus* específico.

No capítulo 3, sobre os materiais e métodos empregados nesta dissertação, apresentamos aspectos históricos e definimos os conceitos que embasam a Lingüística de Corpus, apresentando, ainda, a descrição e definição dos *corpora* de estudo, o programa concordanciador WordSmith Tools e os procedimentos para análise dos dados.

⁵ Preferimos utilizar o termo grau baseando-se na prescrição da gramática tradicional. No capítulo 2, definimos algumas concepções e controvérsias sobre o enquadramento das formações *x-inho* com o rótulo de grau diminutivo.

No capítulo 4, detalhamos as análises e resultados obtidos em função dos objetivos da pesquisa, apresentando uma análise contrastiva entre os *corpora* e, para o *corpus* oral, entre os Modos de Organização narrativo e descritivo.

Por fim, no capítulo 5, apresentamos as considerações finais, retomando o estudo desenvolvido, apresentando suas limitações, contribuições e possíveis implicações, bem como os prováveis encaminhamentos da presente dissertação. A Bibliografia e os Anexos constituem a parte final da dissertação.

CAPÍTULO 2 - REFERENCIAL TEÓRICO

Neste capítulo, apresentamos alguns conceitos teóricos que permeiam e embasam o presente trabalho. Inicialmente tratamos das questões do *x-inho* na língua portuguesa numa visão teórico-gramatical e descritiva, bem como algumas concepções sobre a questão do grau. Em seguida, comentamos os Modos de Organização do Discurso segundo Charaudeau (1992), as idéias teóricas que embasam a noção de contexto segundo a Gramática Sistêmico Funcional de Halliday (1985) e Maingueneau (2002), e a oposição língua falada x língua escrita.

2.1. Visão geral das formações *x-inho*

Esta seção faz um levantamento do estado da arte sobre as concepções e definições que se tem sobre o sufixo *-inho*, baseando-se em noções de caráter descritivo-normativo. A seção começa com *x-inho* enquanto expressão do “diminutivo”.

Para Hakamies (1951, *apud* Ezarani, 1989:16):

Segundo uma tradição que remonta à antiguidade, convencionou-se dar o nome de diminutivos a certos derivados que não implicam necessariamente a idéia de pequenez, mas podem expressar também uma atitude afetiva do sujeito falante. No entanto, o valor diminutivo, com as nuances acessórias, é secundário; isto resulta de um exame dos sufixos diminutivos das línguas indo-européias, pois eles indicam originariamente a semelhança e a pertinência de uma espécie.

Ainda, segundo Ezarani (1989), estudiosos como Hofmann, Conrad e Leuman (cf. Hakamies, 1951, *apud* Ezarani, *op.cit*:16) têm-se voltado para a questão do valor semântico originário dos diminutivos, no intento de esclarecer se é do valor afetivo que decorre a noção de pequenez ou vive-versa. Porém, segundo a autora, até a descrição de seu trabalho, as pesquisas não tinham chegado a resultados satisfatórios.

Segundo Cunha (1994), os sufixos *-inho* e *-ino* provêm do latim *inus*. A forma tipicamente portuguesa é *-inho*; *-ino*, variante erudita, só aparece com valor diminutivo em um restrito número de palavras; *-im* é importação do francês *-in*, ou do italiano *-ino*, através da forma francesa. Compare-se: *tamborim*, do francês *tambourin*; *festim*, do francês *festin*, por sua vez derivado do italiano *festino*.

A formação de diminutivos por derivação sufixal era comum em latim. Entretanto, o sufixo principalmente utilizado (-ũlu- e sua variante -cũlu-), em virtude da sua natureza átona, foi abandonado pouco a pouco em latim vulgar. Nos nomes em que se manteve, perdeu a

função diminutiva e integrou-se no radical, constituindo uma nova palavra simples, que substituiu o seu termo primitivo e, em português, sofreu uma redução do vocábulo grave pela supressão da vogal penúltima átona (*apicula* > *apecta* > *abelha*, em vez de *apis*).

Os dois sufixos básicos diminutivos, em português, são assim: *-inh* (o, a). Lat. vulgar *-in* (u, a). A aplicação é praticamente irrestrita (*livrinho*, *casinha*, etc.) Um e outro apresentam uma variante com *-z-* inicial, como consoante de certos radicais que se integram no sufixo (cf. *rapazinho*, *rapazito*, de *rapaz*). No português moderno, há sensível preferência por essa variante, que fica em justaposição com o termo primitivo numa locução em que tanto este como o sufixo têm flexão: *lobozinho*, *lobazinha*. O sufixo *-inho*, em sua forma feminina *-inha* estabelece, em alguns casos, analogamente, um vocábulo primitivo no masculino: *galo-galinha*, *rei-rainha* (onde o *-i-* tônico do sufixo desviou o *-e-* radical para *-a-* e a vogal assilábica de *rei* foi suprimida).

O sufixo *-inho* (*-zinho*) é de enorme vitalidade na língua, podendo promover na palavra derivada uma ou mais de uma função semântico-pragmática. Essas funções, portanto, podem vir a ser elucidadas pelas características morfo-sintáticas das bases e/ou pela interpretação das formações em contextos discursivos.

Ainda, segundo Cunha (*ibid*), excetuando-se o caso das palavras terminadas em *-s* e *-z*, que naturalmente exigem a forma *-inho* (*piresinho*, *rapazinho*), não é fácil indicar as razões que comandam a escolha entre *-inho* e *-zinho*. Sente-se que muitas vezes a seleção está ligada ao ritmo da frase. Por outro lado, verifica-se uma preferência na linguagem culta pelas formações com *-zinho*, no intuito de manter íntegra a pronúncia da palavra derivante; a linguagem popular, no entanto, simplificadora por excelência, tende para as formações com *-inho*, pois conforme afirma Azeredo (2000:120), sem estudar *xinho* quantitativamente, “A língua coloquial faz amplo uso dos sufixos *-ão* e *-inho*: *bonitão*, *fininho*, *estreitinho*, *pequeninho*.”

Ainda, dentro de uma perspectiva descritiva, Lima (1994) admite a expressão do diminutivo sintético também através dos sufixos, *ito*, *ulo*, *culo*, *ote*, *ola*, *im*, *elho*, além de *inho* e *zinho*. Este último sendo obrigatório quando o substantivo terminar em vogal tônica ou ditongo (cf. *cafezinho*, *paizinho*).

As definições propostas para o diminutivo em algumas gramáticas escolares consultadas (cf. Lima, 1994; Bechara 2000; Cunha, 1972; Sacconi, 1991; Infante, 1995; Ribeiro, 1988;

André, 1975, entre outros) descrevem o diminutivo como uma “diminuição do ser”, isto é, do substantivo, relativamente a seu tamanho normal, embora alguns gramáticos também incluam, com ressalvas, os adjetivos no processo. “As formas diminutivas de adjetivos podem (precedidas ou não de muito, mais, tão, bem) adquirir valor superlativo: *Blusa amarelinha, garoto bonitinho; É bem feiozinho, benza-o Deus, o tal teu amigo!*” (Bechara 2000: 152)

A diminuição pode ser expressa por dois modos distintos, cujos resultados, no entanto, são considerados sinônimos. São eles a sufixação e a colocação de um adjetivo ou advérbio de intensidade junto ao nome. No primeiro caso, diz-se que o diminutivo é sintético (cf. *copinho*); no segundo, analítico (cf. *copo pequeno*). Considera-se também o diminutivo expresso por meio de processo lexical como tendendo a exprimir afetividade (cf. *mãezinha*), ao contrário do diminutivo sintético, cuja tendência é indicar dimensão pura e simples.

O diminutivo raramente é definido de forma explícita nas gramáticas escolares. A descrição de suas manifestações nos nomes e nos adjetivos, no entanto, demonstra que é concebido como “a maior ou menor intensidade que se pode dar à significação das palavras” (Lima, 1994:86). A diferença entre as várias obras reside no tocante à extensão do léxico a qual tal fenômeno é aplicável: somente aos adjetivos; aos adjetivos, nomes e advérbios; aos adjetivos, nomes, pronomes e verbos; a todo o léxico, com exceção de nomes próprios e pronomes.

Embora divergindo, a princípio, a respeito das classes gramaticais a que é aplicável, as obras consultadas dedicam espaço razoável apenas à gradação de nomes e adjetivos, e, em alguns casos, à dos advérbios.

Em algumas obras (cf. Lima, 1994 e Cunha, 1972), a preocupação histórica faz com que sejam listadas como exemplos de grau diminutivo formações que embora possam ter sido originariamente diminutivos, há muito perderam esse valor. Assim, listam-se como diminutivos termos como: *coreto, lingüeta, cartilha, espadim, galinha, rapariga*. É a este tipo de exemplos que Lima (*ibid*) dá o nome de “diminutivos formais” e Cunha, (*ibid*), de “especialização de formas”:

“Muitas formas, originariamente aumentativas e diminutivas, adquiriram, com o correr do tempo, significados especiais, por vezes dissociados do sentido da palavra derivante. Nestes casos, não se pode mais, a rigor, falar em aumentativo ou diminutivo. São, em verdade, palavras em sua acepção normal.”(Cunha,1972:209)

Ressalte-se, ainda, que, segundo Lima (*op.cit.*), o “grau” (diminutivo) para os advérbios, pronomes e verbos é considerado um acidente (uma flexão) excepcional, o que restringe o tratamento desses casos, excluídos os advérbios, quanto muito, à citação de um exemplo: *Irei agorinha mesmo. Elezinho é um encanto! Nenê está dormindinho.*

Cunha (*op.cit.:72*) também aponta a possibilidade de, na linguagem coloquial, os advérbios virem a receber “sufixos diminutivos” (*inho* e *zinho*), ressaltando-se, porém, que tais formações têm valor “superlativo”. Assim, *devagarinho* equivaleria a “bem devagar” e *pertinho* a “muito perto”.

Parecendo não acrescentar nenhuma contribuição às discussões anteriores, Loures (2000) também destaca a carga semântica contida nos “afixos de grau”, conforme a autora assim denomina o sufixo *-inho*. Segundo ela, os sufixos de diminutivo têm como função principal a afetiva e não expressam necessariamente dimensão, mas a afetividade do falante, podendo veicular aspectos positivos (cf. *cachorrinho, colinho*) e negativos (cf. *mulherzinha, leizinha*). A função do diminutivo é de realçar qualidade e/ou quantidade, de acordo com os padrões individuais e subjetivos do falante. Porém, o que se nota é que os valores quantitativos demonstram-se sensivelmente inferiores aos qualitativos. A melhor explicação é, segundo a autora, que os valores qualitativos estão relacionados ao que é agradável ou desagradável, como sentimentos de prazer ou desprazer (cf. *Preciosas horinhas de sono*). Loures (*op.cit.*) analisa a produtividade de vários sufixos do português (cf. *-inho, -ote, -ete, -ito, -ão, -aço*) e observa que o sufixo diminutivo *-inho* teoricamente pode ser afixado a todas as palavras da língua, ou seja, apresenta alto grau de produtividade.

Segundo Rosa (1982) alguns autores como André (1978) e Cegalla (1977) também acrescem a possibilidade de aumentativos e diminutivos de adjetivos, com a observação de que podem comportar-se também como superlativos. “Nesses casos, o sentido que tais adjetivos adquirem não é propriamente de tamanho grande ou pequeno (...): *bonitinho* equivale a muito bonito, *espertalhão* equivale a muito esperto.” (André, 1978: 119 *apud* Rosa 1982:10)⁶

Entretanto, alguns autores como Luft (2002), Lima (*op.cit.*), Bechara (*op.cit.*) e Cunha (*op.cit.*) admitem que, apesar de, em geral, o “aumentativo” e “diminutivo” sintéticos veicularem a idéia de tamanho (diminuição/aumento), nem sempre essa noção é realizada. Os

⁶ André, Hildebrando Afonso de. *Gramática ilustrada*. São Paulo, Ed. Moderna, 1978.

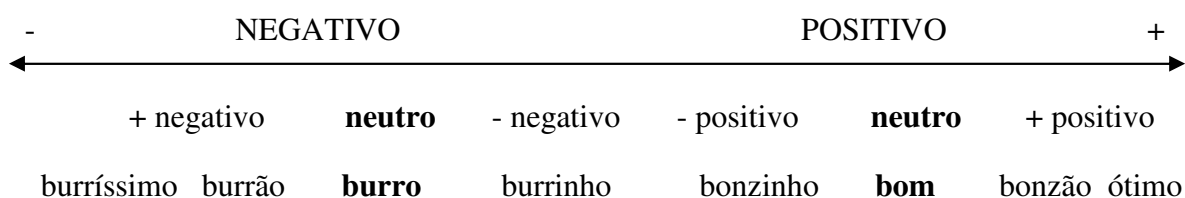
autores acreditam terem os sufixos de grau variadas significações, tais como afetividade (cf. *mãezinha, filhão*) e desprezo ou pejoratividade (cf. *narigão, livreco*).

Fora da idéia de tamanho, as formas aumentativas e diminutivas podem produzir o nosso desprezo, a nossa crítica, o nosso pouco caso para certos objetos e pessoas, sempre em função da significação lexical da base, auxiliados por uma entoação⁷ especial (eufórica, crítica, admirativa, lamentativa, etc.) e os entornos que envolvem falante e ouvinte: ... *coisinha, issozinho*. (Bechara, *op.cit.*,141)

O emprego dos sufixos indica ao leitor/interlocutor que aquele que fala ou escreve põe a língua afetiva no primeiro plano. (...) quer exprimir de modo espontâneo e impulsivo, o que sente, o que o comove ou impressiona. Assim encontra-se no sufixo um meio estilístico que elide a objetividade sóbria e a severidade da linguagem, tornando-a mais flexível e amável, mas às vezes também mais vaga. (Skorge, 1958, *apud* Cunha, 1972:209)

Frota (1985:13, *apud* Ezarani, 1989:21) aponta a existência de uma escala de valores em que podemos situar as palavras num ponto tal que adequadamente traduza aquilo que queremos comunicar, ao nos dizer que sufixos “diminutivos” e “aumentativos” apenas acentuam ou minimizam a carga semântica da base, tenha ela valor pejorativo ou não”. Concordamos, portanto, com Ezarani (*op.cit.*) em acreditar que esta escala deve ser considerada num *continuum* no qual posicionamos as palavras dentro de um maior ou menor grau de intensidade ou dimensão do significado positivo ou negativo de sua base. Assim, a autora exemplifica, conforme quadros 1 e 2 abaixo (adaptado de Ezarani, 1989:22), o grau de intensificação de bases negativas e de bases positivas como sendo bem mais nítidas em adjetivos e o grau de dimensão que ocorrerá nitidamente em bases substantivas.

Quadro 1

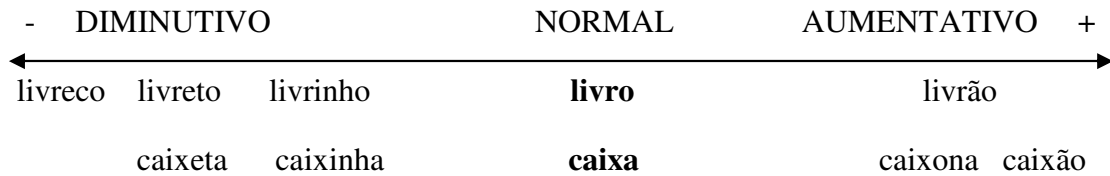


Observa-se, no quadro 1, que as bases neutras *burro* e *bom* de cargas semânticas negativa e positiva, respectivamente, variam em um *continuum* de maior ou menor grau de intensidade, pois, em relação à base *burro*, *burrão* tende a ser mais negativo e *burrinho*

⁷ Apesar de essa dissertação trabalhar com *corpus* oral transcrito, não enfocamos prosódia.

menos negativo, assim como *bonzinho* tende a ser mais negativo e *bonção* mais positivo em relação à base *bom*.

Quadro 2



Da mesma forma, observa-se no quadro 2 que as bases neutras *livro* e *caixa* variam em um *continuum* de maior ou menor grau de dimensão, pois em relação à base *livro*, *livreto* demonstra ser mais diminutivo do que *livrinho* e *caixeta* mais diminutivo do que *caixinha*, assim como *caixão* demonstra ser mais aumentativo do que *caixona*. Ressalta-se, porém, que, conforme veremos no decorrer deste trabalho, as formas *livrinho* e *caixinha* podem adquirir uma função afetiva negativa, dependendo do contexto em que são produzidas.

Para Luft (*op.cit*:144), as idéias de grandeza e pequenez facilmente podem evocar (conotar) as de anormalidade, defeituosidade, afeição, simpatia. Admite, portanto, que os sufixos diminutivos podem ser freqüentemente pejorativos, ou então hipocorísticos (de carinho): *livrinho*, *livreco*, *Pedrinho*, *filhinha*, *benzinho*; etc.

Rocha (1986:17, *apud* Emílio, 2003:12) admite que “o grau diminutivo” é uma categoria morfológica que expressa relação de dimensão, intensidade ou afetividade, que se estabelece entre o termo base e a respectiva forma derivada”. Na definição de Rocha (*op.cit.*), o significado-base mantém-se independentemente das relações expressas pelo diminutivo e, ao ser acrescido do sufixo *-inho* apresenta três problemas: a) dimensão pura, b) dimensão + afetividade, c) dimensão + extensão do significado.

Ainda assim, pode-se ressaltar que as formações em *-inho* podem veicular traços sócio-geográfico-culturais do falante. Ao que tudo indica, a expressão em que ocorre *-inho* parece apresentar função indexical. Nesse sentido, Gonçalves (2000) afirma, sem examinar *corpora* específicos, que o sufixo *-ão* é usado preferencialmente pelos jovens (cf. *prazerão*, *abração*, *povão*). Quando o sufixo *-ão* é utilizado em oposição ao diminutivo *-inho*, caracteriza, normalmente, a linguagem das crianças, pois o seu uso demonstra uma riqueza em recursos expressivos: *molengão*, *travessurinha* x *castigão*.

Verifica-se, portanto, que todas as obras mencionadas corroboram a idéia de que as formações em *-inho* na língua portuguesa não expressam necessariamente dimensão, mas também significação afetiva, pejorativa e intensificadora. Porém, apesar de essas obras admitirem tais possibilidades de uso dessas formações, não explicitam qual das dimensões é mais freqüente, sugerindo, portanto, a idéia de que a função dimensiva venha a ser, ainda, a mais prototípica e a mais produtiva em relação as outras funções. É nesse sentido, como já dissemos, que tentamos levantar e analisar neste trabalho as ocorrências das formações *x-inho*, verificando qual das funções tem a maior freqüência e produtividade dentro de *corpora* escrito e oral, contrastivamente.

2.1.1 Aspectos morfológicos do sufixo *-inho*

Do ponto de vista morfológico, acentuamos que, ao contrário dos aumentativos em *-ão*, os “diminutivos” em *-inho* (e também em *-ito*) não sofrem mudanças de gênero. O “diminutivo” conserva o gênero da palavra derivante: *casa-casinha-casita*. Convém ressaltar, ainda, que as formações populares em que o sufixo *-inho* se junta a participípios, sejam estes irregulares, tornam-se regulares, conforme exemplo abaixo coletado da fala espontânea, observada no nosso dia-a-dia.

Ex: *Aquele prêmio foi muito bem ganhadinho e o dinheiro muito bem gastadinho por mim.*

Apesar do gênero não ser o ponto chave de nosso trabalho, em relação aos aspectos morfológicos, podemos citar Azeredo (2000:111), o qual afirma que o sufixo *-(z)inhol(z)inha* se comporta como unidade autônoma em relação ao gênero, pois é ele, e não o substantivo como um todo, que se flexiona em nomes como *pontezinha* e *pelezinha*, já que os nomes *ponte* e *pele* são de tema em *-e*. O *-a* dos “diminutivos” *portinha*, *ruazinha*, *pontezinha* e *pelezinha* é desinência própria do sufixo.

A estratégia de afixação de *-inho*, como já foi abordado na seção 2 é normalmente utilizada, mesmo não significando necessariamente diminuição de tamanho. Ex: *Ela tem um cachorrinho tão bonito...*, ao contrário de: *Ela tem um cachorro pequeno tão bonito.*

Outra característica morfológica importante apontada por Ezarani (*op.cit.*: 29) é que nem sempre a identidade fonética em final de vocábulo corresponde a presença de sufixo.

Assim não há derivação alguma em palavras como *galinha*, *rainha*, *vizinho*, *toucinho*, etc. Consideramos estas palavras como itens lexicalizados⁸.

Segundo Macambira (1999:35), há a classificação do substantivo sob o aspecto mórfico em que pertence à classe do substantivo toda palavra variável que admite os sufixos *-inho* ou *-zinho*, *-ão* ou *-zão* correspondentes a pequeno e grande respectivamente:

a) *livro*: donde *livrinho* = *livro pequeno*

b) *pé*: donde *pezinho* = *pé pequeno*

Macambira (*ibid*) corrobora também a idéia de que alguns adjetivos, pronomes e advérbios assumem familiarmente os sobreditos sufixos, não, porém, correspondentes a pequeno ou grande.

Nesse sentido, citamos também, numa análise descritiva, Basílio (1989) que compartilha da mesma idéia ao considerar que o “grau”, e em nosso caso as formações em *-inho*, é o caso mais comum – embora não reconhecido como tal – de processo morfológico a serviço da função expressiva da linguagem, corroborando também as idéias de Loures (2000) que afirma que os sufixos “diminutivos” têm como função principal a afetiva, podendo expressar aspectos positivos (cf. *pedacinho*, *nenezinho*) ou negativo (cf. *empregadinha*, *profissãozinha*). Em linhas gerais, segundo Loures (*ibid*), a modalização apreciativa aparece associada às formas em *-ão/-aço*, enquanto a depreciativa tende a se ligar às formas “diminutivas”.

As formações em *-inho* aplicam-se praticamente a qualquer palavra da língua, não reconhecendo exceções. De acordo com Loures (*ibid*), o sufixo *-inho* pode ser empregado a vários tipos de bases, entre elas as pronominais (cf. *euzinho*, *elazinha*), as verbais (cf. *estouzinho*, *dormindinho*), as interjeitivas (cf. *adeusinho*, até *loguinho*, *tchauzinho*). Tal produtividade está associada aos valores afetivos e, nesse caso, as lacunas praticamente não existem (cf. *cheinha*, *bonzinho*, *povinho*, *tipinho*).

Um aspecto levantado por Piza (2001) é que as formações em *-inho*, normalmente, não são responsáveis por mudanças de categoria lexical. O acréscimo de *-inho* a qualquer base não vai alterar a categoria lexical da mesma. Isso significa dizer que essas formações não

⁸ A noção de lexicalização será discutida na seção 2.1.2

apresentam função sintática, característica observada nos processos flexionais. Os únicos casos excepcionais são os itens *doisinho* e *umazinha*⁹, que têm suas categorias modificadas, porém em função de dois tipos de lexicalização sofridas por eles – a categorial e a semântica. Nesses dois exemplos, um numeral passa a substantivo com o acréscimo de *-inho*.

O acréscimo de *-inho* a qualquer base pode caracterizar noções semânticas variadas. A estabilidade semântica dessa formação é um aspecto relevante e amplo. As estratégias de gradação são comumente associadas à forma de dimensionar, aumentando ou diminuindo o tamanho, sem a preocupação de verificar se essas noções são realmente atualizadas nos itens lexicais. Todavia, podemos observar que formações em *-inho* incluem-se também, como já foi abordado, nas intenções emotivas, ou seja, nos valores afetivos e/ou depreciativos.

Outra característica importante observada é que o sufixo *-inho* permite uma convivência harmoniosa com o sufixo aumentativo e com naturalidade a ele mesmo em uma só unidade lexical. Em *calçãozinho*, *cartãozinho*, *portãozinho*, não há restrição alguma quanto à presença de dois diferentes tipos de sufixos. A confirmação reside no fato de que sufixos com a mesma função e de semântica contrária podem se combinar sem problemas. Esse convívio é muito comum nesses casos de lexicalização, mas nada impede que aconteça também em itens não-lexicalizados (cf. *pequeninhinho*, *meninhazinha*). Em *livrinhozinho*, por exemplo, à função diminutiva se ajunta a função de atitude subjetiva positiva ou não, tendo ambas, possivelmente, igual relevância no discurso.

Entretanto, sobre a possibilidade de formações como as assinaladas acima, virem a distinguir *-inho* de *zinho*, Ezarani (*op.cit.*) afirma que, se há distinção, esta se encontra no plano fonético-fonológico e/ou morfológico e não no semântico-pragmático. De fato, em derivações como *meninhinho*, *professorinhazinha*, *engenheirinho* etc., em qual dos elementos (*-inho* ou *-zinho*) se projeta a atitude subjetiva do falante que pode querer veicular positividade, pejoratividade e/ou intensificação e até diminuição? Segundo a autora, pode-se dizer, portanto, que os elementos *-inho* e *-zinho* equivalem-se no plano semântico-pragmático.

