

CADEIAS REFERENCIAIS NO PORTUGUÊS FALADO

Iara Bemquerer Costa

RESUMO: *This paper aims at identifying the set of factors which determine the choice among alternative forms of remissive references in spoken Portuguese. For that purpose, we analysed 100 texts extracted from recordings of oral interviews, of which 50 are narrative and 50 argumentative. The empirical analysis of the forms used to express co-referentiality indicates that the most relevant factor determining the selection of a specific form is the distinction between animate and inanimate referents.*

PALAVRAS-CHAVE: coesão, correferencialidade, oralidade.

INTRODUÇÃO

Pretende-se, neste artigo, retomar o conceito de coesão textual (HALLIDAY & HASAN, 1976), procurando aplicá-lo de forma sistemática ao texto oral. Procura-se em especial colocar em foco as redes de relações coesivas, responsáveis pela indicação de correferencialidade, encontradas de forma recorrente em textos narrativos e argumentativos, com o objetivo de identificar os fatores que atuam na construção de cadeias coesivas ao longo desses tipos de texto.

Os textos estudados fazem parte do acervo do Banco de Dados Lingüísticos VARSUL¹, constituído de entrevistas realizadas entre 1990 e 1995 com falantes de 12 centros urbanos da região Sul. O recorte dos textos estudados no conjunto das entrevistas foi direcionado por um conjunto de quatro critérios. Em primeiro lugar, o início e o final de cada texto foram demarcados a partir do turno conversacional e da unidade temática. Mais precisamente, tomou-se como referência para o recorte de fragmentos das entrevistas, em primeiro lugar, o momento em que o entrevistador propõe alguma questão ao entrevistado e este a aceita, dando início a uma sequência de proposições que mantêm uma unidade temática; de forma análoga, o final do texto coincide em geral com a retomada do

Iara Bemquerer Costa é professora da Universidade Federal do Paraná.

¹ Informações detalhadas sobre a constituição desse banco de dados estão disponíveis em KNIES & COSTA (1996).

turno pelo entrevistador. Como nem sempre as intervenções do entrevistador são determinantes para a mudança do tópico conversacional (muitas vezes elas marcam o interesse pela continuidade do tópico), a unidade temática é um critério complementar à mudança de turno. O segundo critério foi estabelecido com o intuito de obter um conjunto de textos distribuídos de forma equilibrada entre dois tipos: narrativos e argumentativos. Finalmente, tomou-se como terceiro critério a seleção de um material representativo da diversidade regional e étnica que caracteriza o Banco de Dados Lingüísticos VARSUL. Para isso foram escolhidos textos de diversos informantes, distribuídos aleatoriamente entre os municípios documentados. Finalmente, o nível de escolaridade dos falantes foi também um dos critérios nucleares na seleção da amostra, que contém textos de falantes de nível primário (4 a 5 anos de escolaridade) e de informantes com o segundo grau (10 a 11 anos de escolaridade). Mediante a aplicação desses critérios, foi constituído um *corpus* para estudo com 100 textos, dos quais 50 narrativos e 50 argumentativos e, em cada grupo, 25 produzidos por falantes com escolarização primária e 25 com segundo grau.

O *corpus* assim constituído forneceu a base empírica que tornou possível observar as escolhas feitas sistematicamente pelos falantes ao construir a rede de retomadas referenciais que marcam formalmente, na superfície textual, a manutenção do foco. Permitiu também o teste de hipóteses sobre a influência de fatores diversos na escolha da forma de expressão da coesão.

O interesse pela análise sistemática das marcas coesivas no texto surge com o trabalho de Halliday & Hasan (1976). Referência obrigatória nos estudos textuais, esse trabalho clássico colocou em pauta uma série de questões interessantes. Sua pergunta chave – o que faz com que um texto seja texto? – formulada a partir da pressuposição de que os falantes possuem uma competência textual, análoga à competência gramatical chomskyana, resulta na formulação do conceito de coesão textual. A existência da coesão seria condição necessária (e suficiente) para se distinguir um texto, ou um fragmento de texto, de um conjunto aleatório de sentenças. Para Halliday & Hasan, o texto se define por sua coesão, ou seja, por uma rede de ligações intra e inter-sentenciais, que fazem com que a interpretação de cada parte do texto esteja vinculada a outras. Essas ligações, marcadas formalmente na superfície textual, seriam de natureza semântica.