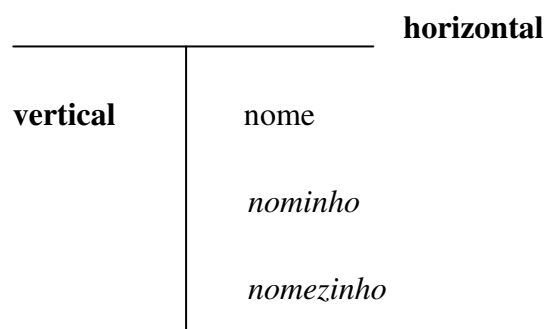
⁹Apesar de aparentemente os vocábulos citados darem a idéia de diminuição numérica, seus significados diferem totalmente do suposto, haja vista tratarem de dois substantivos: *doisinho*, palavra popularmente utilizada para evidenciar o consumo de droga (especificamente “maconha”); *umazinha*, forma também popular de expressar algum tipo de relação sexual tida ou mantida por alguém.

Em relação às formações com *-inho* e *-zinho*, dois exemplos retirados do nosso *corpus* escrito mostram que Ezarani parece não ser muito pertinente em sua afirmação, pois diferentemente do que postula a autora, acreditamos que esses sufixos não são equivalentes no plano semântico-pragmático. Demonstramos esse fato com os exemplos abaixo, em que uma mesma base lexical foi acrescida de *-inho* e *-zinho*, conforme a subjetividade do enunciador.

1. ...com as apresentações no segundo semestre de Digable Planets, Guru e US3 no Brasil, aqui vai mais um **nomezinho** para você começar a se preocupar em ver por aqui: Urban Species. Esse nome é o sinônimo de...(E:550)
2. Roger Avery, parceiro de Tarantino de longa data. Nem chega a ser um grande filme com sua violência algo forçada e péssimo final. Mas tem um **nominho** mágico assinando a produção executiva. (E:549)

No primeiro fragmento, o enunciador está se referindo ao nome de um componente de um dos grupos de Rock que se apresentará no Brasil. Pode-se notar que *nomezinho* traz ao significado-base um valor negativo, depreciativo e, pode-se dizer, irônico. Essa noção é realçada principalmente pelo artigo indefinido “um” que nesse contexto traz uma carga depreciativa, e pelo fato de o referido nome do músico trazer preocupação.

No segundo fragmento, o falante se refere a um filme de má qualidade, porém utiliza *nominho* para se referir a qualidade de um dos produtores. Nesse contexto, *nominho* traz uma semântica positiva, principalmente por aparecer juntamente com a palavra “mágico” que tem conotação positiva. Podemos representar as escolhas acima citadas de acordo com o gráfico 3, abaixo:



Quadro 3: Eixos constitutivos da linguagem

Conforme o gráfico acima, se é verdade, como Halliday (1985) afirma em sua Gramática Funcional Sistêmica, que o uso da língua se dá em dois eixos, o eixo horizontal ou dos sintagmas e o eixo vertical, ou o eixo das escolhas, então deve haver alguma diferença entre a escolha feita pelos enunciadores dos exemplos acima: *nomezinho* e *nominho* são escolhas distintas.

2.1.2. Lexicalização ou idiomatização

São numerosos os exemplos de itens lexicais do português que apresentam morfema de grau, “aumentativo” ou “diminutivo”, esvaziado de sua função essencial de indicar tamanho grande ou pequeno. São exemplos de palavras lexicalizadas ou idiomatizadas: *calção*, *facão*, *portão*, *calcinha*, *caminho*, *armarinho camisola*, *camiseta*, *parquinho*. A prova mais simples de que são formações não transparentes, isto é, lexicalizadas, é que em geral aceitam o sufixo de semântica contrária: *calçãozinho*, *facãozinho*, *portãozinho*, *caminhozinho*, *armarinhozinho*, *camisolão* e até sufixo de mesma função de grau: *portãozão*, *facãozão*, *camisolinha*, *camisetinha*.

Há porém, algumas construções multilexicais em *-inho* que a língua não dá conta de definir, como por exemplo, as encontradas em nosso *corpus* escrito: *dar um jeitinho*, *queridinho do chefe*, *ter um gostinho de quero mais*, *isso tem o dedinho do governo*, *fazer beicinho*. Não há um critério para que possamos caracterizar essas expressões como: provérbio, expressão idiomática, frase feita, clichê, chavão etc. Em nosso trabalho chamamos essas construções de expressões pré-fabricadas “lexical prefabs” usada em inglês por Ermanm e Warren (2000). A nosso ver, esse é um termo que serve para expressões recorrentes e mais ou menos fixas, pois elas não têm a mesma “estabilidade” e “fixidez” das chamados expressões idiomáticas.

Algumas formações em *-inho* parecem estar em um processo de lexicalização na língua, como é o caso de *neguinho* nos exemplos abaixo:

3. ...e na esquina do hotel um “homeless” **neguinho** e magrelinho levantou a placa: Socorro! Minha mulher foi seqüestrada. (E:1063)
4. e::ta... essas coisas todas... né? estava tudo dentro do carro... então tinha milhões:: assim... além do carro... e estava tudo dentro... **neguinho** não tinha tirado... nada... incrível... né? E: mas eu não... não entendi... e aí? ele pegou o carro [e foi andando?] I: [aí ele pegou (ON: 5)

5. Brasil “e a minha pátria. Na minha opinião, acho que neguinho ficou se mordendo o show inteiro para encontrar alguma coisa para estragar (E: 25)

No exemplo 3, *neguinho* traz uma conotação negativa, pejorativa a qual é revelada, principalmente, pelo contexto de sua produção. Porém, nos exemplos 4 e 5, não há essa conotação negativa, pois percebe-se que *neguinho* nesses dois últimos exemplos tem um significado mais geral, representando um grupo de pessoas. Em 4, pode-se dizer que *neguinho* refere-se àqueles que, talvez, tenham roubado um carro, e em 5 a algumas pessoas que, por ventura, talvez quisessem algum pretexto para estragar um show. Como podemos perceber nos exemplos 4 e 5, a formação *neguinho* parece estar em um processo avançado de lexicalização na língua portuguesa, pois produções desse tipo são bastante frequentes no nosso dia-a-dia.

Segundo Rocha (*op.cit.*), algumas formações em *-inho* adquirem um sentido especial em sua forma sintética, podendo representar redução de tamanho ou não, com em: *casquinha* (de sorvete), *tatuzinho* (tipo de crustáceo ou de inseto), etc. Em outras palavras, podemos dizer que a idéia expressa pelas formações em *-inho* dependem do contexto, o que autoriza dizer que a construção morfológica, por si só, não é suficiente para determinar o tipo de informação contida no léxico.

Também, segundo Piza (2001), com as formações em *-inho* pode ocorrer a lexicalização semântica que, na verdade, é um congelamento de sentidos. O que normalmente era empregado para atribuir diminuição ou aumento, passa a caracterizar itens de sentido específico, ou seja, os afixos de “grau” podem apresentar-se desprovidos semanticamente da função original do grau, idiomatizando ou transformando as formas em não-transparentes (cf. Sandmann, 1992 apud Piza, *ibid*). É o que acontece com palavras do tipo *orelhão* e *camisinha*, entre outras que possuem referentes diferenciados em relação a semântica da base. Trata-se de uma interpretação metafórica promovida pelo uso de tais afixos.

2.1.3. Aspectos estilísticos

Acreditamos que as formações em *-inho* estejam ligadas a aspectos estilísticos, pois, em síntese, entendemos que o estilo é resultado de um trabalho de escolha lexical, morfológica, sintática, etc. É na *expressividade*, que vem a ser “a particularidade constitutiva do enunciado, realizada pelo contato entre significação lingüística e realidade objetiva” (Bakhtin,1995:311,apud Emilio, 2003:19), que essa escolha se efetiva. Entendemos que a característica fundamental da expressividade reside na força de persuadir, ou de transmitir

conteúdos desejados, na capacidade apelativa, no poder de gerar elementos evocativos. Os efeitos de sentido produzidos por uma seqüência lingüística, efetuada por meio de uma escolha lexical, dependem do enunciador e do co-enunciador, reais ou pretendidos, porque são sujeitos da interação comunicativa. Assim, uma escolha pode ter efeitos diversos, dependendo de quem a realize.

Portanto, as palavras, como elementos estilísticos do enunciado, realizam-se na interação em função da situação e do interlocutor, que servem de orientação para o locutor o qual, por sua vez, valeu-se de um trabalho de escolha. Essa serve de instrução para o interlocutor, que deverá, a partir dela, estabelecer um efeito de sentido entre ele e o locutor/produzidor responsável pela escolha daquele(s) elemento(s) e não de outro(s). Nesse sentido, citamos Santa'Anna (1989;114):

O diminutivo está na fala de todos, cultos ou ignorantes, e só não aparece com tom afetivo nos textos escritos que têm por meta a objetividade (...) em muitos casos de diminutivo erudito, mesmo a idéia de pequenez passa despercebida.”

Dessa forma, salta aos olhos de todo analista da língua portuguesa, especialmente em suas variantes informais, a enorme freqüência de palavras com formações em *-inho*, com a função de expressar, mais do que tamanho pequeno, como já dissemos, emocionalidade, que pode ser apreço ou despreço. Assim, dificilmente *mulherão*, *carrão* ou *timinho* e *mulherzinha* deixarão de expressar afetividade; no caso de *mulherão* e *carrão* valorativa, no caso de *mulherzinha* provavelmente depreciativa e no caso de *timinho* depende de fatores contextuais: *o time jogou bem e ganhou* (afetiva) ou *jogou mal e perdeu* (depreciativa).

2.1.4. Aspectos lexicográficos

Embora se defenda que toda palavra complexa provida de morfema de “grau” é uma unidade lexical nova ou independente da base, não se defende, às vezes, a prática do registro de todas elas no dicionário, por mais abrangente ou completo que esse pretenda ser. Das palavras citadas no item anterior o Aurélio registra *carrão*, *mulherão* e *mulherzinha*, porém não *timinho*. Também não traz, por exemplo, *golzinho*, *queijinho*, *pastelzinho*. Excetuando *carrão*, a que o Aurélio atribui o significado lexicalizado de “instrumento de pescadores para puxar o barco para a terra”, além do de “carro grande”, não lexicalizado.

2.1.5. Estrutura sintática

Há , ainda, um fato que parece ser digno de registro neste trabalho de peculiaridades das palavras enriquecidas pelo sufixo *-inho*. Por transferir normalmente a palavra-base para outra classe de palavras (*esmiuçar – esmiuçamento*) ou para outra subcategoria da mesma classe de palavras (*marmelo – marmeleiro*) o sufixo é considerado o núcleo ou determinado do produto. *-mento* faz, no exemplo acima, de *esmiuçar*, verbo, um substantivo, e *-eiro* faz do substantivo *marmelo* “fruta” outro substantivo, mas de semântica bastante diversa: “árvore que dá marmelos”.

Em Back e Mattos (1972) nossos sufixos derivacionais, exceto o de “grau”, são chamados raízes, porque ocupam o eixo do vocábulo, são o núcleo ou o sol do produto. O morfema de grau é, juntamente com nossos sufixos flexionais, chamado de sufixo e desempenha função de raio ou adjunto. Mas é feita uma diferenciação importante: o morfema de grau, por sua ausência não ser significativa, é chamado de sufixo facultativo, enquanto as flexões são consideradas sufixos obrigatórios, pois sua ausência é significativa. Exemplificando: em *livro* e *falava* a ausência, respectivamente, do *-s* significa fatalmente singular e 1ª ou 3ª pessoa do singular, em oposição ao plural *livros* e a 2ª pessoa do singular *falavas*; em *carro*, a ausência de *-ão* (*carrão*) ou *-inho* (*carrinho*) não é significativa, isto é, não significa necessariamente grau normal ou “carro que não é grande ou pequeno”. O morfema de grau é, pois, adjunto ou determinante da palavra complexa.

Contrariamente aos outros sufixos, que são raiz ou determinado, o sufixo de grau é raio ou adjunto, característica que compartilha com as flexões em geral. Por outro lado o morfema de grau se associa aos sufixos em geral chamados raízes por seu caráter de facultatividade. Em outras palavras, poderíamos dizer que a presença ou ausência do morfema de grau não é condicionada por fatores sintáticos.

2.1.6. Os traços pejorativo e afetivo

Conforme vimos anteriormente, a formação em *-inho* pode indicar maior ou menor intensidade conferida a uma palavra, no caso específico dos nomes acrescenta-se a possibilidade de demonstrar desprezo, intimidade, afetividade.

Muitas vezes, principalmente no “grau” aumentativo, a enunciação ganha um caráter de intenção depreciativa, isto é, indica que desprezamos o ser: *narigão* (um nariz feio de tão

grande), valentão (um homem que faz ostentação ridícula de ser valente), papelucho (um papel que não vale nada). O “grau diminutivo” também expressa carinho especial pelo ser, sem alusão às suas proporções: *mãezinha* querida, em vez de *mãe* querida.

No primeiro caso, isto é, quando pretendemos indicar desprezo, dizemos que o termo é *depreciativo* ou *pejorativo*; no segundo, quando a intenção é a de expressar carinho, dizemos que o termo é *afetivo*. Seria justificável, no entanto, tratar a expressão quantitativa pura, simples e o valor emotivo de derivados com esses sufixos como um mesmo fenômeno identificável pelo rótulo de grau? Poder-se-ia falar em relação de quantidade nos exemplos considerados pejorativos ou afetivos?

2.2. O sufixo *-inho* visto como grau

Apesar de não ser o foco principal de nossa pesquisa, esta subseção aborda, de maneira sucinta, o tratamento dispensado ao grau em estudos de caráter descritivo-normativo de língua portuguesa, focalizando o tratamento de algumas gramáticas normativas. Apresenta, ainda, a abordagem proposta por Camara Jr., apontando a controvérsia acerca da inclusão do grau em português entre os processos flexionais ou entre os derivacionais. Achamos de caráter relevante salientar essas noções neste trabalho, em virtude do tratamento dado ao sufixo *-inho* que normalmente é considerado como flexão em algumas gramáticas normativas, enquadrando-se no rótulo gramatical: flexão dos substantivos (gênero, número e grau).

Primeiramente, salientamos Rosa (*op.cit.*) que focaliza, em seu estudo sobre a formação dos aumentativos, a problemática que se tem sobre o grau¹⁰ em português. A autora apresenta um sumário do tratamento dispensado ao grau em gramáticas normativas e na abordagem de Mattoso Câmara Jr., abordando os traços pejorativo e afetivo e do problema derivação vs. flexão. Ela reconhece, portanto, que as gramáticas tradicionais ainda permanecem confusas nessa questão, pois não conseguem situar definitivamente o grau como flexão ou como derivação

Os sufixos “diminutivos” têm sido motivo de interesse em vários estudos, em virtude de algumas particularidades. Esses estudos incluem também o fato de o diminutivo, enquanto afixo e na condição morfológica de grau, consistir em flexão ou derivação. De todas as

¹⁰ Concordamos com Rosa ao definir o grau como sendo “*uma categoria que expressa a relação existente entre um significado considerado normal e outro (s) considerado (s) acima, abaixo ou no mesmo nível numa escala de intensidade (muito...pouco) ou de dimensão (pequeno...grande), incluídos os valores pejorativos e afetivos.*” (Rosa, 1982, apud Ezarani, 1989:21)

gramáticas consultadas (Cunha,1972:208; Lima,1995:87 e Bechara, 2000:140) apenas Cunha considera o grau como processo de natureza flexional. Lima (*op.cit.*) e Bechara (*op.cit.*) compartilham a idéia de que a gradação caracteriza-se como processo que se aproxima mais da derivação.

“A NGB , confundindo flexão com derivação, estabelece dois graus de significação do substantivo:

a) aumentativo: homenzarrão

b) diminutivo: homenzinho

A derivação gradativa do substantivo se realiza por dois processos, numa prova evidente de que estamos diante de um processo de derivação, e não de flexão”. (Bechara, 2000:140)

Ainda assim, corroborando Bechara (*op.cit.*) e Lima (*op.cit.*), nas abordagens escolares prescritivas de Sacconi (1991), Infante (1995) e Ribeiro (1998), observa-se forte tendência, explícita ou não, de admitir que o grau dos substantivos apresenta características derivacionais.

Sabe-se, portanto, que a derivação por sufixo também serve para assinalar no nome derivado um grau maior ou menor de dimensão ou intensidade em face do termo primitivo. Em referência aos substantivos cria-se assim a categoria dos “aumentativos” e dos “diminutivos” que se opõem, pelo seu sufixo específico, a uma dimensão considerada normal.

2.2.1. Algumas diferenças entre flexão e derivação

Dentre os estudos morfológicos encontramos, ainda, em relação ao grau, algumas distinções na classificação como um morfema flexional ou derivacional, com base nas características dos processos lexicais e/ou flexionais. Há ainda a seguinte questão: Por que a derivação, o processo morfológico com que formamos palavras, pode ser chamada de voluntária e por que a flexão, o processo morfológico com que adaptamos um item lexical a um contexto, a uma sentença pode ser chamada de natural?

Sabe-se que uma dezena de aspectos têm sido utilizados na intenção de isolar a Morfologia Flexional da Derivacional. Entretanto, na prática, essa divisão parece ser difícil, haja vista serem os critérios relacionados ou mutuamente dependentes. Desse modo, torna-se problemático estabelecer um limite preciso entre as duas “Morfologias”. O mais acertado é encarar o mapeamento dos critérios como tentativa de reconhecimento dos requisitos relevantes na categorização de um fenômeno como flexional ou derivacional.

Segundo Mattoso Camara (2001) o termo “flexão” parece não se adequar às características da própria expressão de grau, que não obedece às condições estabelecidas pelo quadro flexional. O autor afirma o seguinte:

“A expressão de grau não é um processo flexional em português, porque não é um mecanismo obrigatório e coerente, e não estabelece paradigmas exaustivos e de termos exclusivos entre si. A sua inclusão na flexão nominal decorreu da transposição pouco inteligente de um aspecto da gramática latina para a nossa gramática” (Câmara Jr.,2001:83)

Para ele, a freqüente associação do grau a processos flexionais devia-se à existência no latim, dos sufixos *-ior* comparativo e *-issimus* superlativo. Sua utilização dependia da natureza da frase, isto é, havia um contexto sintático que implicava o uso desses sufixos. Por exemplo, para o adjetivo latino Felix <feliz>, existiam duas formas que se mostravam obrigatórias e coerentes em ambientes sintáticos determinados: *felicier* <mais feliz do que> e *felicissimus* <o mais feliz do(a)>. O mesmo não ocorre em português, já que a gradação não é obrigatória e não depende de um contexto sintático para se realizar.

Ainda, assim, esse autor faz a distinção entre morfemas derivacionais e flexionais estabelecendo que os primeiros criam palavras novas e os flexionais adaptam cada palavra às condições específicas de cada contexto. Portanto, a flexão teria um caráter obrigatório e uma sistematização coerente dentro do sistema, para que se preencha a necessidade de uso adequado das formas lingüísticas em determinado contexto, enquanto a derivação ocorrerá ou não, a depender da vontade do falante.

Nesse sentido, para o autor, não há um quadro regular, coerente e preciso para os afixos que expressam gradação intensiva ou dimensiva, pois existe possibilidade de opção quando se usa ou não o vocábulo derivado.

Segundo Bechara (*op.cit*: 140.) a flexão processa-se de modo sistemático, coerente e obrigatório em toda uma classe homogênea, fato que não ocorre na derivação, o que levava o gramático e erudito romano Varrão a considerá-la uma *derivatio voluntária* (não previsível). Nessa ótica ,podemos exemplificar: casa + *inha*: *casinha* (derivação), pois o sufixo *-inho* pode ou não ser utilizado; casa + s: casas (flexão de plural), porém o morfema de plural *-s* dependerá da natureza da frase

Os morfemas gramaticais de derivação não constituem um quadro regular, coerente e preciso. Acresce a possibilidade de opção, para usar ou deixar de usar o vocábulo derivado.

Podemos citar ainda que no plano sintagmático, a flexão provoca o fenômeno da concordância: Ex: móvel novo → móveis novos em oposição a casa nova → *a casinha nova*.

Ainda há outras lacunas entre flexão e derivação. Na derivação, encontramos: *trabalhar* – *trabalhador*, *venda* – *vendedor*, mas não encontramos, *ensinar* – *ensinador*, *estudar* – *estudador*. Nem todos os nomes substantivos portugueses têm um diminutivo correspondente, e os que existem podem ser usados ou não, numa dada frase, de acordo com a vontade do falante. Já os paradigmas flexionais são conjuntos completos fechados, havendo obrigatoriedade e sistematização coerente, pois todos os adjetivos conhecem a forma masculina e feminina, singular e plural; assim como os verbos apresentam paradigmas conjugacionais completos. Alguns substantivos, porém, apresentam alguns fatos marginais como: *pires*, *férias*, *atlas*, palavras que no singular já apresentam o morfema –s (de plural) no final. Há também os não contáveis como: *arroz*, *feijão*.

Na flexão, não há mudança lexical, ou seja, estamos diante de uma mesma palavra. Como exemplo, citamos *mesa* e *mesas*, *cantar* e *cantava* como diferentes formas vocabulares de um mesmo lexema, pois é a natureza da frase que nos faz adotar um substantivo no plural ou um verbo na 1ª pessoa do pretérito imperfeito. Os morfemas flexionais estão concatenados em paradigmas coesos e com pequena margem de variação. Porém se acrescentamos um sufixo derivacional *-inha*, ao substantivo *mesa* teremos *mesinha*, que constitui outro lexema ou uma nova unidade lexical que pode ser usada independentemente da natureza frasal, mas sim da vontade do falante.

Há ainda na derivação a lexicalização, que é uma mudança lexico–semântica, que faz com que o todo não seja mais a simples soma das partes. Ex: *casinha* pode ser uma “casa pequena” ou uma casa muito apreciada ou “latrina”, “privada” que é o significado lexicalizado, ou ainda “bonitinha” que não significa (*bonita-pequena*) em oposição a “*barquinho*” que realmente pode assumir o sentido de “barco pequeno”.

Observa-se portanto que, na flexão, o lexema assume um valor semântico próprio sem nenhuma função poética da linguagem, não sendo passível de inovação, ou seja, a sua criatividade ou produtividade não é um fato constante.

Bechara (*op.cit.*) afirma que, por ser determinada por fatores sintáticos, é a concordância que diz se a flexão é não-voluntária ou *naturalis*. Por exemplo: no sintagma *desejos utópicos*, serão os fatores de concordância que determinarão a forma do adjetivo na

frase, pois poderíamos ter: *desejo utópico*, *vontade utópica*, *vontades utópicas*, *desejos utópicos*, pois a forma a ser usada independe da vontade do enunciador, mas da forma do substantivo, por isso, a flexão é uma forma de derivação não-voluntária ou necessária.

Em outro exemplo temos: *luz piscante*, que tem como outra opção, *luz que **pisca***, demonstrando que a forma escolhida não é fruto de nenhuma regra gramatical, mas sim de uma livre escolha do falante.

Dos pressupostos apresentados, percebe-se que nem sempre é simples classificar as palavras, quer como processos derivacionais ou flexionais, pois os fatos da linguagem são complexos, heterogêneos, dinâmicos e mutáveis no tempo. Não podemos, por exemplo, classificar a flexão *-ando* como apenas flexão de gerúndio (cf. *cantando*, *estudando*), pois também temos esse sufixo em substantivos como *cursando* e *formandos*. Há, ainda, o substantivo “mente”, que é hoje usado como sufixo adverbial, adquirindo outro significado: *abertamente*, *modernamente*. Por isso, na linguagem, é preciso conformar-se com a inexistência, às vezes, de distinções bem definidas.

Sandmann (1992) defende a natureza derivacional de todos os “afixos de grau”. Ele afirma que, no português, existem itens lexicais variados com afixo de grau, aumentativo ou diminutivo, sem indicar necessariamente tamanho (cf. *calção*, *camiseta*). Além disso, Sandmann (op. cit.) aponta para o caráter facultativo desses formativos, ou seja, sua utilização não é obrigatória, o que demonstra comportamento derivacional. O mesmo não ocorre com os sufixos dito obrigatórios, pois sua ausência acarreta problemas, aspecto próprio de elementos flexionais. O autor conclui que a ausência ou a presença do afixo de grau não é condicionada por fatores sintáticos.

Rocha (1994) analisa o “grau” dos substantivos também pelos aspectos mencionados anteriormente. O autor enfatiza a questão da afetividade presente ou não (p. 73). O autor realiza uma divisão dos substantivos afetivos em (a) sufixo subjetivo: *Podem me dar uma **licencinha**?*; (b) sufixo valorativo: (b1) positivo: *Que **carrão** é o seu, hein!* e (b2) negativo: *Ele fez um **papelão** naquela peça!*; (c) sufixo aumentativo: *Você viu o **narigão** daquele homem?*; (d) sufixo diminutivo: *Ele mora naquela **casinha** da esquina*. Todavia, Rocha (op. cit.) afirma ser a flexão um mecanismo morfológico e a expressão de grau acaba por ser feita através de expediente sintático. Desse modo, com os sufixos afetivos estaremos formando palavras novas, ou seja, derivando.

Em linhas gerais, apesar de não estarmos discutindo neste trabalho, como já dissemos, as questões sobre flexão e derivação em português, esse debate que inclui também o sufixo -*inho* demonstra que: (a) o grau parece não ser um processo flexional em português; (b) as principais significações veiculadas pelos afixos de grau não são as dimensionais; e (c) há forte teor de subjetividade em tais construções.

Portanto, em certo ponto, concordamos com Piza (*op.cit*) quando admite em seu trabalho que o grau comporta-se como fenômeno limítrofe, haja vista apresentar características tipicamente flexionais, ao lado de outras marcadamente derivacionais.

2.3. Os modos de organização do discurso

Nesta seção, vamos nos distanciar das questões descritivas da língua portuguesa para nos aproximarmos dos tópicos que permeiam a Análise do Discurso/Texto e a Gramática Funcional. Como primeiro tópico faremos uma breve apreciação sobre o que Charaudeau (1992) convencionou chamar de Modos de Organização do Discurso. Esta inclusão tem a ver com a natureza dos componentes de nosso *corpus* eletrônico, compilado a partir de narrativas e descrições.

A atividade de classificação de textos de acordo com a sua natureza comunicativa é prática que data dos Gregos. Em passado mais recente, e sob as marcas da Gramática Gerativa, as classificações se restringiram ao SIL¹¹ como por exemplo, Longacre (1983), Halliday (1985 et seq.) e Werlich (1975).

A classificação de Werlich (*op.cit.*) incluía a seguinte tipologia: texto narrativo, texto descritivo, dissertação argumentativa, dissertação expositiva e o texto injuntivo. Essa classificação influenciou muitos autores como Fávero e Koch (1987), Torreira (1993), Marcuschi (2002) e Adam (1987).

Compreendendo a prática discursiva como uma atividade de quem recorre a procedimentos para usar a língua a fim de cumprir finalidades comunicativas, Charaudeau (1992) vê reunidos esses procedimentos (ou estratégias) em quatro modos: o Enunciativo, o Descritivo, o Narrativo e o Argumentativo. Cada um destes modos tem uma função de base, que é a finalidade discursiva (enunciar, descrever, narrar, argumentar) e um princípio de

¹¹ Summer Institute of Linguistics. Hoje SIL Internacional. O 1º trabalho começou em 1919 com estudos interculturais dos povos Maias da Guatemala. Seus objetivos eram: aprender as línguas, estudar as estruturas fonológicas e gramaticais, entender os contextos culturais, ajudar na confecção de dicionários.

organização. Essa classificação encontra respaldo também em outros autores, porém com adaptações de nomenclatura. Oliveira (2004), adaptando essa classificação, propõe a denominação **modos de organização do texto**, porém acrescentando os modos expositivo e injuntivo¹².

Marcuschi (*op.cit.*:19-36) dá a denominação **tipos de texto** ao que Charaudeau denomina **modos de organização do discurso**, acrescentando também outros dois modos: expositivo e injuntivo. Adam (1987), por sua vez, propõe um método de análise que consiste em dividir o texto em “seqüências”, cada uma das quais enquadrada num *modo de organização do discurso*, embora sem empregar esse termo, que pertence à terminologia de análise semiolinguística¹³ de Charaudeau (*ibid*).

As funções de base do modo enunciativo intervêm na encenação dos outros três modos, na medida em que mostra a relação de influência do locutor em relação ao interlocutor, o ponto de vista situacional¹⁴ do locutor e testemunho deste face ao mundo. A pessoa do discurso é que essencialmente organiza esse modo, tornando-o presente nos outros. O modo enunciativo comanda os outros modos.

Identificar os seres do mundo de maneira objetiva/subjetiva (Charaudeau,1992:642) é a função de base do modo descritivo. Construir a sucessão das ações de uma história no tempo em torno de uma busca para fazer disso um relato, com seus actantes¹⁵ são funções de base do modo narrativo. Explicar uma verdade numa visão racionalista para influenciar o interlocutor são do modo argumentativo.

Um aspecto característico da semiolinguística de Charaudeau (*ibid*) é a relevância do texto¹⁶ na consideração do discurso¹⁷ com o qual não deve ser confundido, já que um mesmo

¹² Expositivo compreende textos com o predomínio de seqüências analíticas ou então explicitamente explicativas. Injuntivo, também chamado instrucional, procedimental, é o modo que explicita as várias etapas da execução de uma tarefa dispostas em ordem cronológica – “primeiro faça isso, depôs faça aquilo” etc.

¹³ Definição dada por Charaudeau (1992) em análise do discurso: *semio*, de “semiosis”, evocando o fato de que na construção do sentido e sua configuração se fazem através de uma relação forma-sentido (em diferentes sistemas semiológicos), sob a responsabilidade de um sujeito intencional, com um projeto de influência social, em um determinado quadro de ação; *lingüística*, para destacar que a matéria principal da forma em questão – a das línguas naturais.

¹⁴ É de ordem psicossocial, externa à linguagem, embora dela participe. É o lugar onde se constrói um contrato de troca verbal.

¹⁵ Serve para designar os diferentes participantes que estão implicados em uma ação e que nela tem um papel ativo e passivo.

¹⁶ Definido neste trabalho como uma entidade concreta realizada materialmente e corporificada em algum gênero textual.

¹⁷ É aquilo que um texto produz ao se manifestar em alguma instância discursiva; assim o discurso se realiza no texto.

tipo de texto¹⁸ (jornalístico, publicitário, literário etc) pode resultar de um ou vários modos de organização do discurso e do emprego de várias categorias da língua.

O discurso está ligado ao texto e vice-versa. A materialidade do texto (verbal, icônica, gestual) é a expressão da encenação do ato da comunicação. Um texto é o resultado de um ato de linguagem produzido por um sujeito numa situação de troca social contratual (Charaudeau, 1997: 85). Cada texto tem especificidades próprias à instância de realização do discurso. Daí existir uma variedade de tipos de texto. Cada um se mostra com um modo dominante, ou com a combinação dos quatro modos. O texto publicitário, por exemplo, pode revelar esses quatro modos, com tendência ao descritivo e ao narrativo.

Como um de nossos *corpora* de análise está subdividido em um modo narrativo e modo descritivo, convém-nos dar uma visão geral das características apenas desses dois modos, excluindo-se, portanto, o enunciativo e o argumentativo.

2.3.1. Modo de organização descritivo

A preocupação em situar a aplicação dos construtos de seu modelo no domínio escolar é uma constante em Charaudeau, que considera um problema a confusão que, na escola, se faz entre descrever e contar. Qual é, portanto, a diferença entre narrar e descrever, já que podemos descrever ações realizadas? Uma receita de cozinha ou um relato de uma reunião de negócios, por exemplo, podem ser incluídos em qual dos dois modos?

Evidentemente existem, nesse problema, que não é novo, dois aspectos a considerar: a finalidade do texto e seu modo de organização. Num texto, por exemplo, cuja finalidade é informar, podem-se encontrar trechos descritivos ou narrativos. Portanto, a situação de comunicação e as diversas ordens de organização do discurso são aspectos dos quais um texto depende. Numa receita de cozinha, reconhece-se que pode haver a descrição de uma seqüência de ações (fazer, embeber, pegar), ou uma seqüência de atos enunciativos (faça, embeba, pegue).

Por último, há o problema da relação Língua/Texto. Charaudeau (1992:565) pergunta se existe um relação de continuidade entre as categorias da língua e as características discursivas de um texto e se apenas a acumulação, num texto, das marcas de uma categoria

¹⁸ Classificação dada por Charaudeau (2002). Outros autores como Marcuschi (2002) e Oliveira (2004) adotam a nomenclatura Domínios Discursivos.

de língua permite determinar um certo modo de discurso. De fato, constatamos que certas categorias freqüentam mais um modo discursivo que outros, por exemplo, a natureza aspectual de certos tempos verbais como (imperfecto/perfeito) pode distinguir o modo narrativo. Os traços de uma possível caracterização resultam da combinação das marcas textuais e das marcas de outras categorias. Portanto, para Charaudeau, o modo de organização descritivo é a expressão escolhida para nomear o processo discursivo distinto da descrição, que é um resultado ou o texto, onde se podem combinar os processos argumentativos, descritivos, narrativos. Segundo o autor, descrever dá vida aos seres, nomeando-os, localizando-os e atribuindo-lhes qualidades. Desse modo, o narrativo ganha sentido e o argumentativo é auxiliado pelo processo descritivo, que se constrói pelos atos de nomear, localizar, situar e qualificar.

Nomear resulta da operação de construção e estruturação da visão de mundo pelo sujeito, que percebe os objetos pelas suas diferenças e semelhanças e, assim prepara-se para classificá-los. O modo descritivo, em razão disso, é o processo pelo qual o sujeito, *sobredeterminado por características culturais* (Charaudeau, *op.cit.*:660), produz taxionomias, inventários listas que decorrem das finalidades explicativa, informativa e incitativas, que, enfim, atendem a propósitos comunicativos. O processo descritivo conduz à identificação dos seres segundo códigos sociais. Localizar-situar faz um recorte objetivo do mundo quando situa um ser no espaço e no tempo. Também esse ato “depende da visão que um grupo cultural projeta sobre esse mundo” (Charaudeau, *op.cit.*: 665).

2.3.2. Modo de organização narrativo

São muitas as correntes teóricas que se ocupam do mecanismo da narrativa. Fala-se em narrativa, história, discurso, narração e narrado, registros de discurso, modalidades etc., termos que podem variar de significação em função do ponto de vista teórico em que se inserem.

Narrar não é somente descrever uma seqüência de fatos e acontecimentos, como dizem os dicionários. Para que haja narrativa, são necessários um narrador dotado de intencionalidade, isto é, de um querer transmitir alguma coisa (uma determinada representação da experiência do mundo) a um narratário, ambos inseridos num contexto. Narrar é, portanto, uma atividade lingüística desenvolvida em meio a um certo número de tensões e mesmo de contradições, como por exemplo, a tensão entre a ficção e realidade.

Segundo Charaudeau (*op.cit.*) existe a “crença na unidade do ser” que produz o que se convencionou chamar de narrativas místicas. É o caso das narrativas cristalizadas, como os textos sagrados das sociedades primitivas ou modernas; das narrativas alegóricas, com os contos populares, as lendas, os contos de fadas, os evangelhos e certos textos fantásticos e das narrativas que idealizam heróis. Segundo autor, há também a “crença numa realidade plural do mundo e do ser” que produz o que se convencionou chamar de narrativas realistas que retratam, de maneira fragmentada, a dura aprendizagem da vida, reconstruindo, o mais objetivamente possível, a realidade passada.

Ainda, de acordo com Charaudeau (*op.cit.*), não há, evidentemente, uma separação extrema entre os dois tipos de narrativas, já que elas são compostas e podem tomar emprestado a esses dois imaginários. Por outro lado, um ou outro desses imaginários pode se revelar dominante na narrativa, segundo os gêneros ou as épocas. A narrativa é uma totalidade; o narrativo, um de seus componentes. Como a narrativa corresponde à finalidade do que “é narrar”, descreve simultaneamente ações e qualificações, daí englobar o modo narrativo e o modo descritivo.

O Modo de Organização Narrativo se caracteriza por uma dupla articulação. Trata-se não só da construção de uma sucessão de ações segundo uma lógica, como também da organização dessas ações em universo narrado. Pode-se chamar a este segundo processo de organização da colocação em narração, a qual constrói o universo narrado propriamente dito, sob a responsabilidade de um sujeito narrador que se encontra ligado, por um contrato de comunicação¹⁹, ao destinatário da narrativa.

Neste trabalho, pode-se dizer que as narrativas encontradas em nosso *corpus* de análise enquadram-se no imaginário “crença numa realidade plural do mundo e do ser” que, como já dissemos, são narrativas realistas que retratam, de maneira fragmentada, a dura aprendizagem da vida, reconstruindo, o mais objetivamente possível, a realidade passada.

2.3.3. O narrativo/descritivo

Segundo Charaudeau (*op.cit.*), descritivo e narrativo se distinguem pelo tipo de visão do mundo que constroem, bem como pelo papel que desempenha o sujeito que narra ou

¹⁹ Princípio segundo o qual, cada um dos parceiros num ato de linguagem está engajado num processo recíproco (mas não simétrico) de reconhecimento do outro, numa interação que o legitima enquanto tal – o que é uma condição para que o ato de linguagem seja considerado válido. Este princípio é o fundamento do aspecto contratual de todo ato de comunicação.

descreve. O descritivo revela um mundo de existência imutável, que necessita apenas ser reconhecido e mostrado, daí se dizer que organiza o mundo de maneira taxionômica (classificação dos seres do universo), descontínua (sem elos necessário entre si, nem de propriedades entre eles) e aberta (sem início nem fim necessários). O narrativo, ao contrário, faz descobrir um mundo a ser construído no desenvolvimento da sucessão de ações interdependentes, que se transformam num encadeamento progressivo; organiza, portanto, o mundo de maneira sucessiva e contínua, dentro de uma lógica cuja coerência é marcada pelo próprio fechamento (começo/fim).

Em relação aos papéis desempenhados pelos sujeitos que narram ou que descrevem, pode-se dizer que o sujeito que descreve desempenha papéis de observador (vê detalhes); de conhecedor (sabe mostrar e evocar). Isso fica patente em nosso *corpus* descritivo, na medida em que aquele que descreve, e cuja descrição aparece em forma de transcrição, faz um esforço para mostrar com palavras a composição dos detalhes da cena descrita. Já o sujeito que narra desempenha essencialmente um papel de testemunha de como as suas experiências estão encadeadas. Nosso *corpus* narrativo é construído por textos narrados em 1ª pessoa, nos quais também há presença de um esforço para tornar a narrativa interessante e verossímil.

Na próxima seção desta revisão bibliográfica, falaremos sobre um outro binômio que permeia a organização de nossos *corpora* eletrônicos: o binômio língua falada versus língua escrita.

2.4. Oposição Língua falada /Língua escrita

Convém aqui, tecer alguns comentários entre língua falada²⁰ e língua escrita, em virtude dos *corpora* de análise serem de natureza oral e escrita, sendo, portanto, viável fazer uma distinção entre essas duas modalidades da língua, mostrando a relevância de ser ter estudos lingüísticos que, de alguma maneira, contrastem as mesmas para que se possa inferir noções relevantes sobre aspectos constitutivos da língua.

²⁰ Concordamos com Travaglia ao definir que “Para alguns estudiosos tanto faz dizer língua oral quanto falada. Para outros a expressão língua oral pode referir-se mais especificamente ao meio sonoro de produção da seqüência lingüística e podemos ter no oral coisas como a leitura de textos escritos. Assim sendo a expressão “língua falada” identificaria melhor essa variedade da língua com suas características próprias, distintas da língua escrita, cujas características permaneceriam na leitura de um texto escrito.” (Travaglia, 2000:51)

Reportamo-nos um pouco a Saussure, que define como objeto de estudo da lingüística a *língua* (Saussure, 1972:28) e não a fala, de modo que uma *língua* é definida como um sistema de elementos. Nesse sentido, define-se sistema como um conjunto organizado em que um elemento se define pelos outros. Um conjunto é uma totalidade de elementos quaisquer e, se eles estão organizados, isso quer dizer que um elemento está em função dos outros, de modo que a sua função se define em relação aos demais elementos do conjunto. A língua é, portanto, um conjunto de signos em que um signo se define pelos demais signos do conjunto.

Saussure (*op.cit.*) estabelece uma dicotomia entre língua *versus* fala, em que *língua* opõe-se à *fala*, porque a *língua* é coletiva e a *fala* é particular; a *língua* é um dado social e a *fala* é um dado individual. Além disso, considera a *língua* sistemática e a *fala* assistemática. Nessa dicotomia, Saussure separa também os fatos da língua dos fatos da fala em que os fatos da língua dizem respeito à estrutura do sistema lingüístico e os fatos da fala dizem respeito ao uso desse sistema. Diz ainda, que os fatos da língua podem ser estudados separadamente dos fatos da fala e considera também as interferências entre esses dois tipos de fatos. Para ele, uma mudança no sistema pode advir de fatos da fala, porém só são pertinentes para o estudo do sistema da língua quando interferem diretamente nas relações internas entre seus elementos sistematizados.

Observa-se, portanto, que, para Saussure, a fala, enquanto prática da língua, está ligada à inteligência e a vontade de cada indivíduo, comportando suas características físicas e psicológicas e as restrições impostas pelo canal e contexto real de comunicação. Tomada em seu conjunto, essa concepção postula a língua como um fenômeno homogêneo, como um produto que deve ser examinado independentemente de suas condições de produção, adotando um modelo teórico de língua como estrutura e a fala (ou seja, língua em uso) como idiosincrática.

Numa abordagem mais contemporânea, Mattoso Câmara (2001:19) considera uma tremenda ilusão a diversidade muito sutil e falaz que se tem entre a fala e a escrita. Segundo o autor, a língua escrita se manifesta em condições muito diversas da língua oral, porém, é somente a escrita que as gramáticas normativas escolares focalizam explícita ou implicitamente. A fala, porém, segundo o autor, se desdobra numa situação concreta, sob o

estímulo de um falante ou vários falantes outros, bem individualizados. Afirma, ainda, que a escrita não reproduz fielmente a fala, como sugere a metáfora tantas vezes repetida de que “ela é a roupagem da língua oral”. Para ele, a língua escrita tem suas leis e caminho próprios, motivo pelo qual muitos lingüistas relegam a língua escrita para fora de suas cogitações. Nesse sentido, “...há uma diferença fundamental entre esses dois tipos de linguagem”. (Mattoso Câmara, 2001:20).

Segundo Castilho (2003) a constituição da língua falada (*parole*) como objeto científico só se deu muito recentemente na Lingüística. Mais importante, ainda, é a importância desse objeto científico para os lingüistas de *corpus* que se deu durante a segunda metade do século XX.

Castilho (*op.cit.*) vê a língua como uma atividade social, por meio da qual veiculamos as informações, externamos nossos sentimentos e agimos sobre o outro. Assim concebida, a língua é um conjunto de usos concretos, historicamente situados, que envolvem sempre um locutor e um interlocutor, localizados num espaço particular, interagindo a propósito de um tópico conversacional previamente negociado. Assim, uma gramática que entenda a língua dessa forma (como é o caso da Gramática Funcional) procura os pontos de contato entre as estruturas identificadas pelo modelo anterior e as situações sociais em que elas emergem, contextualizando a língua no meio social.

Ainda, segundo esse autor, a linguagem humana é fundamentalmente dialógica, mesmo em sua modalidade escrita. Uma diferença, entretanto, é que na língua falada os usuários estão em presença um do outro, e a construção do enunciado se ressent de maneira acentuada da interação que aí se desencadeia. Uma das óbvias conseqüências disso é que na língua escrita é necessário explicitar as coordenadas espaço-temporais em que se movem os interlocutores, ao passo que, na língua falada, tais coordenadas já estão dadas pela própria situação da fala. Nesse sentido, mesmo não tendo sido abordado pelo autor, ressaltamos, porém, que mesmo numa conversação por telefone ou numa troca descrita por e.mail, apesar de os interlocutores não estarem em presença um do outro, pode-se observar algumas das características da língua oral apresentadas acima.

Para Castilho (*op.cit.*), é habitual reconhecer a existência de duas fases constitutivas da linguagem: uma fase de planejamento pré-verbal, de natureza cognitiva, em que selecionamos o que vai ser dito e analisamos as condições da interação para a veiculação do

que vai ser dito, e uma fase de execução, ou fase verbal, em que codificamos através do léxico e da gramática as idéias consideradas adequadas àquele ato da fala. Uma vez engajada a conversação, procedemos constantemente a ajustes no planejamento anterior.

Na fala, portanto, essas fases de planejamento e execução ocorrem simultaneamente, no tempo real. Elas se dão numa situação discursiva plena, isto é, com todos os usuários em presença, o que interfere diretamente na organização e na execução dos atos da fala. Já na língua escrita, a “audiência” tem uma atuação muito discreta, com pouca probabilidade de interferir nessa organização. Nessa modalidade, como observa Pécora (1980:80, *apud* Castilho 2003:19) ”a primeira pessoa é obrigada a desdobrar-se na segunda”. Em consequência disso, conforme Castilho (*ibid*), na língua falada “tudo vai para o ar”, por assim dizer, fazendo dessa modalidade um excelente meio de reflexão sobre os processos constitutivos da língua. Na língua falada nada se apaga, nem mesmo a própria maquinaria da linguagem, permitindo uma inspeção privilegiada.

Convém, ainda, dar uma noção, sobretudo, de situação interlocutiva e situação monolocutiva, segundo Castilho (*op.cit.*), em virtude de essas noções estarem ligadas, de alguma forma, aos conceitos de fala e escrita.

Em situação interlocutiva, há a percepção imediata das reações do interlocutor pelo locutor, que interage com o interlocutor, o que o leva a antecipar o que quer dizer, a hesitar, a se retificar. O locutor pode fazer uso, também, de gestos e mímica para complementar o que está dizendo. Nesse caso, a configuração verbal comporta algumas características: ordem das palavras, considerada afetiva; construção segmentada da seqüência de palavras em acumulação, quase sem laços lógicos; alternância de termos de valor genérico/específico; economia de palavras e superposição de signos. Tudo isso concorre para a construção de uma configuração verbal.

Em situação monolocutiva, o locutor não está à mercê do interlocutor e pode organizar o que vai dizer de maneira lógica e progressiva. Isso implica que a configuração verbal correspondente a essa situação apresente particularidades opostas à situação precedente: ordem das palavras, dita progressiva; construção contínua e hierarquizada; sucessão de termos de valor semântico progressivo; explicitação, ou não, da entonação e dos gestos, no caso de o canal de transmissão ser gráfico.

Segundo Maingueneau (2002), as velhas oposições entre o oral e o escrito, ou entre o escrito manuscrito e o escrito impresso, não mais subsistem atualmente em sua forma tradicional. As técnicas cada vez mais sofisticadas de gravação e de transporte de informações têm modificado os dispositivos de comunicação e, portanto, o estatuto dos enunciados verbais. De acordo com o autor, o mundo contemporâneo caracteriza-se pelo surgimento de novas formas de oralidade que diferem totalmente da oralidade tradicional.

Maingueneau (*op.cit.*) considera que a distinção entre o oral e o escrito é a categoria “midiológica” mais antiga e mais solidamente ancorada na cultura, porém, ainda segundo esse autor, essa distinção não é tão simples quanto poderíamos pensar; ao utilizá-la, misturamos, na realidade, oposições situadas em planos distintos. Essa oposição refere-se aos suportes físicos: o oral se transmite por ondas sonoras e o gráfico, por signos escritos em um suporte sólido (outrora, em argila, papiros, pergaminhos e, atualmente, em papel).

Tradicionalmente, segundo o mesmo autor, associamos oralidade e instabilidade, escritura e estabilidade: as palavras voam o escrito permanece. Porém, na realidade, nem todo enunciado oral é necessariamente instável; tudo depende de seu estatuto pragmático, isto é, daquilo para que serve. Para ele, o importante não é tanto o caráter oral ou gráfico dos enunciados, mas sim, sua inscrição em formas que assegurem sua preservação.

Maingueneau (*op.cit.*) diz, ainda, que a distinção entre oral e escrito envolve igualmente uma distinção entre enunciados *dependentes* e enunciados *independentes* do ambiente verbal. No primeiro caso, os enunciados são dirigidos a um co-enunciador presente no mesmo ambiente físico do enunciador; no segundo, os enunciados são diferidos, ou seja, concebidos em função de um destinatário que se encontra em um outro ambiente. Os enunciados dependentes do ambiente encontram-se sob ameaça constante do co-enunciador, que a todo momento pode intervir na enunciação em curso. Ele pode também dar força à posição do enunciador, expressando sua aprovação (com sua atitude e com comentários: (“Ah!”, “veja só!” etc.).

Essa noção de enunciados dependentes e independentes pode, a nosso ver, ser comparada à noção de Castilho (*op.cit.*) quanto à situação interlocutiva e monolocutiva e, assim como esse autor faz uma distinção entre essas duas situações comunicativas, Maingueneau também aponta algumas características entre os enunciados dependentes e independentes do ambiente.

No enunciado dependente do ambiente, como já dissemos, os co-enunciadores se encontram no mesmo ambiente e estão vendo um ao outro, verifica-se, portanto, segundo Maingueneau, a ocorrência de: “indicadores não verbais (a mímica, os gestos) acompanhando a fala; elipses quando um objeto está presente no ambiente (você viu...?); inúmeros embreantes, cujos referentes são identificados em relação à situação de enunciação (*eu, aqui, amanhã...*)” (Maingueneau, 2002:75). Nos enunciados independentes do ambiente tende a haver uma auto-suficiência, e há a construção de um sistema de referências intratextual (no interior do texto). Esses enunciados não se apóiam em um ambiente partilhado com o co-enunciador, que não pode interferir na enunciação. O autor aponta algumas propriedades evidentes do texto escrito tais como:

- “...ele pode circular longe de sua origem, encontrar públicos imprevisíveis sem precisar ser modificado a cada vez. Como quem escreve não pode controlar a recepção de seu enunciado, é obrigado a estruturá-lo de maneira a torná-lo compreensível, ou seja, é obrigado a fazer de seu enunciado um texto, no sentido mais pleno do termo;
- no oral, o co-enunciador partilha o mesmo ambiente que o locutor, reage imediatamente à sua entonação, às suas atitudes etc. Não podendo percorrer a arquitetura do enunciado em seu conjunto, ele vai tomando conhecimento dele aos poucos e tem uma consciência muito vaga de sua estrutura. No escrito, por outro lado, ele deve proceder a uma leitura pessoal. Se o locutor encontra dificuldade em controlar o curso de uma interação oral, que implica minimamente a participação de pelo menos duas pessoas, no escrito ele pode impor seu modo de consumo, seu ritmo de apropriação: ler com rapidez que lhe convém, silenciosamente ou em voz alta, com atenção ou diagonal, interromper-se quando quiser;
- a distância que assim se estabelece entre co-enunciador e texto escrito abre um espaço para um comentário crítico ou para análises: o leitor pode sondar o texto, comparar certas partes, de forma a elaborar interpretações;
- um texto escrito pode também ser copiado, arquivado, classificado; a estocagem permite confrontar textos variados e estabelecer princípios de classificação (por temas, gêneros, autores, datas etc.)” (Maingueneau, 2002:79).