A partir da questão proposta inicialmente e da resposta dos autores, outras perguntas surgem. O conjunto de mecanismos de coesão textual proposto é adequado para a caracterização dos recursos coesivos? A coesão é realmente uma condição necessária para que uma seqüência de frases possa ser considerada um texto? Se for necessária, será também suficiente? Até que ponto é possível explicar a articulação textual pela análise do texto em si, sem levar em conta os sujeitos envolvidos na sua produção e compreensão e as condições contextuais em que o mesmo foi produzido?

Vários estudos indicam a necessidade de uma redefinição do ponto de vista a partir do qual os fatores responsáveis pela articulação textual são

analisados. Charolles (1988) dispensa o conceito de coesão e propõe para substituí-lo o de coerência, que leva em conta as representações de mundo do falante/autor e do ouvinte/leitor. Seu conceito de coerência micro-textual recobre os casos estudados por Halliday & Hasan sob o rótulo de coesão e é complementado pela formulação de um conjunto de princípios constitutivos da coerência, que dá conta também do nível macro-textual. Lemos (1992) aponta a dificuldade em explicar o funcionamento textual sem levar em conta que um texto só existe se parte de um sujeito e se dirige a outro. Segundo ela, não há o texto e o não-texto em si, mas o texto ou não-texto para alguém, o que evidencia a importância de se revisar os limites propostos para a análise de questões textuais, mediante a mudança do ponto de vista sob o qual tais questões são enfocadas. Em outras palavras, as dificuldades encontradas na caracterização do texto como objeto de estudo só poderiam ser resolvidas mediante a adoção de uma perspectiva constituída a partir da análise do discurso. Possenti (1988) mostra que uma análise nos moldes de Halliday & Hasan (1976) tem um poder explicativo muito baixo. A identificação dos mecanismos coesivos em um texto nada mais é, segundo ele, que um exercício de leitura. Pode-se ganhar na explicação dos fenômenos textuais passando-se a encará-los como questões de discurso. A escolha de mecanismos de coesão textual é associada, em seu trabalho, ao jogo de imagens (PÊCHEUX, 1969) que o falante/autor tem de si, do ouvinte/leitor e do referente.

Há, por outro lado, revisões que mantêm em grande parte o conceito e os critérios de delimitação do texto presentes em Halliday & Hasan (1976): mantêm a superfície textual como objeto de análise e procuram reformular o conceito de coesão e complementá-lo com outros de natureza igualmente formal. Essa tendência está presente nos trabalhos de Koch, Fávero, De Beaugrande & Dressler, por exemplo. Koch (1988, 1990), Fávero (1991), Koch & Fávero (1994) procuram estabelecer de forma mais clara os limites entre os mecanismos coesivos. Procuram ainda ampliar o leque de recursos de coesão, de modo a dar conta de um número maior de processos que os previstos nas propostas anteriores e marcar a diferença entre dois tipos de recursos coesivos (referenciais e sequenciais). Apontam ainda, sob a classificação de *coesão recorrencial* (FÁVERO, 1991), procedimentos de construção da coesão textual de natureza mais formal (sintática e fonológica) que semântica. De Beaugrande & Dressler (1981) propõem um conjunto de fatores de textualidade complementares à coesão. Todos esses estudos fazem uma revisão que mantém em grande parte os pressupostos iniciais: a superfície textual é o objeto sobre o qual dirigem sua observação e o limite para os fenômenos de coesão. Sob esse ponto de vista, os autores mencionados se aproximam de Halliday & Hasan, embora seja relevante apontar que eles atribuem à coesão uma importância bem menor na constituição do texto do que foi atribuída na proposta original. Já não a consideram uma condição suficiente, e nem mesmo necessária, para que se possa distinguir um texto de uma sequência aleatória de sentenças, mas a concebem como um entre os vários fatores de textualidade.

FORMAS DE EXPRESSÃO DA COESÃO

Em qualquer dos trabalhos mencionados, fica claro que a coesão textual é realizada por um conjunto heterogêneo de mecanismos formais. Há, por um lado, um conjunto de recursos intercambiáveis; por outro, formas que dificilmente poderiam ser substituídas por outras equivalentes. Os recursos coesivos intercambiáveis são em especial os responsáveis pela expressão formal no texto das retomadas referenciais, ou, conforme bem apontam Brown & Yule (1983, p.192), pela expressão da correferencialidade ao longo do texto. Como diz com propriedade Van Dijk (1980, p.33):

A função principal da correferência é a de relacionar participantes de fatos conectados, principalmente por meio de uma relação de identidade. Isto significa que o texto contém o que se poderia chamar intuitivamente uma certa continuidade de foco: continuamos falando da mesma coisa ou pessoa, da qual predicamos propriedades e relações distintas.

Entre os recursos coesivos não intercambiáveis estão, por um lado, os articuladores classificados por Halliday & Hasan como *conjunções*, responsáveis por indicar uma relação semântica ou enunciativa entre as proposições. Estão também os recursos de paralelismo formal, tanto sintático quanto fonológico (classificados por FÁVERO, 1991, como casos de coesão recorrencial), e ainda os casos caracterizados como *substituição*, em que não se tem a ligação um a um entre um sintagma nominal e um elemento que o retoma, mas o uso de um elemento coesivo que faz remissão a trechos relativamente extensos do texto. Os procedimentos de construção da correferencialidade intercambiáveis constituem o centro de interesse deste estudo, pois permitem um controle das condições de uso de cada uma das formas concorrentes, para que se possa avaliar a influência de fatores diversos no seu uso.

Possenti (1988) explora a escolha que o falante/autor faz da repetição, anáfora e coesão lexical em sua relação com as representações imaginárias que ele tem de si próprio, do ouvinte e do referente, mediante a utilização do conceito de jogo de imagens de Pêcheux (1969) e apresenta exemplos em que a ambigüidade gerada pelo uso da anáfora e da coesão lexical só é aceitável no texto porque o autor atribui ao leitor o domínio de um certo universo de referência e o partilhamento de algumas proposições.

A questão que direciona o estudo de Possenti (1988, p.96) está relacionada à escolha, feita pelo falante/autor no momento da enunciação, entre os recursos coesivos intercambiáveis, em cada contexto em que a possibilidade de escolha esteja posta: procura “*perguntar a cada um dos elementos de coesão, porque um deles apareceu e nenhum outro em seu lugar*”. Associando a escolha de um ou outro recurso coesivo com a imagem que o falante/autor tem do seu ouvinte/leitor, Possenti (1988, p.99) afirma que

dependendo da imagem que o locutor faz do interlocutor no momento da enunciação de um discurso, utiliza um ou outro elemento coesivo. Se a imagem for positiva, em termos de capacidade de interpretação do

discurso, utilizará um elemento anafórico; se for negativa, uma repetição.

Essa afirmação tem respaldo no pressuposto de que “a repetição é um procedimento que serve a mais de uma finalidade, mas sem dúvida, parece servir a esta: a clareza de interpretação.” (idem, p.99), o que explica a frequência de seu uso em textos de cunho científico ou didático ou dirigidos ao público infantil.

A associação feita por Possenti entre a escolha de certos recursos coesivos e as imagens que monitoram a organização discursiva pode ter um paralelo no grau de planejamento do texto. Nas entrevistas informais gravadas, que constituem o corpus do presente trabalho, se os fatores apontados forem decisivos para a escolha dos mecanismos coesivos usados, esperava-se encontrar a seguinte gradação em frequência de uso dos mecanismos de coesão textual: repetição, anáfora, coesão lexical (uso de sinônimos, hipônimos, hiperônimos e expressões descritivas). Por outro lado, se o processo de escolarização resultar em uma maior consciência no uso das formas lingüísticas de prestígio, especialmente aquelas que se aproximam da escrita, a frequência de uso dos recursos coesivos apontados acima deveria apresentar uma ligeira diferença entre os textos de falantes com escolarização primária e os de falantes com o segundo grau.