Travaglia (2000) também contrapõe a linguagem falada e a linguagem escrita, considerando-as como variação de modo. Para o autor, a língua escrita constitui um sistema à parte, com características próprias que marcam como um estilo diferente da língua falada, “de tal modo que alguns autores acham que a dificuldade que os alunos têm para escrever não advém do desconhecimento da norma culta ou padrão, mais sim do desconhecimento dessas características próprias do escrito”. (cf. Perini, 1980 *apud* Travaglia, 2000:52). Para o autor a língua escrita e a oral apresentam cada uma um conjunto próprio de variedades de grau de formalismo.

Travaglia afirma ainda, porém sem ressaltar qualquer evidência empírica, que as variedades de grau de formalismo da língua escrita apresentam uma tendência para maior regularidade e geralmente maior formalidade que as da língua falada, todavia o autor ressalta que em cada caso existe uma mesma relação entre os níveis de grau de formalismo propostos para a língua falada e para a escrita. Também lembra que não é válida a distinção que freqüentemente encontra-se enunciada por professores de que a língua falada seria informal e a escrita formal.

Afirma, portanto, que isso não é verdadeiro, pois podemos ter textos altamente formais na língua falada e textos totalmente informais na língua escrita. O autor também ressalva que a língua escrita pode apresentar variantes dialetais, embora estas sejam usualmente pouco numerosas e menos marcantes que na língua falada, porque no escrito desaparecem as diferenças fonéticas, prosódicas e outras. O autor aponta uma série de diferenças entre essas duas modalidades, as quais, a nosso ver, assemelham-se a dos autores já referenciados.

Ainda, como pressuposto teórico para a relação língua falada/língua escrita podemos citar Marcuschi (2000). O autor faz uma relação entre textos falados e textos escritos, apontando que os mesmos se diferenciam nas dimensões lingüística e funcional²¹. A dimensão funcional é constituída por uma série de parâmetros situacionais a ela associados. Esses parâmetros são estabelecidos a partir de um conjunto de características obtidas das formas prototípicas da fala e da escrita. O autor observa, ainda, que alguns textos orais participam de propriedades relativas a parâmetros da escrita e alguns textos escritos participam de propriedades de parâmetros da fala. Estabelece, portanto, os seguintes parâmetros situacionais segundo Biber (1988:44-46):

(1) canal físico caracterizado como unicanal no caso da escrita e multicanal na fala já que ali há também a prosódia e os elementos para-lingüísticos ao lado dos lingüísticos; (2) uso cultural definido pela natureza da aquisição, valor social e influência no *status*; (3) relação entre os participantes sob o ponto de vista da possibilidade de interação face a face, relações pessoais e grau de conhecimento mútuo; (4) relação dos participantes com o contexto; (5) relação dos participantes com o texto e (6) propósitos primários da comunicação que podem ser ou ideacionais no caso da escrita e interacional no caso da fala. (Biber 1988, apud Marcuschi, 2000: 55)

²¹ O autor entende como dimensão funcional a que diz respeito às características da situacionalidade e discursividade do texto, e a dimensão lingüística a que reporta-se à presença de traços gramaticais, fonológicos e lexicais.

Marcuschi coloca também uma questão instigante, de solução estritamente empírica, que é levantada por Biber (1988:24) quando este autor indaga: “será que existe uma dimensão lingüística de traços co-ocorrentes que distingue os textos escritos dos textos falados?” Se uma tal dimensão existir, poderemos distinguir com precisão entre todos os textos escritos de um lado e todos os textos falados de outro. Biber (*op.cit.*) não obteve, durante esses estudos uma prova empírica que desse uma resposta consistente a essa questão, porém, atualmente, esse tipo de dúvida não é mais pertinente, pois o próprio Biber (1999), com colaboradores, produziu, por exemplo, uma gramática da fala e da escrita da língua inglesa, através de exame de *corpora* eletrônicos, na qual atesta que tais diferenças existem e são quantificáveis. E tudo indica que a diferença entre fala e escrita se dá como um contínuo de diferenças que formam um contínuo de variação entre as duas modalidades.

Baseando-se nos pressupostos acima em relação à língua falada e língua escrita, percebe-se que, para os autores apresentados, há concordância, mesmo que haja algumas diferenças entre terminologias ou conceituações, em dizer que existe uma diferença nítida entre esses dois tipos de modalidade da língua, o que nos fez, portanto, optar por trabalhar com as mesmas, através de dois *corpora*, um oral e um escrito, em virtude das particularidades e diferenças existentes, dando com isso, uma maior objetividade ao nosso estudo para a coleta de fatos da língua, tornando-se de fato promissora a análise das relações entre textos falados e escritos mediante comparação das diferentes formas entre si.

2.5. A noção de contexto

Pelo enfoque teórico deste trabalho, ou seja, por ser, como já dissemos, uma abordagem funcional, convém ressaltar nesta seção, mesmo que de maneira sucinta, alguns aspectos relevantes no que tange à noção de contexto. Abordaremos essas noções baseando-se nas concepções da Gramática Sistêmico Funcional (Halliday, 1985) e Maingueneau (2002).

Em estudos sobre a linguagem, pode-se dizer que o contexto tem sido visto sob diferentes óticas, sendo relacionado a comportamentos, à mente ou à sociedade e, conseqüentemente, em diferentes épocas, têm sido enfatizados aspectos lingüísticos, cognitivos ou sociais. A base dos estudos é ver a linguagem ligada ao contexto social e as escolhas lingüísticas como dependentes deste contexto, dele recebendo influências e ao mesmo tempo podendo influenciá-lo (Halliday e Hasan, 1985).

Segundo Oliveira (2002), a noção de contexto tem sido vista por diferentes autores, sob diferentes aspectos, e no momento atual parece ser difícil encontrar uma definição simples, precisa e técnica para esse conceito, já que esse termo assume significados diversos de acordo com os paradigmas de pesquisa adotados (Schiffrin, 1994:363 apud Oliveira, 2002).²² Entretanto, segundo Oliveira (2002):

“ a falta de uma única definição formal, ou mesmo uma concordância sobre o que significa contexto não é uma situação que tenha que ser corrigida. Ao contrário, o fato de que tantos pesquisadores reconheçam a importância do contexto, e estão envolvidos em tentar desvendar como ele funciona, é precisamente a razão por que ele gera a possibilidade de tantos enfoques para análise”(Duranti e Goodwin, 1994:2 apud Oliveira 2002)²³.

Uma das bases teóricas em que este trabalho está inserido é a Gramática Sistêmico-Funcional (doravante GSF), proposta por Michael Halliday (1985). A GSF vê a língua como fruto do contexto sócio-cultural, tendo seu foco no estudo da linguagem em uso.

A GSF pode ser descrita como uma abordagem semântico-funcional, cuja preocupação é explorar como a língua é estruturada para o uso em diferentes contextos (Eggins, 1994: 23-24). Uma das premissas básicas dessa abordagem é que o uso da língua é motivado pelas relações sociais, e que as escolhas léxico-gramaticais realizadas pelos falantes não são aleatórias, mas estão condicionadas pelo contexto. Assim, uma determinada escolha pode ser influenciada ou determinada, ou ainda, influenciar ou determinar as escolhas à sua volta (Thompson, 1996).

Nessa abordagem, a língua é definida como um sistema que oferece ao falante uma série de possibilidades para expressar os seus significados. A gramática é um constructo operacional que organiza as funções realizadas pelo enunciador; os diferentes significados possuem diferentes formas de expressão, cada uma com diferentes probabilidades de ocorrência, dadas as variações de contexto. Ao realizar um significado através de um item lexical ou um fraseado, o enunciador está realizando uma escolha entre tantas outras prováveis (Halliday, 1985). Isso faz com que, no caso de uma análise, essa venha a ter um caráter contrastivo, pois o pesquisador estará sempre comparando as escolhas realizadas pelo

²² Schiffrin, D. (1994). *Approaches to Discourse*. Oxford: Basil Blackwell.

²³ Duranti, A. & Goodwin, C. (1994). *Rethinking Context*. Cambridge University Press.

enunciador com outras disponíveis, de forma a determinar quais os fatores contextuais que levaram o enunciador a optar por “a” e não por “b”.

Segundo Halliday (1985), ao optar por um determinado fraseado, o falante está realizando três tipos de significados simultaneamente:

1. significados relativos às representações de poder e solidariedade, o que engloba as atitudes em relação ao outro e os papéis sociais assumidos;
2. significados relativos à organização do conteúdo da mensagem, relacionando o que se diz ao que foi dito;
3. significados relativos à representação da experiência através da língua, sobre o que se fala e sobre quem ou quem faz o quê para quem.

Cada um desses tipos de significados está relacionado com uma das metafunções da linguagem, sendo que cada metafunção possui um sistema que viabiliza a realização de seus significados. Assim, o item 1 está relacionado à Metafunção Interpessoal; o item 2 à Metafunção Textual e o item 3 à Metafunção Experiencial.

2.5.1. O Contexto de cultura e o contexto de situação

Para a GSF, qualquer realização lingüística está condicionada ao contexto no qual foi produzida. Para Halliday e Hasan (1989:05), o contexto é um fator determinante na produção de um texto, assim, como a situação o é em relação ao discurso nele empregado. Para os systemicistas, o contexto parece “entrar” no texto de forma a influenciar escolhas de palavras e de estruturas (Martin e Eggins, 1997). Assim, os estudiosos da GSF estabelecem dois níveis para o contexto, o *Contexto de Situação* e o *Contexto de Cultura*, os quais funcionam como ferramentas para o entendimento de um texto. Nessa abordagem, para que se possa analisar um texto, é indispensável que seja observado aquilo que acontece à sua volta, encarando o texto e o contexto como interdependentes. (Thompson, 1996).

O Contexto de Cultura pode ser entendido como o pano de fundo onde a interação está inserida, disponibilizando um potencial de realização de significados (potencial semiótico). Dessa forma, Eggins (1994) diz que o estudo do Contexto de Cultura envolve a observação de como a língua é estruturada para o uso. Para tanto, é necessário estudar interações autênticas e completas.

Assim, quando estudamos o Contexto de Cultura, estamos procurando descrever como o propósito geral da interação nos leva a organizar um texto em estágios, procurando alcançar objetivos previstos em nossa cultura. Dentro desse contexto, são estudadas as diferentes manifestações textuais, mostrando como cada um deles se articula. O estudo do Contexto de Cultura passa pela observação e análise dos diferentes gêneros. Na GSF, um gênero pode ser definido como um evento comunicativo, formado por estágios, e que leva à realização de um objetivo social específico, sendo incompleto caso sua realização não ocorra. Essa teoria sugere que textos que realizam diferentes funções em uma cultura serão organizados de forma diferente; uma cultura é vista como um sistema de gêneros (Martin: 2000).

O Contexto de Situação (ou registro) cria significados através de quem somos, o que fazemos e o que dizemos. Seria improvável (ou pelo menos risível) que cariocas produzissem, por exemplo, o seguinte enunciado: *vai ter macaxeira pro café* ou para aqueles falantes que já passaram do 50 anos produzirem frases como: *Mermão, to bolado!*

Contexto de Situação, portanto, está relacionado à situação imediata de realização do texto. Halliday (Halliday e Hasan, 1989) define três variáveis de registro, conforme apresentadas abaixo:

1. **Campo:** o que acontece, a natureza da ação social, aquilo em que os falantes estão engajados;
2. **Teor (relações):** quem está participando do evento e com qual função, quais são suas relações de solidariedade. Estas relações são estabelecidas em três níveis: poder (quem faz o que dentro das relações interpessoais), afeto (o gostar, desgostar, a neutralidade) e o contato (frequência, duração e intimidade do contato social).
3. **Modo:** qual papel da linguagem, refere-se a como a língua está sendo usada (modo de ação ou reflexão, modo falado ou escrito).

Estudos sobre registro estão preocupados em observar como o texto está relacionado ao seu contexto imediato de realização. As variáveis de Modo, Teor e Campo são mapeadas de forma a explicar as razões que levaram o enunciador a realizar determinadas escolhas. Para Halliday (1985), qualquer alteração em qualquer elemento dessa tríade, ocasiona uma modificação no texto como um todo.

Pode-se dizer que há uma correlação direta entre as Metafunções e as variáveis de Contexto de Situação.

Ainda, em relação ao contexto, apresentaremos também algumas concepções de Maingueneau (2002), as quais consideramos também, importantes para este estudo. Segundo esse autor, a reflexão contemporânea sobre a linguagem afastou-se da concepção de se considerar, geralmente, que cada enunciado²⁴ é portador de um sentido estável, a saber, aquele que lhe foi conferido pelo locutor. Esse mesmo sentido seria decifrado por um receptor que dispõe do mesmo código, que fala a mesma língua.

Nessa concepção da atividade lingüística, o sentido estaria de alguma forma inscrito no enunciado, e sua compreensão dependeria essencialmente de um conhecimento do léxico e da gramática da língua; o contexto desempenharia um papel periférico, fornecendo os dados que permitem desfazer as eventuais ambigüidades²⁵ dos enunciados. Na reflexão contemporânea, porém, o contexto não se encontra simplesmente ao redor de um enunciado que conteria um sentido parcialmente indeterminado que o destinatário precisaria apenas especificar.

Com efeito, todo ato da comunicação é assimétrico: a pessoa que interpreta o enunciado reconstrói seu sentido a partir de indicações presentes no enunciado produzido, mas nada garante que o que ela reconstrói coincida com as representações do enunciador.

Compreender um enunciado não é somente referir-se a uma gramática e a um dicionário, é mobilizar saberes muito diversos, fazer hipóteses, raciocinar, construindo um contexto que não é um dado preestabelecido e estável. A própria idéia de um enunciado que possua um sentido fixo fora de contexto torna-se insustentável (Maingueneau, 2002:20).

Ainda, para Maingueneau (*ibid*), não se quer dizer que as unidades lexicais de uma seqüência verbal não signifiquem nada, nem que suas relações deixem de orientar de maneira decisiva a interpretação. O que se quer dizer é que:

fora do contexto, não podemos falar realmente do sentido de um enunciado, mas, na melhor das hipóteses, de coerções para que um sentido seja atribuído à seqüência verbal proferida em uma situação particular, para que esta se torne um verdadeiro enunciado, assumindo em um lugar e em um momento específicos, por um sujeito

²⁴ **Enunciado** se opõe à **enunciação** da mesma forma que *produto se opõe ao ato de produzir*; nesta perspectiva, o enunciado é a marca verbal do acontecimento que é a enunciação. (Maingueneau, 2002: 56)

²⁵ Se dissermos por exemplo, “O cachorro late” ou “Ela está acesa”, o contexto serviria apenas para determinar se “o cachorro” designa um cão particular ou a classe dos cães; a quem o pronome “ela” se refere e se “acesa” é um estado (*a lâmpada está acesa*) ou a um comportamento (*a criança está acesa*)* etc.

que se dirige, numa determinada perspectiva, a um ou vários sujeitos” (Maingueneau, 2002: 20).

O autor aponta, ainda, como tipo de *contextos*, o *cotexto* como sendo as seqüências verbais encontradas antes ou depois da unidade a interpretar que podem fornecer alguns elementos necessários para a interpretação. O recurso ao cotexto mobiliza a memória do intérprete, que vai colocar uma dada unidade em relação a uma outra do mesmo texto. Ressaltamos que é exatamente isso o que faz a Linguística de Corpus: o cotexto é uma pista para entendermos o texto e o contexto.

Segundo o autor, *a priori* nunca há uma única interpretação possível para um enunciado e é preciso explicar quais os procedimentos do destinatário para chegar à interpretação mais provável, que será aquela que se deve preferir em tal ou qual contexto.

A partir dos pressupostos apresentados neste subitem, este trabalho se coloca na posição de fazer uma análise de fatos da língua, em nosso caso, as formações *x-inho*, baseando-se nos contextos de produção, ainda assim, pelo fato de concordarmos com Maingueneau (*op.cit*) ao considerar que “...o conhecimento de língua está longe de ser suficiente para interpretar um enunciado, sendo preciso recorrer a procedimentos pragmáticos ligados ao contexto” (Maingueneau, *op.cit.*:29).

CAPÍTULO 3 - PROCEDIMENTOS METODOLÓGICOS

Neste capítulo, será apresentada a metodologia de pesquisa adotada neste estudo. Primeiramente, serão reiterados os objetivos que norteiam o trabalho. A seguir apresentaremos a Lingüística de Corpus de forma histórica e conceitual calcada em Berber Sardinha (2004), além de enfatizar algumas informações importantes que embasam o trabalho como um todo. Em seguida, será traçado o perfil do *corpus* de estudo e, finalmente, serão descritos os procedimentos metodológicos de análise de dados, bem com as ferramentas computacionais utilizadas nesta investigação.

3.1. Lingüística de Corpus

3.1.1. Breve Histórico da Lingüística de Corpus

Segundo Berber Sardinha (2004), a tradição de observarmos os fenômenos lingüísticos a partir da compilação de dados empíricos de linguagem remonta há muitos anos e, dentre os vários exemplos, podemos citar o trabalho de Kading em 1897. Esse autor, utilizando um *corpus* de alemão contendo 11 milhões de palavras, verificou a freqüência da distribuição e a seqüência das letras no idioma. Na metade do século XX, concomitantemente às preocupações chomskyanas sobre “competência” (língua e não a fala), Randolph Quirk liderou a fundação do “*Survey of English Group*” em 1959, tendo iniciado a compilação do último grande *corpus* coletado de forma manual (posteriormente digitalizado). A partir das amostras de linguagem obtidas, Quirk e seus colaboradores criaram, dentre mais de 200 publicações, gramáticas de suma importância como *A Grammar of Contemporary English* de 1972 e *A Comprehensive Grammar of Contemporary English* de 1985.

No final dos anos 50 com o lançamento de *Syntactic Structures*, Chomsky prepara a lingüística para um novo paradigma no qual a intuição do falante nativo é suficiente para a análise da linguagem, não sendo necessário coletar dados de outros falantes. O impacto causado pela onda Chomskyana e sua premissa de desvinculação entre a análise lingüística e a observação de dados empíricos, obscureceu o lançamento do primeiro *corpus* eletrônico de linguagem escrita do mundo, o *corpus* Brown.

Com a invenção dos computadores nos anos 80, aumentou a capacidade de analisar e armazenar *corpora* maiores e de se executarem tarefas mais complexas. Conseqüentemente

diminuíram as possibilidades de erros na verificação dos dados e reacendeu-se a visão empirista da linguagem, dando origem ao que conhecemos modernamente como Lingüística de Corpus. O desenvolvimento dos computadores e Lingüística de Corpus estão intimamente ligados, uma vez que grandes quantidades de dados necessitam ser analisados rápida, correta e eficazmente, além de prestarem-se à verificação de outros pesquisadores.

Pode-se dizer que o surgimento da Lingüística de Corpus está relacionado com o aumento da capacidade dos computadores em armazenar dados e com o crescimento dos tamanhos dos *corpora* (Leech, 1991) e que teve como ponto de partida o processamento eletrônico de grandes amostras de linguagem, sendo o *corpus* Brown, com um milhão de palavras do inglês americano, seu marco inicial.

Depois desse, houve a compilação de um número cada vez maior de *corpora* eletrônicos com tamanhos cada vez maiores. Entre eles, Berber Sardinha (*op.cit.*) destaca o LOB (Lancaster-Oslo-Bergen) de 1978, com um milhão de palavras, composto por textos escritos em inglês britânico, que consiste de 500 amostras de textos subdivididos em quinze categorias da língua escrita, e também o Birmingham *Corpus*, que em 1987 já tinha vinte milhões de palavras e hoje possui mais de 480 milhões, provenientes do inglês britânico escrito e falado. Esse *corpus*, chamado de Bank of English, é um dos projetos mais representativos de *corpus* de monitoramento da linguagem, tendo sido levado a cabo pelo projeto Cobuild, liderado por John Sinclair. Esse tipo de *corpus*, como o Bank of English, é relevante para pesquisas na área de lexicografia, já que pode proporcionar ao pesquisador a oportunidade de observação da ocorrência de novas palavras na língua, a morte ou o acréscimo no uso de tantas outras ou a observação de novos significados de palavras já conhecidas. Berber Sardinha (*op.cit.*) aponta também outro *corpus* contemporâneo de dimensões grandes é o BNC (British National *Corpus*) de 1995 com cem milhões de palavras do inglês britânico tanto falado como escrito.

Para Leech (1991), desde os anos 60, a Lingüística de Corpus tem aumentado gradualmente seu campo de atuação e influência e quase se tornado uma corrente lingüística em si própria. O autor destaca entre as principais vantagens a possibilidade que a Lingüística de Corpus tem de encontrar informações específicas e testar hipóteses lingüísticas. Para Leech (*op.cit.*), nenhum *corpus* contém todas as informações lingüísticas. Entretanto, conforme coloca Fillmore (1992, *apud* Berber Sardinha, 2000) “...não há nenhum *corpus* que contenha todas as informações que eu quero” mas “todo *corpus* me ensinou coisas sobre a

linguagem que eu não teria descoberto de nenhum outro modo”. Nesse sentido, o pesquisador pode deparar-se com um mundo de informações e fatos novos que não foram previamente hipotetizados por ele , revelando características da língua com muito mais credibilidade.

3.1.2. Conceitos centrais à Lingüística de Corpus

Conforme mencionado na Introdução deste capítulo, a Lingüística de Corpus é uma área que estuda a linguagem por meio de grandes quantidades de dados empíricos, mediante o uso de computador. Convém destacar, portanto, o que se considera por observação empírica. Sob um ponto de vista filosófico, empirismo é uma abordagem ligada à psicologia, declarando que o desenvolvimento da teoria deve estar relacionado a fatos e experiências observáveis, ou que todo conhecimento humano advém da experiência. (Dictionary of Language Teaching and Applied Linguistics, 1992:123).

Um contraste entre empirismo e racionalismo pode melhor esclarecer a definição acima. A teoria racionalista baseia-se em comportamentos artificiais e em julgamentos introspectivos, assim como em afirmações teóricas baseadas em reflexões sobre a fala emitidas pelo próprio pesquisador. Portanto, a grande diferença entre racionalismo e empirismo consiste na observação dos dados, sejam eles artificialmente induzidos ou naturalmente ocorrentes, respectivamente. Assim, a abordagem empírica serve de base para a Lingüística de Corpus, uma vez que a própria palavra *corpus*, nesse contexto, significa coleção de textos que ocorrem naturalmente.

Convém destacar, ainda, como já dissemos, a importância e a necessidade do uso de computadores na Lingüística de Corpus, pois esses trazem como principais vantagens para a investigação da língua, a rapidez e a precisão das informações obtidas sendo possível a utilização de programas como os concordanciadores²⁶, em que o usuário pode fazer a busca de uma palavra específica, trazendo todas as ocorrências da palavra selecionada no *corpus* de estudo e a observação das palavras que co-ocorrem com essa forma selecionada . Além disso, há a possibilidade de anotação no *corpus*, isto é, o acréscimo de informações como por exemplo, se uma palavra pertence à classe gramatical dos verbos, substantivos ou preposições e anotações dos mais diferentes tipos, dependendo da pesquisa que está sendo realizada²⁷.

²⁶ O Concordanciador WordSmith será discutido mais adiante na seção 3.3

²⁷ Um *corpus* sem anotação é considerado “cru”.

Kennedy (1998) também adiciona à utilização de computadores, a rapidez, a exatidão, a confiabilidade estatística e a habilidade de manusear grandes quantidades de dados. Com isso, os computadores têm permitido aos lingüistas trabalhar com uma grande variedade de textos e, dessa forma, procurar generalizações sobre a língua, não restritas ao uso de textos específicos ou à intuição de um lingüista. A quantificação da língua através de estudos baseados em *corpus* tem ajudado a renovar ou fortalecer a ligação entre a descrição lingüística e outras áreas de estudo, como, por exemplo, o ensino de línguas estrangeiras (Granger, 1998).

Assim, para os estudiosos desse ramo da Lingüística, *corpus* não é simplesmente uma coleção de textos usada como base para se desenvolver uma pesquisa de análise ou descrição lingüística.

Berber Sardinha (2004:18) cita a definição de “corpus”, feita por Sánches (1995:8-9) como a mais completa:

Um conjunto de dados lingüísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso lingüístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise.

Ainda segundo Berber Sardinha (*Ibid*), trata-se da definição mais completa por contemplar vários pontos importantes, quais sejam

- “(a) a origem: os dados devem ser autênticos
- (b) o propósito: o corpus deve ter a finalidade de ser um objeto de estudo lingüístico
- (c) a composição: o conteúdo do corpus de ser criteriosamente escolhido
- (d) a formatação: os dados do corpus devem ser legíveis por computador
- (e) a representatividade: o corpus deve ser representativo de uma língua ou variedade
- (f) a extensão: o corpus deve ser vasto para ser representativo”

Já para McEnery e Wilson (1996) o termo *corpus* deve considerar quatro fatores principais: em primeiro lugar, amostra e representatividade, isto é, na construção de um *corpus* deve-se considerar que a amostra seja representativa das tendências de uma

determinada variedade da língua, incluindo suas proporções. Esse último item, como veremos na discussão de nossos resultados, é particularmente importante.

A segunda característica destacada pelos autores é quanto ao fato de ser legível por computadores, sendo este fator considerado indispensável para a investigação da língua, uma vez que através desse instrumento é possível investigar informações sobre traços lingüísticos de áreas diversas como traços morfológicos, morfo-sintáticos e semânticos.

Outro fator relevante destacado pelos autores é a possibilidade que a pesquisa realizada através de *corpus* eletrônico pode trazer quanto a uma descrição objetiva da língua, além da possibilidade de descoberta de fatos novos. A última característica destacada pelos autores é a referência padrão, isto é, um *corpus* é considerado como referência padrão da variedade da língua que se representa.

Assim, vários *corpora* têm sido coletados com diferentes propósitos, representatividades e formatos. Dentre os diferentes tipos de *corpora* encontrados, destacam-se o *corpus* geral e o especializado. Um *corpus* geral é uma coletânea de textos utilizada para explorar a linguagem e é também usado como base para respostas a perguntas específicas sobre vocabulário, gramática ou estrutura do discurso de uma língua como um todo. Deve, portanto, conter textos de diferentes gêneros e domínios, incluindo textos falados e escritos, particulares ou públicos (Kennedy, 1998:20).

Já um *corpus* especializado é desenvolvido para atender às necessidades específicas de um trabalho de pesquisa em particular, de acordo com seus objetivos, como por exemplo, uma coletânea de textos de propaganda de escolas de línguas, coletado com o intuito de se verificar a variedade lexical e gramatical mais utilizada por esse gênero. Assim, um *corpus* especializado geralmente define claramente a população estudada, os objetivos de pesquisa, o gênero ou gêneros contidos no *corpus* e pode especificar o período de tempo de sua compilação. Dessa maneira, os resultados obtidos não caracterizarão a língua estudada como um todo, apenas em relação ao *corpus* de estudo específico (McEnery & Winson, 1996:64). Cabe ressaltar que nosso *corpus* escrito (toda a Folha de São Paulo publicada durante o ano de 1994) pode ser caracterizado como *corpus* geral por conter textos escritos²⁸ não específicos a

²⁸ O *corpus* oral, no entanto, é proveniente do *corpus* Discurso & Gramática e foi coletado a partir de uma pré-divisão dos participantes em idade, sexo e nível intelectual.

determinados grupos ou gêneros discursivos, sendo, portanto, passível de se explorar alguns aspectos da linguagem, inferindo-se generalizações e particularidades.

Sinclair (1991) também identifica a linguagem como um sistema probabilístico em sua pesquisa. Esse autor não dissocia as ocorrências lexicais das gramaticais. Para ele, há um vínculo entre suas ocorrências, isto é, um padrão que carregará o sentido do que está sendo enunciado e que dependem das escolhas selecionadas pelo falante.

A preocupação da Lingüística de Corpus está em proporcionar evidências de que determinados itens ou estruturas, sejam eles lexicais e/ou gramaticais, apesar da possibilidade de sua ocorrência, conforme proposto pela teoria de Chomsky (1957), não ocorrem com a mesma frequência ou aleatoriamente, apresentando-se, dessa forma, ligados ao contexto de sua ocorrência. Assim, tais possibilidades têm sido investigadas pela Lingüística de Corpus e suas descobertas crescentemente utilizadas em diferentes pesquisas de diversas áreas da lingüística.

Kennedy (1998:8) aponta para a ênfase de estudos que priorizam a probabilidade de ocorrência de estruturas e itens lexicais, dividindo os pesquisadores da Lingüística de Corpus em quatro áreas: 1) aqueles que coletam e produzem *corpora*; 2) aqueles que desenvolvem ferramentas para a análise dos *corpora*; 3) os lingüistas descritivos, os quais usam os *corpora* computadorizados para descrever a linguagem; e 4) aqueles que aplicam as descrições lingüísticas baseadas em *corpus* em diversas áreas, como o ensino de línguas e a tradução entre outras. A pesquisa aqui apresentada encaixa-se nos itens 3 e 4 acima.

Como se pode observar, a Lingüística de Corpus pode ser utilizada para uma enorme variedade de pesquisas lingüísticas, tais como “pesquisas sobre palavras isoladas, características gramaticais, linguagem de homens e mulheres, aquisição de linguagem por crianças, estilo de determinado autor, padrões de registro” (Biber, et al, 1998:11)

Ainda de acordo com Biber, a Lingüística de Corpus pode ser utilizada em qualquer área de investigação lingüística empírica como:

- Lexicografia – associações lingüísticas e não lingüísticas. Além de fornecer os significados das palavras, os dicionários podem apresentar os usos mais comuns, a frequência e o contexto de determinadas palavras (Biber et al, 1998:11). Como o Longman Dictionary of Contemporary English, baseado do BNC.

- Gramática – como a gramática de referência desenvolvida pelo próprio Biber: *Longman Grammar of Spoken and Written English (1999)*.
- Sociolinguística – dialetologia e registro, co-ocorrências de padrões em diferentes registros.
- Aquisição da linguagem.
- Estilística – estilos de diferentes autores podem ser investigados através dos períodos históricos, assim como o desenvolvimento desses registros.
- Linguística educacional – confecção de materiais e atividades para colocar o aprendiz em contato com a língua em uso.

3.1.3. Metodologia ou abordagem?

Uma das questões amplamente discutidas dentre os teóricos da Linguística de Corpus relaciona-se ao status desta. A Linguística de Corpus é simplesmente uma base metodológica ou uma área de pesquisa? (Tognini-Bonelli, 2001).

McEnery e Wilson (1996:02) não consideram a Linguística de Corpus como sendo um ramo da linguística, à maneira da sintaxe ou da sociolinguística, pois tais disciplinas descrevem/explicam algum aspecto do uso da linguagem, enquanto que a Linguística de Corpus parece funcionar como uma metodologia que pode ser usada em quase todas as áreas sem se constituir como uma área em si mesma.

Já Biber, Conrad e Reppen (1998:09) conferem à Linguística de Corpus o status de abordagem complementar e recomendam que ela seja utilizada concomitante a abordagens mais tradicionais. Nesse sentido, pode-se enquadrar esta dissertação a qual parte de pressupostos teórico-gramaticais sobre as formações *x-inho*, confrontando-os com uma análise mais funcional através de uma abordagem baseada na Linguística de Corpus.

Kennedy (1998:07), oferecendo um outro posicionamento, considera que a Linguística de Corpus não deve ser vista como uma teoria de linguagem contrapondo-se a outras como, por exemplo, a Gramática Transformacional, mas sim como uma fonte de evidência que pode servir a qualquer teoria linguística que aceite a riqueza de dados empíricos. O autor também

determina que o escopo da Lingüística de Corpus pode ser definido em termos do que as pessoas fazem com os *corpora*.

Por fim este trabalho corrobora a posição de Hoey (1997), citado por Berber Sardinha (*op.cit.*), que considera a Lingüística de Corpus como uma “forma de se chegar à linguagem”, uma abordagem e não somente um instrumental.

3.1.4. Corpus-driven e corpus-based

Pelas características teóricas e metodológicas deste trabalho, convém, nesta subseção, definir dois tipos de abordagens que norteiam a Lingüística de Corpus as quais serão definidas abaixo, sendo, portanto, uma dessas abordagens de fundamentação importante para esta dissertação.

Um corpus pode ser usado de diferentes maneiras para validar, exemplificar ou construir uma teoria da linguagem. Diferentes termos têm sido usados por diferentes estudiosos, mas todos estão centrados em torno de uma distinção básica. O termo mais freqüentemente usado é “corpus-based”, em oposição a “corpus-driven”.

3.1.4.1. Abordagem conduzida por *corpus* (corpus-driven)

Na abordagem conduzida por *corpus* (“corpus driven”), o lingüista usa um *corpus* além da seleção de exemplos, para sustentar um argumento lingüístico ou para validar um discurso teórico. Nesta abordagem, o compromisso do lingüista é com a integridade dos dados como um todo, e a descrição de qualquer fenômeno lingüístico tem o objetivo de ser abrangente o suficiente com relação à evidência do *corpus*. O *corpus* é, portanto, visto como muito mais que um depósito de exemplos ou um arquivo de textos para apoiar teorias já existentes. Qualquer teoria tem de refletir diretamente a evidência extraída do *corpus* propriamente dito.

A novidade desta abordagem pode não ter sido prevista 20 anos atrás, pois não havia consciência da necessidade de construirmos uma abordagem direcionada por *corpus*, e os pequenos *corpora* da época produziam grande quantidade de informações interessantes quando vasculhados através da estratégia baseada em *corpus* (Johanson & Hofland). Foi apenas gradualmente que os pesquisadores reconheceram que a descrição tradicional de pré-corpus da linguagem estava sendo implicitamente questionada pela evidência de *corpora* maiores. O projeto Cobuild, já descrito na seção 3.1.1, que reflete a posição de Sinclair na

teoria da linguagem e na metodologia descritiva pode ser visto como o primeiro estudo em *corpus* dirigido à lexicografia.

O trabalho de *corpus* possui uma abordagem empírica à descrição do uso da língua, assumindo um contexto, a teoria funcional dos significados e fazendo uso de novas tecnologias. Esta definição não se aplica necessariamente à abordagem baseada em *corpus* (“*corpus-based*”), que veremos adiante, onde o relacionamento entre um item e seu contexto não é tomado como sistemático e determinante na definição das categorias lingüísticas. A abordagem direcionada por *corpus* (“*corpus-driven*”), por outro lado, combina bem com a definição que objetiva derivar categorias lingüísticas sistematicamente de padrões recorrentes e a distribuição de freqüências que emerge da língua em contexto.

A plataforma de pesquisa estabelecida por Sinclair como direcionada por *corpus* tem começado a publicar trabalhos referência de gramática – *The Cobuild Grammar Patterns*, que apresenta seleção em nomes, verbos e adjetivos, mostrando uma semântica correlacionada, o que era obscurecida quando a perspectiva era exclusivamente gramatical.

3.1.4.2. Abordagem baseada em *corpus* (“*corpus-based*”)

Estudos “*corpus-based*” é reivindicação de muitos trabalhos que se relacionam ou são extraídos de um *corpus*. O termo “*corpus-based*” é usado para fazer referência a uma metodologia que se utiliza prioritariamente de *corpus* para explicar, testar ou exemplificar teorias e estudos da linguagem.

Pode-se dizer que os lingüistas “*corpus-based*” têm uma relação de confiança entre dados e teorias que eles acreditam ser fundamentalmente adequadas. Eles distinguem e analisam *corpus* através dessas categorias e peneiram os dados. O *corpus* é considerado importante porque, em algumas ocasiões, indica onde devem ser feitas correções secundárias e ajustes no modelo adotado e, naturalmente, pode ser valioso como fonte de evidência quantitativa. Estudos “*corpus-based*” servem para refinar categorias pré-existentes, mas nunca está realmente em uma posição de desafiá-las.

Portanto, as duas metodologias às quais estamos nos referindo em relação ao trabalho com *corpus*, ou seja, “*corpus-based*” e “*corpus-driven*”, refletem duas posições opostas deste assunto. Enquanto o lingüista “*corpus-based*” testa uma teoria tentando separar, unificar e reduzir os dados; o lingüista “*corpus-driven*” constrói categorias a partir dos dados.

Nossa dissertação, por seu escopo e enfoque pode ser denominada de “corpus-based”. Partiremos de algumas categorias existentes em relação à formação do sufixo *-inho* e testaremos essas categorias em evidência proporcionada por dois *corpora*.

3.2. Tamanho e representatividade de um corpus

Assim como o modo de abordar dados varia, o tamanho dos *corpora* tem variado bastante. Por exemplo, há *corpora* pequenos, com poucos milhares de palavras, até coleções com dezenas de milhões de palavras. A construção de um *corpus* com um milhão de palavras, foi, no passado, um sonho e esse tamanho já foi substituído por *corpora* bem maiores, como o BNC e o COBUILD. Ao mesmo tempo, a importância dos *corpora* pequenos está sendo cada vez mais enfatizada como fontes legítimas e informativas de análise lingüística e de aplicação pedagógica (Ghadessy *et. Al*: 2001).

Porém, pouco se tem concluído quanto aos critérios mínimos para definição do tamanho ideal de um *corpus* geral de uma língua para que ele seja representativo. Berber Sardinha (2004) sugere uma classificação baseada na observação dos *corpora* utilizados segundo quatro anos de conferência de Lingüística de Corpus, de acordo com a tabela 1 abaixo:

Tamanho em palavras	Classificação
Menos de 80 mil	Pequeno
80 a 250 mil	Pequeno-médio
250 mil a 1 milhão	Médio
1 milhão a 10 milhões	Médio-grande
10 milhões ou mais	Grande

Tabela 1: Classificação relativa do tamanho de corpus de Berber Sardinha (2004:346)

Tal classificação está baseada na observação dos *corpora* efetivamente empregados pela comunidade de lingüistas de corpus segundo estudos apresentados nas principais conferências de Lingüística do Corpus. (Berber Sardinha, 2004).

O tamanho e representatividade de um corpus têm estado totalmente associados. Sinclair (1991) afirma que quanto maior for o corpus, mais representativo da língua em uso ele será. Porém, a compilação de um corpus geral totalmente representativo de uma língua é considerada estritamente impossível, pois o tamanho de um corpus que tenha essa pretensão é infinito, uma vez que a língua está constantemente mudando – palavras novas são criadas e a frequência do uso de certas expressões varia com o decorrer do tempo e o contexto.

Enfatizamos, portanto, que a contribuição de Chomsky para a Lingüística de Corpus está exatamente onde se baseavam suas críticas, no fato de a língua não ser finita. Segundo Chomsky (1957), um *corpus* não pode representar uma língua em sua totalidade; conseqüentemente, ele não pode ser representativo dessa língua. Porém, para se estudar o infinito, um pesquisador cria uma amostra. Essa amostra, no caso da lingüística, é o *corpus* de estudo. Além disso, assim o faz, porque a linguagem, como já dissemos no capítulo 1, é probabilística.

3.3. Seleção e descrição dos *corpora*

Acreditamos que as estruturas lingüísticas obedecem às intenções e objetivos do falante que com elas constróem e/ou reiteram o seu discurso para informar e/ou tentar convencer o seu interlocutor, com que interage em maior grau.

Entendemos que qualquer porção significativa em *x-inho* atribui-lhe significado funcional que, por sua vez, é indexado à língua como reflexo do ato da fala. Por essas razões, tentamos descrever a situação comunicativa em que a formação *x-inho* ocorre e, para isso foram utilizados dois tipos de *corpora*, um de natureza oral e outro de natureza escrito, para que pudéssemos levantar a frequência e distribuição dessas formações em cada corpus e fazer uma análise contrastiva entre os mesmos para se verificar a relevância dessas formações nessas duas modalidades da língua: oral e escrita.

O primeiro *corpus* utilizado para esta pesquisa, é o conjunto de entrevistas que compõem o *Corpus Discurso & Gramática* (Votre e Oliveira: 1995), constituído de entrevistas orais e escritas de alunos de CA infantil, CA supletivo, 4º série do ensino

fundamental, 8º série do ensino fundamental, 3º série do ensino médio e último ano do ensino superior com idades entre 17 e 42 anos e de ambos os sexos. As entrevistas basearam-se nos seguintes Modos de Organização:²⁹

- Narrativa de experiência pessoal: em que se pede ao informante que narre uma história triste, alegre ou interessante que tenha ocorrido com ele.
- Narrativa recontada: em que se pede ao informante que narre uma história triste, alegre ou interessante que tenha ocorrido com alguém que ele conheça.
- Descrição de local: em que se pede ao informante que descreva um lugar da sua casa onde ele mais goste de estar ou um lugar que ele goste de passear ou brincar.
- Relato de procedimento: em que se pede ao informante se ele sabe fazer alguma coisa, como cozinhar, por exemplo, e pede-se a ele que ensine ao entrevistador.
- Relato de opinião: em que se pede ao informante que dê sua opinião referente a algum assunto ligado ao interesse do aluno.

Em função dos objetivos desta pesquisa, formulamos a hipótese de que as formações *x-inho* ocorreriam em maior profusão na modalidade oral, principalmente, nos modos narrativo e na descritivo.

Para tanto, transformamos esse *corpus* principal em dois *subcorpora* eletrônicos, isto é, foram compilados um *corpus* de respostas de natureza narrativa e outro *corpus* de respostas predominantemente descritivas, contemplando os Modos de Organização do Discurso (Narrativo e Descritivo), segundo a classificação de Charaudeau (1992), explicitada no capítulo 2. Consideramos como Modo de Organização Narrativo a “narrativa de experiência pessoal” e a “narrativa recontada” e como Modo de Organização Descritivo a “descrição de local” e o “relato de procedimento” (este principalmente por estar ligado ao que chamamos descrição de processos). Excluímos o “relato de opinião” por considerar este mais ligado ao Modo Argumentativo. Ambos os *subcorpora* foram salvos em arquivos txt, para poderem ser lidos pelo programa de análise textual WorsdSmithTools (Scott,1998) que será descrito mais adiante.

²⁹ Preferimos usar o termo Modo de Organização por estar enquadrado em uma de nossas bases teóricas, sendo que o termo original utilizado na construção do *corpus* refere-se a tipos de discurso.

Quanto às variáveis adotadas, consideramos apenas, para nossa análise, os níveis de escolaridade CA Infantil, CA Supletivo, 4ª e 8ª séries ensino fundamental, 3º ano ensino médio e último ano do ensino superior, excluindo-se, portanto, a faixa etária e o sexo, tendo, porém, a hipótese de que essas duas variáveis, de alguma maneira, podem influenciar no emprego das formações *x-inho*. Para cada nível de escolaridade, foram selecionados 8 informantes, sendo 4 masculinos e 4 femininos, totalizando 48 informantes.

Ressaltamos, ainda, que apesar desse *corpus* ser de natureza oral não levamos em conta os elementos prosódicos (alteamento e abaixamento de voz, trecho acelerado e ralentado, pausas, pico entonacional, intensidade, duração etc.), em virtude de nossa base teórica estar centrada na Linguística de Corpus que busca, principalmente, evidências lingüísticas no contexto e, ainda, por não ter acesso às gravações das entrevistas e, mesmo assim, por não poder medir operacionalmente com aparelhos esses limites prosódicos, apesar de considerar que, para uma análise da língua falada, a prosódia é de suma importância.

Não buscamos adivinhar quais as intenções do falante, posto que estas só ele mesmo conhece, mas buscamos na sua fala os indicadores lingüísticos do papel funcional destas formações. Consideramos, portanto, o significado contextualizado, ou seja, o interesse não se restringiu ao item lexical apenas, mas ao item somado ao contexto temático.

O segundo *corpus* adotado nesta pesquisa é o *corpus* de natureza escrito, CETENFolha do periódico Folha de São Paulo do período de janeiro a dezembro de 1994³⁰, com todos os seus cadernos. É um *corpus* de extratos de textos eletrônicos da Folha de São Paulo, compilado pelo Núcleo Interinstitucional de Lingüística Computacional da Universidade de São Carlos (NILC), criado pelo Projeto Processamento Computacional do Português. O *corpus* é obtido através do site www.linguateca.pt mantido pelo Ministério da Ciência e Tecnologia de Portugal. O CETENFolha está dividido em extratos, classificados por semestre e cadernos. Cada extrato divide-se em parágrafos e frases e os títulos e autores dos artigos mantidos³¹.

É importante ressaltar que qualquer comparação entre os *corpora* oral e escrito, tem que ser fundamentada em uma normalização, isto é, os dados têm que ser transformados em

³⁰ O CETENFolha contém o periódico Folha de São Paulo de 1994 a 1999. Devido ao escopo de nossa dissertação de mestrado, usamos somente o ano de 1994 da Folha.

³¹ Usamos a versão não dividida do *corpus*.

percentagens. Portanto, os *corpora* utilizados neste trabalho apresentam a seguinte configuração, conforme tabela 2 abaixo:

<i>Corpus</i>	Ocorrências (tokens)	Formas (types)
Oral	67.861	6.156
Escrito	10.360.605	152.732

Tabela 2: Quantidade de ocorrências de formas lexicais nos diferentes subcorpora

3.4. O programa WordSmith Tools

Nesta seção, descrevemos o programa computacional selecionado para uso na pesquisa.

Há várias ferramentas de análise para *corpora* eletrônicos (Berber Sardinha, 2004). Dentre elas, foi escolhido o programa computacional WordSmith Tools. Esse programa, criado por Mike Scott e comercializado pela Oxford University Press (Scott, 1998), é de simples utilização, além de rápido e poderoso. Além disso, de acordo com Berber Sardinha (*op.cit.*), o programa atende às exigências de análise lingüística por ser de fácil manuseio, por apresentar flexibilidade e pela possibilidade de contextualização das palavras selecionadas para análise. O programa coloca à disposição do analista uma série de recursos, os quais, se bem usados, são extremamente úteis e poderosos na análise de vários aspectos da linguagem. Entre esses aspectos, estão a composição/densidade lexical e a temática de textos selecionados, além de, através da comparação entre dois *corpora*, possibilitar o levantamento das palavras-chave de cada um dos *corpora* analisados. A única exigência do programa é que os textos compilados para fazerem parte do *corpus* de estudo estejam disponíveis em formato que possa ser lido por computador, ou seja, txt ou em ASCII.

Tendo como ponto de partida alguns princípios da lingüística de *corpus*, WordSmith Tools oferece várias ferramentas. Uma delas, o *Concord*, tem a função de concordanciador (nome pelo qual ficou conhecida essa função de extração) e extrai concordâncias a partir de um ou mais *corpora* selecionados. Em outras palavras, o programa disponibiliza ao usuário, a partir de uma palavra de busca específica, todas as ocorrências dessa palavra, juntamente com o seu contexto (palavras ao redor). O ponto de partida dessa ferramenta é o conceito de KWIC (ou palavra-chave em seu contexto). Em outras palavras, é no contexto que se constroem os significados de qualquer palavra. Assim, a partir do alinhamento que o

programa dá a palavra de busca, é possível ver se essa mesma palavra estabelece padrões de aparecimento em determinadas estruturas lexicais ou gramaticais repetidas. Abaixo, uma típica tela contendo palavras de busca com o sufixo *-inho*, todas alinhadas ao centro e mostradas em ordem alfabética.

The screenshot shows the WordSmith Tools software interface. The main window displays a concordance table with the following columns: N, Concordance, Set, Tag|Word No., File, and %. The table lists various words ending in '-inho' such as 'bonzinho', 'brinquedo', 'cachorrinho', and 'caminhãozinho', along with their respective frequencies and file references.