Os dois tipos de textos escolhidos para estudo correspondem a formas diferenciadas de organização da sequência de tópicos ao longo do texto. Nos textos narrativos, as redes de retomadas referenciais correspondem, na maioria das vezes, à apresentação e sucessivas remissões aos personagens centrais participantes dos eventos narrados. Nos textos argumentativos, as formas de retomada referencial dependem da estratégia argumentativa adotada, particularmente do uso de estratégias centradas em exemplos particulares ou na generalização. Assim, uma amostra constituída por textos narrativos e argumentativos apresenta, certamente, a possibilidade de ocorrência de cadeias referenciais com características suficientemente diversificadas para que se possa testar que fatores semânticos, formais, pragmáticos ou sociais são relevantes para a escolha da forma como a correferencialidade é expressa.

Os dois fragmentos de textos narrativos abaixo, ambos produzidos por falantes com nível primário de escolarização, contêm exemplos de cadeias referenciais com configurações encontradas sistematicamente nesse tipo de texto. No fragmento 1, a informante relata a perseguição de seu marido por um lobisomem. Foram destacadas as referências feitas ao marido (animado e específico) no episódio.

Texto 1

Uma vez *o meu marido*₁ vinha do quartel e aqui embaixo, é aqui onde tem a Rio Sulense, ali. Ali tinha a casinha alta, que é dessas casas que tinham aqueles pilares altos de madeira (inint)² e ele

² (inint): Codificação utilizada na transcrição para *ininteligível*.

estava sentado em cima. *Ele*₁ diz que quando *ele*₁ vinha vindo, saiu um cachorro grande dali, *ele*₁ disse assim: “Ah, isso... isso aí é lobisomem.” E veio pra... mordendo *ele*₁ e correndo atrás *dele*₁. Aí *ele*₁ diz que *ele*₁ pousou os talheres assim, e quando o lobisomem veio em cima *dele*₁, *ele*₁ fez assim com os talheres, aí o lobisomem foi embora. Aí *ele*₁ disse: “Abre a porta, abre a porta, que o lobisomem vem atrás de mim.” Disse que era um cachorrão desse tamanho. E chegava meia-noite, o homem saía, se vestia de lobisomem. (SCFLP08 - linhas 448-468)³

No fragmento 2, a informante, ao narrar um episódio de sua infância, tem como foco de atenção na narrativa uma sombrinha (um referente inanimado e específico). Foram assinaladas as referências feitas a esse objeto.

Texto 2

Até que chegou um dia que eu ganhei *uma sombrinha linda*₁. *A sombrinha*₁ era maravilhosa, a coisa mais linda. Aí, queria porque queria ir com *a sombrinha*₁ pro colégio, né? E a mãe não deixou porque *Ø*₁ era muito bonita e a mãe já sabia, né? que a bagunça era demais, assim, né? Aí eu disse: “Ah, não, mãe, deixa eu ir com *a sombrinha*₁, deixa.” (inint) deixou eu ir com *a sombrinha*₁. *A sombrinha*₁ era a coisa mais lindinha. Aí eu assim: “Ah, hoje eu vou me vingar de todo mundo.” E quando as pessoas vinham tirar a forra comigo, aí eu tacava *a sombrinha*₁. Final, eu cheguei em casa só com aquela parte de dentro da *sombrinha*₁. Os arames deu tudo... tudo... (inint) Só o cabo da *sombrinha*₁. Aí, eu tive que passar na casa da minha avó antes pra... pra... Sabe como é, né? eu ia chegar em casa com *a sombrinha*₁... (SCFLP01 - linhas 282-302)⁴

Os trechos acima mostram dois tipos de configuração das cadeias referenciais em textos narrativos. No primeiro, após a apresentação do personagem (animado e específico), todas as remissões ao mesmo são feitas com o uso do anafórico *ele*. No segundo, após a apresentação do objeto tomado como foco da narrativa (inanimado e específico), as referências ao mesmo são feitas com o uso sistemático da repetição, havendo apenas um caso de elipse. Esses exemplos sugerem que, ao contrário do que aponta Possenti, a opção por uma ou outra forma possível de se marcar a correferencialidade no texto pode ser influenciada de forma sistemática por fatores semânticos, sintáticos ou pragmáticos, que podem ter um peso maior que a imagem do interlocutor.

³ As referências às entrevistas são feitas conforme a codificação adotada no Banco de Dados Lingüísticos VARSUL (Cf. KNIES & COSTA, 1996). SCFLP08, Santa Catarina, Florianópolis, entrevista 8.

⁴ SCFLP01, Santa Catarina, Florianópolis, entrevista 1.

No texto 3, temos um recorte de texto argumentativo que toma como foco um referente animado e genérico: um empresário qualquer, diante do aumento do salário mínimo.

Texto 3

Porque um... *um cara*₁ que é, vamos supor, entende? *o cara*₁ que é... *ele*₁ é um (inint). Então, *ele*... *ele*₁ pega... *ele*₁ empreita a empreiteira pra fazer alguma coisa (inint). Então, há um mês atrás *ele*₁ pegou uma empreitada, tem quinze empregados. Agora, há um mês *ele*₁ pagava noventa mil, ou o máximo que *ele*₁ pagava eram duzentos mil, e agora com o salário novo? Agora *o cara*₁ (?) lá não vai reajustar o preço que *ele*₁ fez da empreitada que *ele*₁ pegou e *ele*₁ é obrigado a pagar o salário novo pros empregados. Então, o que *ele*₁ vai ter que fazer? Mandar metade embora e fazer a outra metade trabalhar o dobro, pra poder pagar. Então isso aí é uma safadeza isso aí. Não adianta aumentar salário pra... pra agradar uns e... e a metade do... um montão ficar desempregado. Que que adianta? Duzentos e trinta de salário e a metade desempregada, resolve alguma coisa? Não resolve nada. (PRCTB07 - linhas 980-1000)⁵

Comparando-se os textos 1 e 3, vê-se que têm em comum o fato de tomarem como foco referentes animados e de escolherem a anáfora como forma de marcar as retomadas referenciais.

No texto 4, o referente é genérico e inanimado.

Texto 4

É, o dia de hoje, é muito mais fácil que fazer (inint) uvas sem... *o vinho*₁ sem uva, né? e a química está aí pra destruir, né? O problema que está *o vinho*₁ por causa disso, né? porque hoje o vinho se teria... não teria *o vinho*₁ sobrando nem nada. Se *o vinho*₁ fosse feito de uvas puras, entende? puro *vinho*₁ de uva, não teria problema nenhum desses, estaria faltando *vinho*₁. Toda essa sobra de *vinho*₁, esse preço baixo é a desmoralização do *produto*₁, entende? que essas cantinas grandes começaram a trabalhar com química, (inint). Pra fazer um litro de *vinho*₁, pra vocês saberem, precisa um quilo e trezentos gramas de uva, né? senão não dá pra fazer *o vinho*₁ puro, por menos é impossível, né? Então, eles fazem um litro de *vinho*₁ com meio quilo de uva, e olha lá se ainda que depois não adultera $\emptyset$ ₁ com outras coisas mais, né? Então esse problema é que surgiu, entende? que é a desmoralização do *vinho*₁. (RSFLC10 - linhas 95- 115)⁶

⁵ PRCTB07, Paraná, Curitiba, entrevista 7.

⁶ RSFLC10, Rio Grande do Sul, Flores da Cunha, entrevista 10.

Este texto reitera a observação feita em relação ao texto 2: o fato de o referente retomado ser inanimado pode ser relevante para o uso da repetição nas suas retomadas ao longo do texto. Os fragmentos de textos tomados como exemplos sugerem que um critério relevante para a escolha entre as duas formas de retomadas referenciais escolhidas para exemplificação – a repetição e a anáfora – é a oposição entre referentes animados e inanimados.