N	Concordance	Set	Tag Word No.	File	%
124	e até um pouco de romance entre um skinhead bonzinho (?) e uma integrante amalucada do g	p	656.276	etenf~1.0	6
125	onzinhos? Esperem um pouco, porque o bonzinho sou eu.» Desse modo, s	j	3.603.742	etenf~1.0	33
126	pantalho, por exemplo! Rarárá. Tô bonzinho hoje, né? É que o candidato FH	j	5.738.817	etenf~1.0	53
127	, Ruth afirmou que o PFL mudou «não porque é bonzinho , mas porque perdeu poder» .	j	3.471.213	etenf~1.0	32
128	anhar dinheiro nas apostas, mas tem um irmão bonzinho , que guarda um segredo .	p	2.157.080	etenf~1.0	20
129	io, Ruth disse que o PFL mudou «não porque é bonzinho , mas porque perdeu poder» .	j	680.436	etenf~1.0	6
130	com uma visão maniqueísta . Ninguém é bonzinho , ninguém é mauzinho . Tenho t	j	9.338.925	etenf~1.0	86
131	co» e disse que o partido mudou «não porque é bonzinho , mas porque perdeu poder» .	j	4.076.827	etenf~1.0	38
132), disse ontem que o PFL mudou «não porque é bonzinho , mas porque perdeu poder» . S	j	4.724.343	etenf~1.0	44
133	conhecido por gostar desse tipo de rock inglês brinquedo e ensimesmado, mas o Blur vai be	i	412.057	etenf~1.0	96
134	nte . O rapaz escolheu roteiro bem brasileirinho para mostrar aos três amigos um	i	5.945.191	etenf~1.0	55
135	isso funcionar teremos sempre um brasileirinho se enfiando nos pódios por aí .	j	8.619.856	etenf~1.0	80
136	k Rio, «revela a a escola/ de samba paradoxal/ brasileirinho / pelo sotaque/ mas de língua inter	j	2.123.660	etenf~1.0	20
137	quase em casa e sorri quando fala de seu novo brinquedo . Esse sorriso é a maior am	p	6.181.156	etenf~1.0	57
138	rescindível hoje em dia», regulamenta . O brinquedo eletrônico tem uma função interes	p	62.391	etenf~1.0	1
139	uo é ladrão, tem amantes, é homossexual, usa brinquedo . Não se vê o mesmo pr	j	909.937	etenf~1.0	8
140	de mostrar seus dotes . O saxofonista de brinquedo armou histrionices . O trombo	d	6.133.366	etenf~1.0	57
141	cara?» Os olhos escuros e brilhantes e o brinquedo de ouro são mostrados no primeiro r	d	8.216.710	etenf~1.0	76
142	dicando o fotógrafo do jornal: «Com esse brinquedo , no meu governo não entrava» e deu	j	8.996.966	etenf~1.0	83
143	sfrutáveis pimpolhas, que o espionavam por um buraquinho da cabine do barco? A situaç	d	9.016.755	etenf~1.0	83
144	ao óvulo, entendido? Nenhum outro buraquinho da pele ou do corpo pode substituir	d	511.298	etenf~1.0	5
145	a menstruação . Pois é por esse mesmo buraquinho que o espermatozóide pode entrar .	d	3.520.030	etenf~1.0	32
146	ta . O sotaque simples, os olhos azuis, o cabelinho branco, o perfil magro, a fé: tínhamos	p	6.580.248	etenf~1.0	61
147	A gente pega o rato, joga na fogueira e raspa o cabelinho . Depois abre a barriga, tira as	d	731.929	etenf~1.0	99
148	alhasdas ecoavam com Vieira imitando veado de cabelinho curto . Collor de boca aberta .	d	103.086	etenf~1.0	92
149	nhor Giovanni Arnolfini e sua mulher Giovanna o cachorrinho no primeiro plano, maçãs numa ca	d	376.015	etenf~1.0	3
150	em t-shirt de corte estreito; acima, desenho de cachorrinho DISNEYLÂNDIA	d	5.322.098	etenf~1.0	49
151	rasil x Inglaterra, na Copa de 1962 no Chile, um cachorrinho invadiu o gramado, driblou vários jo	d	650.824	etenf~1.0	98
152	r exemplo, traduzir como «Choque» o nome do cachorrinho da heroína da «Madeixa», quando	d	5.640.480	etenf~1.0	52
153	emitiu um som de gemido ou ganido, como um cachorrinho e aí sumiu.»	d	9.364.207	etenf~1.0	86
154	emito um som de gemido ou ganido, como um cachorrinho . Tradução de Clara All	d	8.181.882	etenf~1.0	76
155	foi definitivo . Ele estava tão enferrujado... caidinho . Mas eu já encontrei um substit	i	7.624.732	etenf~1.0	70
156	o tipo «ele não pode viver sem você», «ele está caidinho , você pode fazer o que quiser com ele,	m	6.214.405	etenf~1.0	57
157	militar está feia, e o pior é ter que aceitar tudo caladinho . Virgínia Félix, casada	i	7.561.426	etenf~1.0	70
158	s em quantidade, estatuetas egípcias, um lindo camelinho chinês falsificado .	d	6.875.469	etenf~1.0	63
159	franquias, a Shred-it . Cada franquía é um caminhãozinho equipado com máquinas de gra	p	6.871.412	etenf~1.0	63

Figura 1: Tela do programa *WordSmith Tools* mostrando linhas de concordâncias de palavras terminadas em *x-inho*.

Através do mesmo programa concordanciador, foi possível a ampliação das linhas de observação, isto é, em vez de se observar o item selecionado inserido em um contexto de apenas uma linha com cinco palavras à esquerda e cinco à direita, optou-se pela visualização de um contexto ampliado com um ou mais parágrafos. Essa opção foi feita para uma melhor visualização do emprego das formações *x-inho*, podendo-se caracterizar melhor suas diversas noções.

O concordanciador também permite a extração de ‘colocados’, isto é, os itens lexicais mais freqüentes que ocorrem com determinado item de busca. Isto pode indicar se o item de busca ocorre mais freqüentemente sozinho ou em grupos lexicais compostos de multipalavras, por exemplo.

Outra ferramenta computacional que integra o programa WordSmith Tools é o *Wordlist*, que gera uma lista de frequência a partir de um item ou mais textos selecionados. O programa disponibiliza ao pesquisador informações quanto ao total de palavras *token* (palavra inglesa que pode ser traduzida como *item* ou *ocorrência*, refere-se às palavras que podem ser encontradas em um texto) do *corpus* de pesquisa e também quanto ao número de palavras *type* (termo utilizado para se referir às palavras diferentes existentes em um texto) juntamente com dados estatísticos dados. O pesquisador pode obter a lista de palavras tanto em ordem alfabética quanto em ordem de frequências. As listas de palavras, bem como as frequências indicadas nas mesmas, podem oferecer ao pesquisador um raio-x dos textos que compõem seu *corpus*. Além disso, é possível proporcionar resultados reveladores quanto à descrição do contexto das palavras selecionadas, por meio do exame de co-ocorrência, o que foi primordial neste trabalho para a análise e interpretação das diversas noções das formações *x-inho*.

Para melhor compreensão da potencialidade do programa, listamos abaixo as dezesseis funções desempenhadas pelas três ferramentas do programa WordSmith Tools citadas por Berber Sardinha (2004), mantendo entre parênteses os nomes em inglês usados no programa:

a) Wordlist:

1. Lista de palavras individuais (Wordlist);
2. Lista de multi-palavras (Wordlist, cluster activated);
3. Lista de palavras de consistência individuais (detailed consistency);
4. Lista de dimensões e densidade lexical (statistics). Apresenta várias contagens relativas aos textos do *corpus*, tais como o tamanho em itens (tokens) e formas (types), números de parágrafos, comprimento do parágrafos, etc

b) Concord:

1. Concordanciador (Concordance);
2. Lista de colocados (Collocates);
3. Lista de agrupamentos lexicais (Clusters);
4. Lista de padrões de colocados (patterns);
- 5 Gráfico mostrando ONDE no texto ocorre palavra de busca (plot).

c) KeyWords:

1. Lista de palavras-chave (keywords);
2. Banco de dados de listas de palavras-chave (database);
3. Lista de palavras-chave chave (key keywords);
4. Lista de palavras-chave associada (associates);
5. Lista de agrupamentos textuais (clumps);
6. Gráfico de distribuição de palavra-chave (keyword plot);
7. Listagem de elos entre palavras-chave (keyword plot links).

Quadro 4: Funções desempenhadas por cada uma das três ferramentas do programa de análise WorsSmith Tools, (adaptado de Berber Sardinha (2004).

Para análise dos *corpora* de nosso trabalho foram utilizados os recursos a) 1 e b) 1 do quadro acima, ou seja, lista de palavras individuais e concordâncias.

3.5. Passo a passo da extração e análise.

O primeiro passo da compilação foi dar um comando ao programa para que extraísse todas as ocorrências com sufixo *inho* de cada um dos *corpora* separadamente. Nosso objetivo era coletar, dentro do *corpus*, todas as ocorrências da formação *x-inho* e analisá-las de acordo com sua função afetiva, diminutiva ou intensificadora. Uma vez analisada cada ocorrência nosso objetivo era quantificá-las para depois comparar esses números e distribuição.

Como já descrevemos no presente trabalho, além do estudo das frequências, o estudo de *corpus* eletrônico com ajuda de programa concordanciador permite o levantamento sistemático de padronização, ou seja, das regularidades formadas pela ocorrência de itens coocorrentes. Desta forma, o primeiro comando dado ao Concordanciador foi o de buscar **inho*, onde o asterisco funciona como toda e qualquer base. Após a extração de todas as linhas de concordâncias nas quais constavam as formações *x-inho*, em cada um dos *corpora* foram eliminados todos aqueles itens lexicalizados terminados em *-inho*, como *caminho*, *espinho*, etc. Uma vez feita essa limpeza, foram conservados somente os itens verdadeiramente compostos de uma base+sufixo.

No *corpus*, oral foram extraídos 157 itens terminados em *-inho* que, após a limpeza, foram reduzidos a 131. No *corpus* escrito, o número total e preliminar de itens em *-inho* foi de 10.544. As ocorrências verdadeiras de base mais sufixo foram 897.

Para a análise da padronização, foi necessário dar um segundo comando ao concordanciador para que listasse as palavras de busca (neste caso todas as palavras em **inho*, depois da limpeza mencionada acima). Por fim, cada uma dessas instâncias foi analisada de acordo com a função desempenhada no co-texto. Para cada linha, foi então atribuída uma identificação de classificação: d para diminutivo, p e n para função afetiva positiva e negativa, m para usos metaforizados ou lexicalizados.

A partir da compilação de todos os 'd', todos os 'p', etc. de nossas classificações foi possível dar um último comando ao concordanciador para que extraísse somente aquelas listas contendo uma única função. Por fim, com as listas já prontas quantificamos não só a distribuição de itens formados em *x-inho* por *corpus*, mas também fizemos um contraste entre os *corpora*.

Neste capítulo, fizemos uma breve revisão sobre o que seja Linguística de Corpus e seus conceitos mais importantes. Mostramos também o passo a passo de nossa análise. Passamos agora então ao nosso capítulo analítico.

4. ANÁLISE E INTERPRETAÇÃO

Este capítulo visa a descrever os passos da análise empreendida quando submetemos os dados aos processos de quantificação através do programa *WordSmith Tools*.

Nosso objetivo consiste em coletar, dentro dos *corpora*, as ocorrências da formação *x-inho* e analisá-las de acordo com os propósitos descritos no capítulo 1, quais sejam: a) investigar quantitativamente em um *corpus* de natureza oral a frequência das formações *x-inho*; b) identificar que tipo de função *x-inho* está exercendo dentro desse *corpus*; c) verificar qual a distribuição dessas funções dentro do *corpus* oral, se o subdividirmos em Modos de Organização Narrativo e Descritivo; d) investigar quantitativamente em um *corpus* de natureza escrita qual a frequência e distribuição por função das formações *x-inho* dentro dele; e) como podemos comparar probabilisticamente os dois *corpora* em termos da distribuição e frequência das formações *x-inho*.

4.1. Contornos estruturais para identificação das funções *x-inho*

Primeiramente, vamos apontar, nesta seção, os contornos estruturais das formações *x-inho* dentro de um contexto discursivo circundante, mostrando os parâmetros sintáticos e semânticos para o reconhecimento se uma forma em *x-inho* tem função dimensiva, intensificadora, afetiva ou outra.

Concordamos com Ezarani (*op.cit.*) ao presumir que *x-inho* – como qualquer outra expressão de grau – pode sinalizar o núcleo estrutural da atitude subjetiva do falante, no sentido de que evidencia, de imediato, no discurso, a presença de subjetividade. Essa, por sua vez, pode também dispor-se nas estruturas circundantes. Deste modo, identificamos as funções de *x-inho* nos *corpora*, não como item isolado, mas, sim, a partir da investigação de seus precedentes e conseqüentes dentro de seu co-texto e contexto.

Sabemos, porém, que a incidência das funções sobre determinada classe gramatical dá-lhes uma caracterização morfológica básica. Dessa forma, postulamos que+ a função diminutiva ocorrerá somente em substantivos, mas não obrigatoriamente em todos os tipos de substantivos. A função intensificadora reside morfológicamente em adjetivos, advérbios e pronomes, considerando o seu posicionamento no *continuum* quantitativo de intensificação. As funções positiva e pejorativa, a rigor, podem ocorrer em qualquer classe gramatical, mas em nossos dados, reside principalmente nos substantivos.

Ressaltamos, ainda que referenciamos os exemplos retirados dos *corpora* com a indicação E, ON ou OD correspondentes ao *corpus* escrito e aos *subcorpora* oral narrativo e oral descritivo mais a sucessão de um número que indica o número da concordância dentro do tipo de *corpus*. Desta forma, em (E.66) por exemplo, lê-se: exemplo retirado do *corpus* escrito, concordância 66. Em (ON.17), lê-se: exemplo retirado do *corpus* oral narrativo, concordância 17 e, em (OD.25), lê-se: exemplo retirado do *corpus* oral descritivo, concordância 25.

4.1.1. A Função diminutiva

A estratégia básica inicial de reconhecimento da função diminutiva consiste em fazer corresponder à forma sintética a forma analítica do diminutivo, assim *aparelhinho* é um aparelho pequeno, *barquinho* é um barco pequeno e *saquinho* é um saco pequeno em:

- 6 - O executivo terá acesso a agência de notícias, bancos de dados, reservas de passagens e até videogames num aparelhinho que poderá ser carregado no bolso. (E.66)
- 7 - ue eu sei fazer bem é gelatina... eu deixo/ eu boto um pouquinho de água pra ferver... e:: aí eu de/ quando a água ferve eu pego... o saquinho de gelatina... despejo no/ despejo num vasilhame... boto um pouquinho de água fervendo... boto um pouco de água gelada...(OD:118)
- 8 - ... ela ficou lá... os pais dela... e os avós dela... estavam em casa... só eles que tinham ido pra praia... depois... eles fizeram um barquinho lá... não... primeiro eles saíram num jet-ski... que a minha colega falou... foram pra uma ilha lá... não sei aonde... (ON.30)

Um outro indicador desta função é a presença do adjetivo *pequeno* ou outro correlato em referência à formação:

- 9 - De repente, você não está mais fazendo um cruzeiro num navio equipado. Está num barquinho pequeno, remando contra a maré. (E. 69)
- 10 - ... e que você tenha achado muito engraçada ou muito triste... I: essa história foi a minha mãe que contou... ela falou que... tinha um garotinho pequenininho... de dois meses... ele estava no carrinho... balançando uma fralda... e lá... na casa dele...(ON:21)

Também em função diminutiva estão os substantivos cujo conteúdo semântico da base contém por si só o traço [+ redução]. A presença do sufixo acentua a idéia de diminuição. Nesse caso, pode-se dizer que o indicador geral da função diminutiva seria a sua centralização em substantivos capazes de conservar o conteúdo semântico que se expressa na base e/ou no sufixo com ou sem irradiação para o contexto circundante.

- 11 -Tenho certeza de que muita gente, assim como eu, gostaria que vocês reservassem um pedacinho do jornal para dar sugestões de novos lançamentos de filmes. (E.579)
- 12 - to ((riso)) como é que é o nome? Miojo Lamen... então... é só você botar na agüinha pra ferver... medir aqueles copinho... tem a instrução... tá? no... no pacotinho... aí você bota os três copinhos de água... bota pra ferver... bota aquele pozinho... lá... e pronto... está feito o macarrão... não tem mistério...(OD:01)
- 13 - ta um::/ uma tira de uns::... trinta centímetros por::... cinco ou seis... de cada cor... eh::... um palitinho de churrasco cortado ao meio ou então um palitinho mesmo de:: pirulito... durex... ah... e papel crepom verde também pra fazer as/a folhinha... bom... você vai franzir essa tira colorida... (OD.10)

4.1.2. Função afetiva positiva

A proximidade quase sempre afetiva entre o falante e o objeto de seu afeto transparece, estruturalmente, nos substantivos em que a função positiva se centraliza e também no contexto circundante.

Consideramos como indicadores estruturais no contexto circundante para indicação da função positiva o seguinte:

a. referenciadores de 1ª pessoa:

- 14 - a porta... tem um espelho atrás da porta... tem umas pastas destas de material de desenho... assim... grandes... aí aquilo ali é... uma bagunça... meu cantinho da bagunça... assim... onde eu vou jogando tudo... e só::... assim... tem um tapete na sa/ no meio do quarto... todo colorido também...(OD.09)
15.não acho graça no barato da cocaína; prefiro o meu baseadinho...(E:78)

Quanto à ocorrência 15, observa-se que *baseado* traz uma conotação negativa, ou seja, referir-se ao uso de algum tipo de droga, porém o falante trouxe uma carga positiva ao seu enunciado, evidenciado, principalmente pelo pronome de 1ª pessoa “meu”; a construção com *-inho* apenas enfatizou essa carga positiva.

b. reiteradores de afetividade:

16. Daquele tempo lembro com alguma nitidez do JK, do barulhinho gostoso dos motores dos tempos dos DKW E dos Malzoni. (E.73)
- 17 -. ... então ele é gostoso... então pra você ler um livro... você:: ouvir um walkman... passar uma revista... ou então até você () assim... você está num cantinho tão gostoso... que todo mundo chega... vai e procura o cantinho... e meus dois irmãos quando casaram tentaram copiar esse cantinho ((riso). (OD:13)

Nesses exemplos, podemos observar que os termos *tão* e *gostoso* reiteram a afetividade expressa pelas palavras construídas com *-inho*.

c. repetição do objeto de afeto:

18 - e o seu *dinheiro* ela juntava pra minha avó... porque minha avó era muito pobre... e não tinha *dinheirinho* de aluguel... então ela dava... pra juntar... e aí cabia pro... (ON.26)

19 - Gomes disse que, para não deixar o *galo* sozinho, decidiu também comprar sua companheira. É um *galinho* caipaira, ainda novinho que precisa da companheira. (E.355)

Motivada pelo conteúdo semântico de bases que contêm o traço [+ afeto] para o falante, a função positiva tende também a residir em substantivos que indicam:

a. nomes de parentescos próximos:

20 - Com quase três anos, o *filhinho* de Lisa, Cristofher, não demonstra interesse algum por futebol.(E.311)

21 - A minha mãe não via a hora de encontrar de novo o *netinho*, afirmou Renata. (E.548)

b. nomes de seres vivos e/ ou partes deles:

22 - a gente pode brincar com ele... eh::... até também tem alguns bichinhos que a gente gosta muito... e alguns que a gente até tem medo... tem pônei... que a gente pode levar... eh a gente pode levar eh::... eh::... eh::... o seu... *gatinho*... po/ gatinha... (ON)

23 - O *galinho* ajudava acabar com as ervas daninhas do quintal. (E.356)

c. hipocorísticos:

Entende-se por hipocorísticos, “qualquer palavra criada por afetividade com intenção de carinho”(Camara Jr., 1973:217, apud Ezarani, 1989).

24 - Aos 10 anos, Christiane aprendeu a beijar <de língua>. Foi com Sigfried, um *gauchinho* de 10anos que dizia, cara de bobo: Eu te amo guria. (E.369)

25 - Romário, o *galinho* de Vila da Penha. Os Deuses são Nanicos. (E.354)

Observa-se, portanto, que em relação à função positiva, o contexto circundante (que neste trabalho denominamos cotexto) parece exercer o papel relevante de promover o surgimento de *x-inho*.

4.1.3. Função afetiva negativa

Segundo Frota (1985:13), os sufixos aumentativos e diminutivos apenas acentuam ou minimizam a carga semântica da base, tenha ela valor pejorativo ou não. Acresce, portanto, que os sufixos diminutivos minimizam cargas positivas da base, dando à formação um certo valor depreciativo, como em *bonzinho*, *simpatiquinho*, mas também, podem, por outro lado, minimizar a carga pejorativa da base, como em *burrinho*, *bobinho*. (Frota, *ibid*)

Por limitar a investigação ao âmbito exclusivo da palavra, passaram despercebidas por Frota (*ibid*) outros indicadores estruturais da função pejorativa de *x-inho*, que não seja somente a base da formação.

Neste trabalho, consideramos também como noções pejorativas aquelas que se estruturam em:

a. presença de dêiticos de 2^a e 3^a pessoa:

- 26 - Se o senhor deseja comprar alguma coisa a prestação, eu sugiro que o senhor guarde esta prestação, guarde *esse dinheirinho* por dois ou três meses, que logo mais os preços vão ter segurança de que não vai acontecer nenhum aumento. (E.256)
- 27 - Desconfio que eles nem sabem fazer isso, jogar defensivamente. Mesmo assim, com *aquele goleirinho*. (E.371)
- 28 - Um dia , o “espírito” fugiu de casa e quando o descobriram, ele estava lá dentro no escuro, e era *aquele pretinho* de quem se via só, os olhos brancos, e por isso o chamaram de “Espírito da Garagem da Bosta” (E.684)

b. presença de indefinidor:

- 29 - ...perguntaram a Jackie uma lista de suas experiências em viagens que poderiam ajudar o marido. Ela fez então *um diáriozinho* de 141 palavras em uma folha amarela. (E.254)
- 30 - ...entre a defeituosa MP da URV e as conversões para o Congresso e o Judiciário. O que se pode depreender do episódio é muito mais do que *um artiguinho* comporta. (E.39)
- 31 - só tem *um grupinho* que nega nosso sucesso de qualquer jeito, dizendo até que a gente nunca tocou lá fora. (E.390)

Nos exemplos 29, 30 e 31 acima, o artigo indefinido “um” mais a formação *x-inho* aliados ao contexto, nos permite identificar essas formações como pejorativas.

c. presença de estruturas reiteradoras da pejoratividade:

- 32 - E os Raimundos, você já sabe: são de casa. A única bola fora é o Sabbath com *esse cantorzinho caído* aí, mas tudo bem. (E. 165)
- 33 - Quayle é um reacionário de direita da escola antiga. Ele foi um republicano capaz de alegrar o *coraçõzinho sujo* de Nixon. (E. 230)
- 34 - ...como gostávamos de chamar as canções e os sambinhas que perpetrávamos naquela época de tamanha ingenuidade, que supúnhamos que *aquele mundinho corrupto* e egoísta se esfacelaria diante de uma sextilha e acorde dissonante. (E.537)
- 35 - A greve envolve um problema ético. Só os alunos de graduação, pequenos fascistas com ódio dos imigrantes, aqueles que se atrevem a vir gozar os benefícios do esporte e das aventuras. Eta, *povinho pequeno!* Não tem os horizontes libertários que sempre caracterizou o surfe. (E.679)

Nos exemplos 32, 33, 34 e 35 acima, as palavras *esse*, *caído*, *sujo*, *aquele* e *corrupto* reiteram e enfatizam a carga negativa que o enunciador atribui às formações em *-inho*.

4.1.5. Função intensificadora

Muitos autores têm se empenhado em provar ou negar a existência de uma função de superlativização em formações com *-inho*. Barrera (1986:31, apud Ezarani, *op.cit.*:61) considera que o “diminutivo” move-se nos eixos do que chama “diminuição objetiva, afetividade e superlativização”, coloca que o que existe “é uma gradação, uma relação de grau e não de oposição.”

Oliveira (1962:80) afirma que a impressão de valor superlativo que por vezes é atribuído ao “diminutivo” é à existência de outros intensificadores na frase (entoação, repetição, etc.). O autor não nega, portanto, que, em muitos casos, é fácil passar de um valor afetivo para um valor intensificador e afirma:

O sufixo diminutivo não é, de maneira nenhuma, um processo de superlativização, mas a atitude emotiva provocada no sujeito por uma qualidade intensa, pode traduzir-se numa apreciação de caráter subjectivo, que encontra expressão no *-inho*. (Oliveira,1962:83)

Citamos também Skorge (1963:263, *apud* Ezarani, *op.cit.*: 62) que chama a isto “superlativo da opinião subjectiva” ilustrando com a possibilidade alternante de se dizer a mesma coisa em “A obra está *perfeitinha*” e “A obra está *perfeitíssima*”.

Observa-se, portanto, que os autores corroboram a idéia de haver algo de intensidade nas formações *x-inho*, mesmo não admitindo que esse algo de intensidade se configure em uma função do sufixo. Dessa maneira, se algo de intensidade é expresso nas formações *x-inho*, ainda que subjetivamente, não pode ser desconsiderada, devendo, portanto, ser analisada a sua natureza e ser verificada a sua importância semântico-gramatical, ao lado dos demais processos de intensificação que já são subjetivos intrinsecamente.

Concordamos com Ezarani (*op.cit.*) ao considerar como intensificação a possibilidade de mover uma palavra de um ponto para outro num *continuum* da sua significação, para situar subjetivamente sua quantidade intensiva dentro de áreas consideradas sócio-culturalmente positivas ou negativas, através da utilização de sufixos superlativos e, mais coloquialmente, dos sufixos *-ão* e *-inho* em função de intensificação.

A estratégia inicial para identificar a função intensificadora consiste em poder fazer corresponder à formação *x-inho* a forma base acompanhada de qualquer palavra compatível com o seu significado que expresse a quantificação de intensidade. Isto ocorre, na maioria das vezes, nos adjetivos:

36. Nelson... um cara super extrovertido... que sacaneia muito... então o cara entrou ali... estava tipo assim (há) uns dois meses... o cara está **perdidinho** ((riso)) então o outro cara está naquela sacanagem toda... “aí:: não sei quê...” e o meu clima de serviço sempre foi assim...(ON:13)
37. festa lá... eh... também tem o quarto da/ eh... do patrão e da patroa dela... a cozinha é enorme... eh... aí... eh... tem... sabe? aqueles banheiros **bonitinho** lá... todo cheio de... banheira assim... ih... adorei... aí tem... eh/ e também tem o quarto de empregada... lá... eh... lá é legal de ficar.. (OD:41)
38. Ninguém gosta dessa verdura quando é pequeno. Depois não há quem resista ao seu sabor **amarguinho** e frescor característico. (E:23)

Em 36, a formação *x -inho* pode ser interpretada como *bem perdido, bastante perdido*. Em 37 e 38 *x -inho* também pode ser perfeitamente interpretado como *bem bonito* e *bem amargo*. Pode-se notar, ainda, que nesses exemplos há também a noção afetiva positiva, principalmente pelo contexto em que essas formações foram produzidas. Em 37, percebe-se essa afetividade positiva pelas expressões *ih... adorei, lá é legal de ficar..* . Em 38, a afetividade positiva é percebida pela expressão *não há quem resista, frescor característico*.