Para testar essa e outras hipóteses sobre os fatores que podem levar o falante a escolher uma ou outra forma de expressão da correferencialidade no texto oral não planejado, optou-se por fazer um levantamento exaustivo das cadeias referenciais identificadas ao longo dos 100 textos que constituem o *corpus* definido para este estudo. Como se trata da comparação de uso de um conjunto de formas intercambiáveis, observadas a partir de alguns critérios previamente definidos, os dados foram submetidos a uma análise estatística com o uso do pacote de programas VARBRUL⁷.

RETOMADAS REFERENCIAIS: ESCOLHA ENTRE FORMAS ALTERNATIVAS

Para a observação sistemática das retomadas referenciais feitas ao longo de cada texto, foram consideradas as seguintes possibilidades: repetição, anáfora, elipse, expressão descritiva:

a) Repetição:

Minha mãe tinha um cachorro₁ bem grande, sabe? Chamava-se Nero. Os bandidos entraram, o cachorro₁ avançou com tudo.

b) Anáfora:

Minha mãe tinha um cachorro₁ bem grande, sabe? Chamava-se Nero. Os bandidos entraram, ele₁ avançou com tudo.

c) Elipse :

Minha mãe tinha um cachorro₁ bem grande, sabe? Chamava-se Nero. Os bandidos entraram, Ø₁ avançou com tudo.

d) Expressão descritiva (coesão lexical, segundo HALLIDAY & HASAN):

Minha mãe tinha um cachorro₁ bem grande, sabe? Chamava-se Nero. Os bandidos entraram, o bicho₁ avançou com tudo.

É importante ressaltar que só foram incluídos no levantamento os casos de elipse em que esta poderia ser substituída por qualquer um dos outros tipos estudados. Cada caso de elipse foi submetido a um teste de substituição pelas demais formas de retomadas referenciais antes de ser incluído no levantamento estatístico, para averiguar se a ocorrência da elipse se devia a condicionamentos estruturais, que tornassem seu uso obrigatório ou se havia a possibilidade de escolha de outra forma no mesmo contexto. Só foram submetidos à análise estatística os casos em que se pode reconhecer, ainda que de forma grosseira, a

⁷ PINTZUK, Susan. VARBRUL Programs. 1988. Inédito.

correferencialidade ao longo da cadeia de remissões. No texto 2, por exemplo, é possível garantir que o objeto *sombrinha* mantém-se o mesmo nas diversas remissões? De forma análoga, no texto 4, só é possível considerar que há correferencialidade nas remissões a *vinho* a partir de uma concepção frouxa de *identidade de referência*.

Cada ocorrência de retomada referencial foi então observada a partir de sete grupos de fatores que poderiam influir na escolha da forma como a correferencialidade é expressa: 1) a característica $[\pm \text{animado}]$ do referente; 2) a especificidade; 3) as características formais do sintagma nominal retomado; 4) a função sintática do elemento coesivo; 5) o paralelismo na forma da retomada referencial; 6) o tipo de texto; 7) o nível de escolaridade do informante.

O detalhamento sobre a caracterização de cada um desses grupos de fatores e sobre sua atuação na escolha das formas de retomada referencial é apresentado a seguir. Dada a limitação deste texto, optou-se por apresentar somente os resultados quantitativos dos fatores que se mostraram relevantes para a escolha das formas de expressão das retomadas referenciais.⁸

A retomada de referentes animados e inanimados

O primeiro fator examinado como possível condicionante da escolha das formas de retomada referencial diz respeito a uma característica semântica: a distinção entre referentes animados e inanimados. Os exemplos de textos citados acima apontam para a importância dessa propriedade na escolha entre a anáfora e a repetição. Para a análise desse fator, o ponto de partida foi a distinção feita por Oliveira e Silva (1991, p.93-4), em estudo sobre o uso dos pronomes possessivos de terceira pessoa no português do Brasil, entre referentes: *humanos*; *inanimados*; *animais*; *conjunto inanimado com elementos humanos*. A distinção *humano x animal* não foi levada em conta neste estudo, uma vez que na amostra estudada há poucos textos que tomam animais como foco, e nesses a forma pela qual a retomada referencial é feita assemelha-se às retomadas de referentes humanos. As demais distinções utilizadas pela autora foram incorporadas à análise estatística. Assim, foi feita a distinção sistemática entre referentes animados (*minha mãe*, *Joaquim*, *a cobra*), inanimados (*a casa*, *a sombrinha*) e conjuntos inanimados com elementos humanos (*clube*, *repartição*, *governo*). A influência da característica $[\pm \text{animado}]$ na escolha da forma de expressão da correferencialidade nos textos orais está refletida na Fig.1.