Pode-se encontrar a função intensificadora também nos advérbios:

39. Em Palo Alto já de manhã **cedinho**, um sinal perturbador. Desde que chegamos aqui na Califórnia, o céu tem-se estendido azul, diria que enjoativamente azul, sem um fiapo de nuvem. (E:185)
40. no microfone ((riso)) todo mundo olhando assim... ninguém entendia nada... gente pertinho de mim assim... todo mundo **pertinho** da caixa de som... ninguém entendia nada do que eu falava... aí... fui falando... falando... até a hora que eu n (ON:09)
41. a água pra ferver... bota o ovo dentro... aí... quando tiver assim borbulhando... aí você tira o ovo... tira a casca e come... bota sal e come... E: **rapidinho**... né? legal... Narrativa de experiência pessoal E: Alexsandro... queria que você me contasse uma história que tivesse acontecido... (OD: 20)
42. Que só são encontrados na Christopher Street, no Village ou **lazinho** mesmo... (E:451)

Em 39, 40 e 41 as formações *x-inho* também podem ser interpretadas como intensificadoras, pois equivalem a *bem cedo*, *bem perto* e *bem rápido* respectivamente. Em 42 o enunciador enfatiza o lugar, intensificando o advérbio *lá* através da formação *lazinho*, que interpretamos como *somente nesse lugar*, *realmente nesse lugar* ou *bem nesse lugar mesmo*.

Por vezes podemos encontrar a intensificação também nos pronomes e substantivos:

43. o... aí depois eu vou ficar () quando ele estiver quase secando... aí vou ver se está du::ro... se está mole... se não tiver... aí eu boto mais um **pouquinho** d'Água... se tiver com pouco sal... eu boto... E: ah... obrigada... já aprendi a fazer arroz.. tá vendo? ((riso)) legal...(OD:15)
44. e estava/ todo mundo xingando a professora “professora... a senhora quer reprovar a gente...” e tal... aí quando chegou no **finalzinho**... cadê? ela tinha se esquecido de mim... da... da minha/ aí não tinha mais prova pra me dar... ela “ih:: falta você (ON:37)

Em 43 e 44 as formações podem ser interpretadas como *bem pouca* e *bem no final*, respectivamente

Além de configurar-se em classes gramaticais determinadas, a função intensificadora apresenta outros indicadores estruturais de sua presença, tais como:

a. advérbios de intensidade:

45. os ovos... o óleo... o leite... farinha... fermento... fermento... tá? aí você coloca tudo no liquidificador... aí bate... até ficar aquele creme... bem... **bem fininho**... tá? aí depois disso

... você unta uma forma... aí depois de untar... você coloca uma parte daquela massa... que você preparou no liquidificador... e... (OD:58)

46. Conheça as novidades da Comdex/Spring aldeia, na área Kreanak, na divisa do Estado de Minas Gerais com o Espírito Santo, tem muita menina e menino. Toda manhã a turma sai **bem cedinho** e entra na canoa para atravessar o rio. (E:184)

47. O show está **bem ensaiadinho**. Das 28 músicas gravadas na sexta-feira, 13 devem estar na versão em LP. (E:299)

No exemplo 47, além estar caracterizado a função intensificadora, principalmente, pela presença do advérbio de intensidade “bem”, podemos ressaltar que, neste exemplo, a opção pela formação em *-inho* não descarta a possibilidade de o enunciador querer expressar também uma atitude positiva, o que nos permite dizer que a função intensificadora, em alguns casos, pode coocorrer com a função afetiva positiva.

b. ratificador do conteúdo semântico da formação representado na estrutura por um vocábulo:

48. ... ih... adorei... aí tem... eh/ e também tem o quarto de empregada... lá... eh... lá é legal de ficar... dá pra/ dá até pra dormir que é tudo **escurinho assim**... é bem/ é legal... já brinquei de pata-cega lá com a minha prima aí... E: brincou de quê? I: pata-cega... aquele negócio que tampa o olho aí fica lá na escuridão... aí/ é? e um (OD:84)

49. buna... E: na rua Itabuna? a rua Itabuna fica onde? I: fica... E: fica aqui em São Cristóvão? I: fica aqui em São Cristóvão **mesmo... pertinho... quase** perto da escola... um pouquinho... fica a rua Itabuna... E: legal... Ana Maria... tem alguma uma história... alg (ON: 3)

50. Que só são encontrados na Christopher Street, no Village ou **lazinho mesmo**... (E:451)

c. pronome indefinido indicando completude:

51. A situação do militar está feia, e o pior é Ter que aceitar **tudo caladinho**. (E:157)

52. stante...ö ((risos)) aí... resolvemos fazer a estante... aí ela é toda ajeitadinha assim... porque tem lugar pra... pra lapiseira... tem lu/ sabe? **todo/ tudo certinho** assim mesmo... meio... meio designer... aí... tem a prancheta que eu comprei há pouco tempo também... que ela até está sem forrar...que el (OD:102)

53. No entanto, para reciclar a roupa deve-se misturar sem medo. Você pode pôr um blazer preto **todo quadradinho** com uma blusa mais moderna. Não muda aforma, mas muda a maneira de usar. (E: 695)

54. mais... quando chegou na hora do cerimonial... aí o rapaz do som armou o som... tudo... escolhemos a música... **tudo direitinho**... na hora do cerimonial mesmo de falar... quem ia

falar? eu... aí... comecei a falar... ninguém entendia nada ..(ON:38)

d. proximidade de bases e, conseqüentemente, de quantidade significativa no continuum da intensidade:

55. Parecia a Convenção dos Democratas. Festivo, tudo muito organizado! Cívico mesmo. Aliás, civicuzinho!.. (E:211)

e. repetição de *x-inho*:

56. Eu estou igualzinho ao presidente Itamar, igualzinho ao ministério, igualzinho ao 150 milhões de brasileiros, que não sabem coisa alguma sobre este plano. (E: 402)

Ressaltamos, ainda, que outro indicador de intensificação é a prosódia que pode promover fatos como o de destacar a altura, intensidade e/ou duração do /i/ do *-inho* e o de acentuar afetivamente sílabas da formação e/ou da unidade entonacional em que ela se insere, por exemplo. Porém, os fatores prosódicos não fizeram parte de nossa análise, pelos motivos já apresentados na introdução desta dissertação.

4.2 . Interpretação dos dados

Como já descrevemos no presente trabalho na seção 3.1, além do levantamento de frequências lexicais, uma investigação de *corpus* eletrônico com auxílio de programa concordanciador, como o *Wordsmith Tools*, permite o estudo sistemático de padronização, ou seja, das regularidades formadas por itens co-ocorrentes.

Para a análise dessa padronização, foi necessária a observação dos contextos de ocorrências do item em questão (*x-inho*), isto é se o *corpus* é oral narrativo ou descritivo ou se é escrito, bem como a observação cuidadosa do contexto, ou seja, os itens lexicais circundantes da palavra terminada em *-inho* que se quer analisar. Utilizamos para isso, a lista de concordâncias que, como já dissemos no capítulo 3, é uma lista das ocorrências de um item específico, dispostos de tal modo que a palavra de busca (aquela em que se tem interesse de investigar) aparece centralizada na página. A palavra de busca é acompanhada do seu co-texto original, isto é, das palavras que ocorrem junto com ela no *corpus*.

O primeiro passo para a análise das concordâncias é a produção delas. Para isso, utilizamos os comando do “Concord” do programa *WordSmith Tools*, descrito no subitem 3.4. Selecionamos a palavra de busca, que no nosso caso era **inho* (asterisco inho), isto é, toda e qualquer forma que terminasse em *inho*, circundada por um horizonte de 5L para 5R, ou seja, de cinco palavras à esquerda e cinco à direita da palavra de busca. Para analisarmos com maior exatidão o item formado com *x-inho*, foi necessário por vezes expandir o contexto para um horizonte maior. O resultado final das linhas de concordâncias, sem a expansão do horizonte, mostram-se no anexo 1 (concordâncias *x-inho*).

Uma primeira extração do total de ocorrências das formações *x-inho* nos *corpora*, baseando-se nos critérios apresentados acima, revelou os seguintes números dessas formações conforme tabela 3 abaixo:

TIPO DE CORPUS	NÚMERO TOTAL DE ITENS		FORMAÇÕES X- <i>inho</i>
CORPUS ORAL	MODO NARRATIVO	25.469	39
	MODO DESCRITIVO	18.668	118
CORPUS ESCRITO	10.360.605		10.544

Tabela 3: Nr total de itens e frequência das formações *x-inho* nos *corpora*

Ressaltamos, porém, que para a análise das funções exercidas pelas formações *x-inho*, tanto no *corpus* oral como no escrito, foram excluídos os itens lexicalizados, expressões metafóricas, nomes próprios e itens que estão em processo de lexicalização, conforme exposto no capítulo 2. Desta forma, o quadro acima não foi considerado para análise das funções expressas pelas formações *x-inho*. A tabela 4 abaixo mostra o quadro de ocorrências no *corpus* oral após a exclusão dos itens citados. A análise das frequências do *corpus* escrito será explicitada mais adiante.

Tipos de ocorrência	Narrativo		Descritivo	
	Nr. de Ocorrências	%	Nr. de Ocorrências	%
Formações <i>x-inho</i>	30	0,11	101	0,54
Itens lexicalizados	09	0.03	17	0.09
TOTAL	39	-	118	-

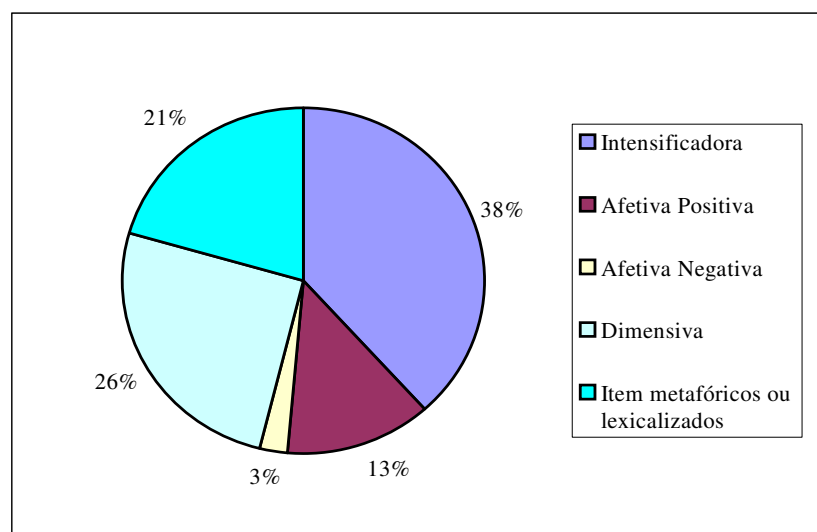
Tabela 4. Ocorrência de *x-inho* no *corpus* oral por modo

Pelos dados apresentados na tabela 4, verifica-se que, pelas percentagens, essas formações ocorrem com maior frequência no Modo Descritivo (cerca de quatro vezes mais do que no Modo Narrativo), o que não corresponde à nossa intuição, pois, a princípio, julgamos que as formações *x-inho*, carregadas de avaliação, em geral tendem a ocorrer com maior frequência no discurso narrativo, principalmente em narrativas de experiências pessoais de caráter oral, como é o caso dos dados selecionados para esta dissertação.

Isto se dá, pois, quando falamos de nós mesmos ou sobre eventos nos quais estamos envolvidos, tendemos a formatar nossas narrativas usando a categoria que Labov (1972) denominou Avaliação, isto é impregnando o fato narrado de subjetividade, com o intuito de realçar algum aspecto ou conteúdo da narrativa para torná-la mais interessante para o narratário.

Da análise das 131 ocorrências da formação *x inho* encontradas no *corpus* oral, após excluirmos as formações lexicalizados, expressões metafóricas, nomes próprios e itens que estão em processo de lexicalização, conforme já citado acima, pôde-se fazer a seguinte depuração, por ocorrências e por funções, em cada Modo de Organização do Discurso, conforme os gráficos abaixo:

Gráfico 1: Noções expressas por *x- inho* no *Corpus* Oral, Modo Narrativo



Baseados na análise e classificação individual de cada ocorrência, passamos a discutir o perfil das formações *x-inho* de cada um dos modos do *corpus* oral.

A noção intensificadora foi a mais freqüente em todo o *corpus* oral. Esse fato corrobora a idéia de que ao narrar, principalmente experiências pessoais ou recontadas (característica de nosso *corpus* narrativo), tendemos a usar a avaliação e enfatizar o fato narrado para atrair a atenção do narratário. Este fato contribui para o aparecimento das formações *x-inho* cuja função é intensificadora e que, conforme exemplos abaixo, deixa claro a intenção de realçar algum aspecto da narrativa.

57. Nelson... um cara super extrovertido... que sacaneia muito... então o cara entrou ali... estava tipo assim (há) uns dois meses... o cara está **perdidinho** ((riso)) então o outro cara está naquela sacanagem toda... “aí: não sei quê...” e o meu clima de serviço sempre foi assim... (ON: 13)
58. viu... aí o que que aconteceu? a gente... demorou ali um tempo... pra trocar o pneu... quando a gente trocou... foi tudo bem... demorou um **pouquinho**... né? aí a gente entrou no carro... estava tudo molhado... os papéis ((riso)) tudo molhado... o carro... a gente ent/encharcada.. (ON:23)

Podemos observar também, um número significativo de ocorrências da função dimensiva. Esse fato, porém, não é relevante em nosso *corpus* narrativo, pois as ocorrências se deram, principalmente em itens repetidos e em bases que já dão a idéia de pequenez, conforme *garoto* e *pequeno* nos exemplos abaixo:

59. ... desmaiou... aí chamaram o médico... ela foi pro trabalho/ a avó dele... foi pro trabalho da mãe dele... chegou lá... e levou o **garotinho** e a moça pro médico... nesse caminho pro médico... aí eles encontraram... um:/: eles pegaram um engarrafamento... (ON:06)
60. ou a mulher... aí ela começou a correr... correr... e falou/ ele estava com/ ele queria pegar o meu primo que estava com ela... que é **pequeninho**... mas aí ela correu... correu... entrou no prédio... e pegou o amigo dela e falou pra... ficar com ela... que tinha alguém (ON:18)

A tabela 5 mostra as ocorrências das noções dimensivas no *corpus* oral narrativo.

Itens	Número de ocorrências
Garotinho	5
Pequeninho	3
Barquinho	1

Tabela 5: Ocorrência da noção dimensiva

A tabela acima nos permite intuir que a noção dimensiva pode estar relacionada, principalmente, à descrição, pois, a nosso ver, isso se explica pelo fato de esta noção estar relacionada a características de substantivos.

Quanto às noções Afetiva Positiva e Negativa, há uma frequência pequena em relação às noções intensificadora e dimensiva, o que, parece-nos ter uma explicação pelo fato de que a proximidade, quase sempre afetiva entre o enunciador e o objeto do seu afeto transparece, estruturalmente, e com maior frequência nos substantivos em que essa função parece se centralizar.

Nesse sentido, seria mais evidente a ocorrência dessa noção com maior frequência no Modo Descritivo, pois, como já dissemos, nesse modo apontamos, principalmente, características de seres e objetos concretos que, por serem da categoria gramatical dos substantivos, são passíveis também de receberem uma carga expressiva de afetividade.

Conforme tabela 5, foram observadas 3 ocorrências da noção Afetiva Positiva e 1 ocorrência da Afetiva Negativa no Modo narrativo, o que pode nos revelar que, a afetividade expressa pelas formações *x-inho*, indicando tonalidade apreciativa, carinhosa, de cortesia etc. é mais frequente do que a afetividade que reflete depreciação, desdém, ironia, etc. A escolha por *x-inho* afetivo foi realizada no nosso *corpus* narrativo, na maioria das vezes, para expressar apreciação positiva em relação a algo, principalmente quando o enunciador quer refletir sua sensibilidade. A nosso ver, isso não ocorre com a noção negativa, pois, dentro de um contexto, nos parece que, para expressar afetividade negativa o enunciador pode usar uma palavra com base *x-inho* ou com base normal. Os exemplos abaixo demonstram algumas ocorrências dessas noções no *corpus* oral narrativo.

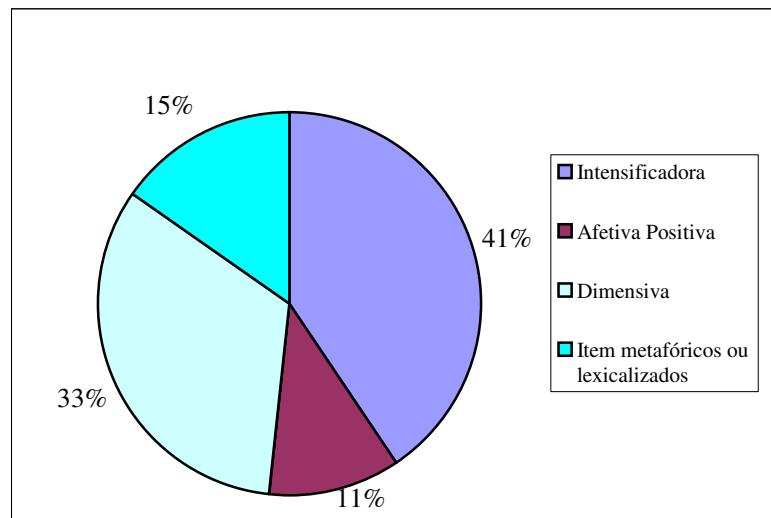
61. mês atrás... então o pessoal... numa sexta-feira... todo mundo “ah:: uma cervejinha... (vamos beber) uma **cervejinha**... vamos fazer um **churrasquinho**” e tal... tal... tal... todo mundo muito alegre... muito brincalhão... e eu fui pra:/: saí da academia... comi o churrasco... (ON:7)
62. empurrando um ao outro... ele sem querer encostou... na ((riso)) na mulher... e a mulher virou-se pra ele e falou assim “ô **rapazinho**... pára de tirar sarro comigo...” não sei quê “você não tem o que fazer não... é?” e aí ele me disse “pô... Fábio... (ON:16)

Em 61, podemos observar a carga afetiva positiva expressa nas bases substantivas *cerveja* e *churrasco*, reiteradas principalmente pelo co-texto em que aparecem as palavras *alegre* e *brincalhão*.

Em 62, podemos observar que, pelo contexto da ocorrência da formação em *-inho*, é evidente a expressividade de uma carga afetiva negativa à base substantiva *rapaz*. No entanto, se o enunciador optasse pelo uso da base normal *rapaz*, sem o acréscimo de *-inho*, a nosso ver, não mudaria a carga negativa, principalmente pelo co-texto da ocorrência (um rapazinho que *tira sarro*), o que corrobora a nossa concepção expressa anteriormente de que, em alguns contextos, para expressar afetividade negativa o enunciador pode usar uma palavra com base *x-inho* ou com base normal.

A partir da análise de cada uma das ocorrências do *Corpus Oral Descritivo*, de acordo com o gráfico 2 abaixo, podemos notar as seguintes características:

Gráfico 2: Noções expressas por *x-inho* no *Corpus Oral*, Modo Descritivo



Pela análise do gráfico 2, verifica-se que as noções intensificadora e dimensiva também são as mais produtivas. Em relação à noção intensificadora, há uma frequência bem maior do que a verificada no Modo Narrativo. Este fato, portanto, nos parece previsível pelo escopo do nosso *corpus* descritivo, em que as descrições referem-se a locais e processos. Nesse sentido, a alta frequência das noções intensificadora e dimensiva são pertinentes, pois, como já dissemos, ao descrevermos (locais, ambientes, pessoas, processos, etc.), podemos avaliar o objeto da descrição sob a perspectiva de suas características (físicas, psicológicas), bem como de sua localização espacial.

Nessa ótica, algo pode ser pequeno, grande ou pode estar em um *continuum*, como já dissemos, no qual posicionamos as palavras dentro de um maior ou menor grau de quantidade

intensiva ou dimensiva do significado positivo ou negativo de sua base e, ao nos referirmos a essas atitudes de valores, podemos empregar as formações *x-inho* como noção intensificadora ou dimensiva. Nesse sentido, o Modo Descritivo nos parece ser mais propício a essas construções.

A alta frequência da função dimensiva no Modo de Organização Descritivo corrobora a idéia de que essa função, como foi explicitado no capítulo 2, está relacionada ao grau de dimensão que ocorre, principalmente, em bases substantivas.

Partindo desse pressuposto, esse fato parece ser pertinente, pois quando descrevemos características físicas do objeto da descrição, essas características se relacionam, principalmente, a substantivos, e é nesse Modo de Organização do Discurso que nos preocupamos em apontar atributos da pessoa ou da coisa descrita, ou melhor dizendo, do objeto da descrição (seres humanos, animais, processos, seres inanimados, cenas e locais).

Sob essa perspectiva, avaliamos sob um *continuum* de tamanho, utilizando, portanto, para a noção de maior/menor, pequeno/mais pequeno, redução/mais redução as formações *x-inho* com a noção dimensiva.

Verificamos pelo gráfico 2 que a noção afetiva também mostrou-se menos freqüente do que as noções intensificadora e dimensiva, porém com uma frequência inferior a do Modo Narrativo, o que não foi compatível com a nossa intuição, pois, como já dissemos anteriormente, acreditamos que o Modo de Organização Descritivo parece-nos mais propício à afetividade, pelo fato dessa noção estar ligada, principalmente, a substantivos, e é, portanto, na descrição que se verifica uma maior probabilidade de se encontrar essa categoria gramatical.

Em relação à noção afetiva negativa, o *corpus* oral descritivo não apresentou nenhuma ocorrência, porém, a nosso ver, isso não descarta a possibilidade de que essa noção também seja produtiva nos Modos de Organização do Discurso apresentados.

Quanto ao *corpus* escrito, o programa extraiu um total de 10.544 ocorrências da formação *x-inho*. Ao investigarmos quantitativamente essas ocorrências os dados revelam a seguinte frequência dessas formações conforme tabela 6 abaixo:

PALAVRAS	TOTAL DE FORMAÇÕES <i>X-inho</i>	%
10.360.605	10.544	0,10

Tabela 6: Total das formações *x-inho* no *corpus* escrito

Ressaltamos que, para a análise das funções exercidas pelas formações *x-inho* no *corpus* escrito, também foi feita a exclusão de itens considerados não analisáveis, do mesmo modo que foi feito com o *corpus* oral. Verificamos, portanto, que o sufixo *inho* no *corpus* escrito da Folha de São Paulo é usado 9647 vezes em nomes próprios como Robinho, Ronaldinho, etc – sendo este talvez no discurso jornalístico o uso mais freqüente desse sufixo.

A tabela 7 abaixo mostra o quadro de ocorrências no *corpus* escrito após a exclusão dos itens citados:

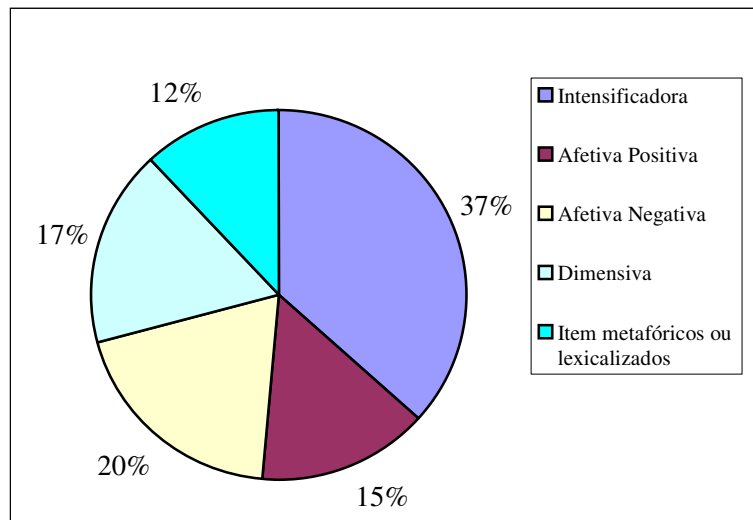
Itens	Ocorrências	%
Formações <i>x-inho</i>	897	0,008
Itens excluídos	9647	0,093
Total	10.544	0,10

Tabela 7 : Ocorrência de *x-inho* no *corpus* escrito

Considerando-se apenas para nossa análise o que chamamos formações *x-inho*, propriamente ditas, isto é, se retirarmos todos os 9647 substantivos próprios terminados em *inho* conforme tabela acima, verificamos que a freqüência real dessas formações é muito baixa em relação ao número de palavras do *corpus* (0,008% de ocorrências). Esse fato parece demonstrar que, em geral, na linguagem escrita as formações *x-inho* tendem a ser menos produtivas em comparação com a linguagem oral (0,11% no *corpus* narrativo e 0,54% no *corpus* descritivo, conforme tabela 4) em virtude das características próprias dessa modalidade, conforme explicitado no item 2.4 deste trabalho.

Da análise das ocorrências selecionadas, de acordo com o gráfico 3, podemos observar as funções expressas pelas formações *x-inho* na modalidade escrita:

Gráfico 3: Noções expressas por *x- inho* no *Corpus Escrito*



Pela análise do gráfico acima, notamos que, apesar de a ocorrência de formações *x-inho* no *corpus* escrito ser bem inferior ao *corpus* oral (narrativo e descritivo), há um número representativo de ocorrências das funções intensificadora e afetiva, concorrendo ambas quase com a mesma frequência, fato este, não observado nos *corpora* oral. Isso nos pode ser explicado pelo fato de o *corpus* escrito não ter sido dividido nos diversos gêneros jornalísticos que a Folha de São Paulo engloba.

Dessa forma, o *corpus* escrito inclui desde editoriais às críticas de filmes e às colunas sociais e humorísticas. Entretanto, mesmo sem a divisão em gêneros, os achados nos levam a supor que não há restrições quanto ao tipo de linguagem (oral ou escrita) para que essas funções ocorram.

Percebe-se, que mesmo na modalidade escrita formal, o enunciador pode demonstrar sua subjetividade com a possibilidade de mover uma palavra de um ponto para outro no *continuum* da sua significação para situar algo dentro de áreas consideradas sócio-culturalmente positivas e negativas, bem como demonstrar uma atitude afetiva de valor - funções que podem ser expressas pelas formações *x-inho*.

Um fato interessante que se pode observar ainda é em relação à frequência da função Afetiva Negativa, que mostrou ser mais produtiva do que a Afetiva Positiva, o que parece não corresponder à nossa expectativa. Entendemos, porém, que essa frequência maior da função Afetiva Negativa pode ser explicada pelo fato de que estamos lidando com discurso de jornal

– a função de um jornal, via de regra é criticar, o que já explicaria a função afetiva negativa. O interessante, portanto, é que o jornal esteja lançando mão de um sufixo *-inho* para expressar essa crítica.

Ressaltamos, portanto, que para esclarecer melhor esse fato, seria viável analisar o gênero em que se dão as ocorrências da função afetiva, principalmente a negativa, nesse tipo de *corpus*, pois entendemos que, dentro do tipo de texto jornalístico (que é o caso de nosso *corpus* escrito), podemos encontrar vários gêneros textuais prototípicos (notícia, reportagem, editorial, política, artigo etc.), que seriam passíveis de favorecer a ocorrência desse tipo de função veiculada por *x-inho*. Essa análise, entretanto, não faz parte de nosso trabalho, pois analisamos, apenas, as instâncias de *x-inho* no *corpus* escrito como um todo, sem subdivisões.

Em relação à função dimensiva, houve uma frequência inferior as demais funções, porém mais significativa do que as do *corpus* narrativo, pois, pela análise das ocorrências, observou-se que as mesmas se deram, principalmente, em bases substantivas diversas e não repetidas, diferentemente do que ocorreu no *corpus* oral narrativo. Os dados também nos revelam que a frequência dessa função no *corpus* escrito, se comparado ao *corpus* oral descritivo, mostra-se bem inferior, demonstrando, que a função diminutiva parecer ser prototípica da descrição.

Da análise das ocorrências selecionadas, de acordo com os gráficos 1, 2 e 3 acima, podemos fazer a seguinte comparação probabilística entre os *corpora* em termos de distribuição e frequência das formações *x-inho*.

Comparando-se o número de frequências das formações *x-inho* entre os *corpora*, os dados revelam que essas formações são mais produtivas na modalidade oral da língua, o que corrobora a idéia de que na modalidade escrita formal – que é o caso do nosso *corpus* escrito – as formações *x-inho* tendem a ser menos produtivas, principalmente pelas diferenças entre essas duas modalidades, conforme apresentado no capítulo 3.

Em relação às funções expressas por essas formações, a intensificadora foi a mais produtiva em todos os *corpora*, sendo que no *corpus* oral narrativo e no *corpus* escrito o número de ocorrências foi quase idêntico (38% e 37% respectivamente). No *corpus* oral descritivo, a frequência mostrou-se um pouco superior aos demais *corpora* (41%), sugerindo com isso que o Modo Descritivo parece ser mais característico para esse tipo de função.

A função dimensiva mostrou-se bem mais produtiva no *corpus* oral descritivo, corroborando as características da descrição explicitadas neste trabalho, pois quem descreve pessoas, animais, objetos, processos etc., pode valer-se das características dimensionais daquilo que é descrito. Ainda assim, os dados parecem revelar que essa função tende a ser mais prototípica da descrição.

Um fato importante que pôde ser observado foi a alta frequência da função Afetiva, principalmente a negativa, no *corpus* escrito, superando as dos *corpora* orais. Ainda, assim, o fato de a função Afetiva Negativa ter uma frequência maior do que a Positiva no *corpus* escrito contraria nossa intuição, pois acreditamos que essa função tende a ser mais produtiva na modalidade oral pelos motivos já apresentados neste trabalho.

Esse fato, portanto, nos remete a uma análise e estudo mais apurados dessas frequências, a fim de se verificar quais os fatores que justificam ou condicionam as mesmas dentro desse tipo de *corpus*, o que pelo escopo desta dissertação não foi possível realizar, ficando, portanto, para uma pesquisa posterior.

Em relação às formações *x-inho* que caracterizamos como itens metafóricos, lexicalizados, em processo de lexicalização e polilexicais, conforme explicitados no capítulo 2, houve uma frequência considerável nos *corpora*, principalmente no *corpus* oral narrativo, porém não podemos considerar que esse tipo de *corpus* seja o mais prototípico a essas construções, pois teríamos de analisar cada forma isolada, dentro de cada *corpus*, para poder chegar a uma conclusão mais pertinente sobre esse fato que os dados nos revelaram.

Baseando-se nas análises e interpretações feitas a partir dos dados e gráficos apresentados acima, faremos no capítulo 5 as considerações finais que julgamos interessantes para este estudo.

CAPITULO 5 - CONSIDERAÇÕES FINAIS

Vimos na revisão bibliográfica desta dissertação que em alguns compêndios prescritivos e descritivos, bem como em alguns estudos sobre a língua portuguesa o sufixo *-inho* é classificado como flexão de grau diminutivo. Sem a pretensão de estabelecer conclusões definitivas e tendo a consciência de que há muito para se estudar em relação a formação *x-inho* do português do Brasil a partir de *corpus* eletrônico, passamos agora a resumir os resultados apresentados neste trabalho. Retomaremos os objetivos do trabalho, que visou a estudar a distribuição e frequência dessas formações de dois *corpora* distintos: um de natureza oral e outro escrito, conforme os pressupostos abaixo.

O estudo apresentado descreveu padrões relativos à frequência das funções expressas pelas formações *x-inho* em um *corpus* de natureza oral, o qual foi subdividido em dois *subcorpora* (narrativo e descritivo), segundo Charaudeau (1992), e em um *corpus* escrito. Essas funções, porém, só puderam ser identificadas pelas características morfo-semânticas de suas bases e/ou pela interpretação dessas formações, levando-se em consideração seus contextos discursivos e seus cotextos mais amplos. A pesquisa teve como ponto de partida aquilo que as gramáticas explicitam sobre o sufixo *inho* e a investigação de evidências baseadas no uso da linguagem cacalda na Linguística de *Corpus*. O trabalho teve, portanto, um caráter inovador, pois até o momento não encontramos nenhuma outra pesquisa que analisasse esse fenômeno da língua sob essa ótica.

Um levantamento de todas as ocorrências dentro dos *corpora* e subsequente análise das mesmas sugere que o emprego das formações *x-inho* tendem a refletir, com mais frequência, outras atitudes por parte do enunciador do que puramente focalizar a dimensão. As funções expressas pelo sufixo (afetiva, intensificadora e dimensiva) competem entre si sob a mesma face de *x-inho*, porém quantitativa e probabilisticamente se distribuem diferentemente.

Nesse sentido, os dados nos mostraram que a função principal veiculada por *x-inho* é a intensificadora, mostrando-se muito mais relevante do que a função dimensiva (diminutiva). Tal fato parece contrariar alguns compêndios prescritivos e descritivos que, apesar de veicularem a possibilidade de essas formações expressarem atitudes subjetivas, não as quantificam nem estabelecem suas distribuições na linguagem oral ou escrita, deixando, além disso, a idéia de que a função dimensiva venha a ser a mais produtiva.

A análise de nossos dados também nos revelou que a função dimensiva é mais produtiva no *corpus* oral, principalmente no Modo Descritivo, o que, como já foi dito neste trabalho, corrobora nossa hipótese de ser a descrição mais passível de se encontrar esse tipo de função.

Percebemos também um fato interessante que se refere à função afetiva, a qual teve uma frequência maior no *corpus* escrito, principalmente a afetiva negativa. Esse fato, como já foi explicitado na análise dos dados, merece ser aprofundado visto que os resultados podem ter sido influenciados pela não divisão do *corpus* escrito em gêneros distintos.

Os dados mostraram que não há restrição de modalidade da língua para a ocorrência das formações *x-inho*, porém ficou evidente neste trabalho que a modalidade oral investigada tende a ser bem mais propícia a essas formações.

Naturalmente, este trabalho não pretendeu apresentar exhaustivamente as formações *x-inho*. Algumas questões de interesse devem fazer parte dos encaminhamentos da pesquisa. Dentre as mais corriqueiras, há necessidade de verificar a distribuição das funções expressas por essas formações em termos de classes gramaticais, para se identificar qual a classe mais propícia as mesmas. Os dados de distribuição e frequência dos itens em *-inho* podem ser subdivididos ainda em termos de sexo e de fatores sociais, como o nível de escolaridade e a faixa etária na avaliação da produtividade e função dessas formações. Outro ponto que pode servir de desdobramentos para a presente pesquisa seria a subdivisão do *corpus* escrito nos diversos gêneros que o compõem para se poder esclarecer, dentre outros aspectos, a relevância ou não da alta frequência da função afetiva nesse tipo de *corpus*. Há que ser feito, num estudo posterior não coberto por esta dissertação, um levantamento da frequência e distribuição de formações em *x-inho* que podem ser classificados como itens metafóricos, lexicalizados e polilexicais, para identificar suas funções mais comuns entre os falantes, tanto na modalidade oral como na escrita, bem como, num estudo sobre lexicalização propriamente dita, o traçado dos processos de lexicalização de algumas formações mais populares de *x-inho*, como é o caso de *neguinho*, já exemplificado neste trabalho.

A prescrição gramatical muitas vezes limita nosso entendimento sobre certos aspectos da linguagem. Por isso, o propósito deste trabalho foi apontar a possibilidade de outra perspectiva na compreensão de um fenômeno da língua. Desta forma, a presente dissertação mostrou, através de toda a parte teórica e da análise dos dados colhidos dos *corpora* apresentados, que existe uma relevância em relação ao uso das formações *x-inho*, fazendo-

nos crer que a descrição gramatical pode se beneficiar quando à complexidade da língua em uso é permitida uma abordagem que inclui os vários níveis desse mesmo uso, como foi o caso do *-inho* como atitude avaliativa.

Portanto, os resultados apresentados são somente alguns indicativos de que uma abordagem mais funcional deve ser levada em consideração na explicação de alguns fenômenos lingüísticos, e que a formação *x-inho* não deve ser considerada apenas no âmbito da morfologia, como elemento indicador de dimensão, que é o que mais destacam as gramáticas normativas.

Esperamos, portanto, com o presente trabalho, ter contribuído para o desenvolvimento de estudos na interface da Lingüística de Corpus com a Língua Portuguesa que levem em conta aspectos cotextuais e contextuais para análise dos diversos fenômenos dessa língua, pois, com este trabalho, evidenciamos existir uma real necessidade de investigar o Português sob aspectos que transcendam um olhar puro e simples sobre o sistema da língua.

REFERÊNCIAS BIBLIOGRÁFICAS

ADAM, Jean-Michel.(1987) *Types de séquences textuelles élémentaires. Patiques.* Metz,56:554-79.

ALVES, J. M. (1990) *Neologismos.* São Paulo: Ática.

ANDRÉ, Hildebrando Afonso de. (1978) *Gramática ilustrada.* São Paulo: Ed. Moderna.

AZEREDO, J.C. (2000) *Fundamentos da gramática do português.* Rio de Janeiro: Jorge Zahar.

BACK, E. e GERALDO, M. (1972) *Gramática construtural da língua portuguesa.* São Paulo: FENAME.

BAKCTIN, M. (1995) *Marxismo e folosofia da linguagem.* São Paulo, Hucitec.

BARRERA, L. F. de (1986) Empleo del diminutivo em algunos estados centrales de Venezuela. In: *Actas Del V Congresso Internacional de La Asociación de Linguística y Filologia de América Latina (ALFAL).* Cracas, Universidad Central de Venezuela,p. 321-326

BASÍLIO, M. (1989) *Teoria lexical.* São Paulo: Ática.

_____.(1980) *Estruturas lexicais do português: uma abordagem gerativa.* Petrópolis: Vozes.

BECHARA, E. (2000) *Moderna gramática da língua portuguesa.* 37. ed.. Rio de Janeiro: Lucena.

BERBER SARDINHA, A. P. (2004) *Lingüística de corpus.* Barueri: Manole.

BIBER, D. *et al* (1998) *Corpus linguistics – Investigating Language Structure and use.* Cambridge: Cambridge Iniversity Press.

BIBER, D. (1988) *Variation across speech and writing.* Cambridge. Cambridge University Press.

CAMARA JR., J. M.(2001) *Estrutura da língua portuguesa.* 34. ed. Petrópolis:Vozes.

_____. (1973) *Dicionário de Filologia e gramática*. Rio de Janeiro: J. Ozon.

CASTILHO, Ataliba Teixeira de. (2003) *A língua falada no ensino de português*. 5. ed. São Paulo: Contexto.

CEGALLA, Domingos Paschoal. (1977) *Novíssima gramática da língua portuguesa*. São Paulo: Ed. Nacional.

CHARAUDEAU, P. & MAINGUENEAU, D. (2002) *Dicionário de Análise do Discurso*. Paris: Seuil

CHARAUDEAU, P. (1997) *Médias Recherches*. France: Nathan.

_____. (1992) *Grammaire du sens et de l'expression*. Paris: Hachette.

CHOMSKY, N. (1957) *Syntactic Structures*: New York: Mouton.

CUNHA, C. F. da. (1972) *Gramática da língua portuguesa*. Rio de Janeiro: MEC/FENAME.

DURANTI, A. GOODWIN, C. (1994) *Rethinking context Cambridge*: Cambridge University Press.

EGGINS, S. (1994) *An introduction to systemic functional linguistics*. London: Pinter Publishers.

EGGINS, S. e MARTIN, J. R. (1997) "Gêneros y registros del discurso". In: DIJK, Teun van (org.). *El discurso como estructura y proceso*. Estudios del discurso: Introducción multidisciplinaria 1, Barcelona: Gedisa, p 335-371.

EMÍLIO, Aline. (2003) Diminutivo x grau normal: um fenômeno estilístico no enfoque da abordagem variacionista. *Revista da ABRALIN*, vol. II, nº 1, p. 9-49.

ERMAN, B. E WARREN, B. (2000) "The idion principle and the open-choice principle". *Text*, 20. P. 29-62

EZARANI, E. S. (1989) *Formações x-inho na fala carioca*. Dissertação de Mestrado em Letras. Rio de Janeiro: UFRJ/Faculdade de Letras.

FILMORE, C. (1992) Corpus linguistics or computer corpus linguistics. In: *Directions in corpus linguistics*. Proceedings of nobel symposium 82, Stóckholm, Ed. Jan Avartvik, 35-60. Berlin/Nova York, De Gruyter.

FROTA, M. P. (1985) *A expressão do pejorativo em construções morfológicas*. Dissertação de mestrado. Rio de Janeiro: UFRJ.

GHADESSY, M.; HENRY, A.; ROSEBERRY, R. L. (orgs.) (2001) *Small corpus studies and EFL: theory and practice*. Amsterdã/Filadélfia: Jonh Benjamins.

GONÇALVES, C. A. V. (2000) *Sufixos intensivos: reforço prosódico e função indexal*. Comunicação apresentada no Work-Shop “Mosfopragmática da (hiper) expressividade”. Rio de Janeiro: PUC-RIO, mimeo.

GRANGER, S. (org.) (1998) *Leaner english on computer*. New York: Longman.

HAKAMIES, R. (1951) *Étude sur l'origine et l' evolution du diminutif latin et survie dans les langues romanes*. Helsinki: Academiae Scientiarum Fennicae.

HALLIDAY, M. A. K. (1985a) *An introducion to funccional grammar*. London: Edward Arnold.

_____. (1985b) *Spoken and written language*. Oxford: Oxford University Press.

HALLIDAY, M. A. K. e HASAN, R. (1989) *Language context and text: aspects in a social-semiotic perspective*. Oxford: Oxford University Press.

HOEY, M. (1997) “From concordance to text atructure: New uses for computer *corpora*”. In: LEWAANDOWSKA – TOMASZCZYK, B.: MELLA, P. J. (orgs.) *PALC 97 – Pratical applications in language corpora*. Lodz, Lodz University Press.

INFANTE, U. (1995) *Curso de gramática aplicada aos textos*. Rio de Janeiro: Scipione.

KENNEDY, G. (1998) *An introducion to corpus linguistics*. Londres: Longman.

KOCH, I. G. V. e FÁVERO, L. L. (1987) *Contribuição a uma tipologia textual*. Letras & Letras: Uberlândia.

LABOV, W. (1972) *Language in the inner city: studies in the Black vernacular*. Philadelphia: University of Pennsylvania Press.

LAPA, M.R. (1998) *Estilística da língua portuguesa*. São Paulo: Martins Fontes.

LEECH, G. (1991) *Corpora*. In K. Malmkjaer (ed) *The Linguistics encyclopedia* (pp. 73-80). London: Routledge.

LIMA, C. H. R. (1994) *Gramática normativa da língua portuguesa*. 32. ed. Rio de Janeiro: José Olympio.

LONGACRE, R. (1983) *The grammar of discourse*. New York: Plenum.

LOURES, L. H. (2000) *Análise contrastiva de recursos morfológicos com função expressiva em francês e português*. Tese de Doutorado em Linguística. Rio de Janeiro: UFRJ/Faculdade de Letras.

LUFT, C.P.(2002) *Moderna gramática portuguesa*. 2. ed. São Paulo: Globo.

MACAMBIRA, José Rebouças.(1917) *A estrutura morfo-sintática do português*. 9.ed. São Paulo: Pioneira, 1999.

MAINGUENEAU, D. (2002) *Análise de textos de comunicação*. Tradução de Cecília P. de Souza-e-Silva, Décio Rocha. São Paulo: Cortez.

MARCUSHI, L. A. (2002) *Gêneros textuais: Definição e funcionalidade*. In: DIONÍSIO, Ângela Paiva, MACHADO, Anna Raquel e BEZERRA, Maria Auxiliadora (orgs.) *Gêneros textuais e ensino*. Rio de Janeiro: Lucerna.p.19-36.

_____ (2000) *Gêneros textuais: O que são e como se classificam*. Recife: texto mimeografado.

MARTIN, J. R. (2000) *Grammar meets genre: Reflections on the Sydney School*. Inaugural Lecture on the Sydney University Arts Association.

McENERRY, T. e Wilson, A. (1996) *Corpus linguistics*. Edinburgh: Edinburgh University Press.

MONTEIRO, J. L. (1987) *Morfologia portuguesa*. Campinas: Pontes.

MORENO, C. (1977) *Os diminutivos em -inho e Zinho e a delimitação do vocábulo nominal no português*. Porto Alegre: UFRGS. Dissertação de mestrado.

NEVES, M. H. de M. (1997) *A Gramática Funcional*. São Paulo: Martins Fontes.

OLIVEIRA, H. F. (2004b) Os gêneros da redação escolar e o compromisso com a variedade padrão da língua. In: HENRIQUES, Cláudio Cezar & SIMÕES, Darcília, orgs. *Língua e cidadania: novas perspectivas para o ensino*. Rio de Janeiro, Europa, p.183-193.

OLIVEIRA, M. M. (1962) *Processos de intensificação no português contemporâneo* (a entonação, processos morfológicos e sintáticos). Lisboa: Centro de Estudos Filológicos.

PÉCORA, A. (1980) *Problemas de redação na universidade*. CAMPINAS, Editora da UNICAMP.

PERINI, M. A. (1980) *A contribuição da lingüística no ensino de língua portuguesa*. Belo Horizonte, UFMG. Mimeo.

PIZA, M. (2001) *Gênero, número e grau no continuum flexão/derivação em português*. Rio de Janeiro: UFRJ/Faculdade de Letras. Dissertação de Mestrado em Língua portuguesa.

RIBEIRO, M. P. (1998) *Gramática aplicada da língua portuguesa*. 10. ed. Rio de Janeiro: Metáfora.

RICHARDS, J. C. ,PLATT, J. e PLATT, H. (orgs.) (1992) *The Longman Dictionary of Language Teaching and Applied Linguistics*. London: Longman.

ROCHA, L. C. (1994) *Flexão e derivação no português*. Cadernos de pesquisa NAPq. Belo Horizonte: FALE/UFMG, nº 19.

_____ (1986) *Estudo da produtividade dos nomes diminutivos em -inho (-zinho)*. Trabalho final apresentado à professora Margarida Basílio na disciplina Tópicos Especiais, código LEF 818. Rio de Janeiro: UFRJ, Faculdade de letras.

ROSA, M. C. (1982) *Formação de nomes aumentativos: estudo da produtividade de alguns sufixos portugueses*. Dissertação de Mestrado em Lingüística. Rio de Janeiro: UFRJ/Faculdade de Letras.

SACCONI, L. A.(1991) *Nossa gramática: Teoria*. 11. ed. São Paulo: Atual.

- SANCHES, A. (1995) "Definición e historia de los corpus". In: SANCHES, A. et al. (orgs.). *CUMBRE: corpus lingüístico de Español contemporáneo*. Madri, SGEL. p.7-24.
- SANDMANN, A. J. (1989) *Morfologia geral*. São Paulo: Contexto.
- _____ (1992) *Morfologia lexical*. São Paulo: Contexto.
- SANTANN, M. N. (1989) *Introdução à estilística: expressividade da língua portuguesa*. Rio de Janeiro: Ed. Universidade de São Paulo.
- SAUSSURE, F. (1972) *Curso de lingüística geral*. São Paulo: Cultrix.
- SCHIFFRIN, D. (1994) *Approaches to discourse*. Oxford: basil blackwell.
- SCOTT, M. (1998) *WordSmith Tools version 3*. Oxford: Oxford University Press.
- SKORGE, S. (1963) *Os sufixos diminutivos em português*. In: Boletim de Filologia, tomo XVI, fascículos 3 e 4. Lisboa, Centro de Estudos Filológicos.
- SINCLAIR, J. McH. (1991) *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- THOMPSON, G. (1996) *Introducing functional grammar*. Edward Arnold.
- THORNDIKE, E. L. (1921) *A Teacher`s wordbook*. Nova York: Columbia Teachers Colege.
- TOGNINI-BONELLI, E. (2001) *Corpus linguistics at work*. Amesterdã/Atlanta: John Benjamins.
- TORREIRA, R.Q. (1993) *O parágrafo e o texto*. Rio de Janeiro: UFRJ. Tese de Doutorado.
- TRAVAGLIA, L. C. (2000) *Gramática e interação: uma proposta para o ensino de gramática no 1º e 2º graus*. 5 ed. São Paulo: Cortez.
- VOTRE, S. e OLIVEIRA (1995) *Corpus Discurso e Gramática*. UFRJ.
- WERLICH, E. (1975) *Typologie des textes: Entrouf eines linguistisches Models zur Grundlegung einer Textgrammatik*. Heidelberg.
- ZANOTTO, N. (1985) *Estruturas morfológicas do português*. Porto Alegre: EDUCS.

RESUMO

As formações *x-inho* são descritas, na maioria das gramáticas de Língua Portuguesa como contendo noções dimensiva e afetiva. Entretanto, essas mesmas gramáticas não incluem os fatores extralingüísticos e contextuais nos quais os enunciadores estão inseridos quando optam por uma formação em *x-inho*. Sob esta perspectiva, tem-se no presente trabalho, o objetivo de investigar a produtividade das formações *x-inho* em dois *corpora* eletrônicos: um oral, subdividido em dois *subcorpora* contendo respectivamente narrativas e descrições e um escrito, oriundo exclusivamente das variadas seções e cadernos de um jornal de grande circulação e qualidade. A dissertação quantifica as ocorrências das formações *x-inho* em cada um dos *corpora*. Em seguida cada uma dessas ocorrências é analisada para se verificar que tipo de noção (dimensiva, afetiva positiva ou negativa, intensificadora, etc) ela contém. Por fim são contrastados os dados de frequência e dispersão de cada uma das noções encontradas para cada um dos *corpora*. A metodologia de nossa análise está centrada na área de investigação lingüística denominada Lingüística de *corpus*, que serve de base para que os dados colhidos sejam analisados e interpretados.

ABSTRACT

The items ending in *-inho* are described in the majority of grammars of Portuguese as conveying two notions, namely affect and dimension. However, the same grammars do not seem to include either the extralinguistic or contextual factors in which speakers are inserted when they opt for a word ending in *-inho*. The aim of the present work thus is to investigate the productivity of such items in two electronic *corpora*: one of an oral nature which is further subdivided into two sub-*corpora* containing narratives and descriptions and a second one compiled exclusively from the various sections of a widely read quality newspaper. The dissertation quantifies the various instances of items ending in *-inho* in each of the *corpora*. Next, each of these occurrences is analysed and classified to check which notion (dimentio, positive affect, negative affect, intensification) they convey. Last the results of both frequency and dispersion counts are contrasted for each of the *corpora*. The methodology of our analyses is centered on the area known as Corpus Linguistics, which provides a basis for the data to be compiled and interpreted.