⁸ A análise completa está no relatório do projeto “A articulação textual do português falado”, apresentado ao CNPq.

Influência da distinção animado/inanimado na forma das retomadas referenciais (VARSUL)

	Repetição		Anáfora		Expressão descritiva		Zero	
	%	Peso relativo	%	Peso relativo	%	Peso relativo	%	Peso relativo
Animado	317/1807 18%	0,096	1195/1807 66%	0,372	111/1807 6%	0,337	184/1807 10%	0,195
Conjunto inanimado com elementos humanos	47/89 27%	0,297	33/89 37%	0,281	3/89 3%	0,231	6/89 7%	0,191
Inanimado	220/285 77%	0,415	34/285 12%	0,114	7/285 2%	0,153	24/285 8%	0,318
TOTAL	584/2.181 27%		1262/2.181 58%		121/2.181 6%		214/2.181 10%	

Fig.1

Pode-se observar inicialmente que, embora se tenha a possibilidade de fazer as retomadas referenciais através de quatro tipos, na oralidade a polarização se dá entre a repetição (27% do total de ocorrências) e a anáfora (58%). A retomada por zero é, em grande parte condicionada por fatores estruturais, que permitem esse tipo de retomada em contextos em que não se tem a possibilidade de ocorrência das demais formas consideradas neste estudo. O baixo índice de utilização da elipse como forma de retomada referencial (10% do total de casos incluídos no levantamento estatístico) tem que ser relativizado: tem-se, sim, esse baixo índice em contextos em que o zero é intercambiável com a repetição, a anáfora e as expressões descritivas.

Como se vê na Fig. 1, o uso de expressões descritivas é muito pequeno para todos os tipos de referentes (6% do total de casos). Isso pode ser uma indicação de que esse tipo de retomada referencial seja característico dos textos escritos. Possenti (1988, p.95), afirma:

em cento e cinquenta narrativas orais pesquisadas encontrei este dado curioso: a coesão léxica está praticamente ausente (pode-se concluir?) da oralidade, exceto por casos como “o cara”, “o homem”, etc., que, a rigor, são mais propriamente pronomes, isto é, anafóricos, do que outros nomes com certa relação semântica com os termos que correferem. Seria, então, a coesão léxica uma das características do “estilo escrito”?

Observando nas narrativas a escolha dessas expressões descritivas, vemos que em geral elas pouco contribuem para a apresentação do referente, o que ratifica a observação de Possenti. Em geral, o que se tem são formas que acrescentam muito pouco às informações contidas na expressão usada para colocar em foco o referente: *a menina, o médico, o rapazinho, a mulher*; às vezes a referência é feita com o nome próprio do personagem ou com uma indicação da mudança do seu estatuto no decorrer da narrativa: *meu noivo - meu marido*.

Há apenas um caso em que o uso de expressões descritivas representa uma real contribuição informativa ao texto: as retomadas ofensivas. Por exemplo, numa

narrativa em que a falante relata sua discussão com a amante do marido⁹, esta é apresentada como *aquela praga* e retomada com expressões como *diabo*, *aquela puta*, *essa praga*. Em outro texto, quando a falante narra um episódio em que seu pai perdeu todos os bens devido à desonestidade de um sócio¹⁰, o personagem é apresentado como *um homem* é retomado como *esse cara*, *aquele cachorro*.

Ao contrário das hipóteses levantadas por Possenti (1988), os textos examinados indicam uma especialização de funções na linguagem oral: as anáforas servem principalmente para a referência a seres humanos e animais, enquanto as repetições têm preferência quando se trata de fazer a remissão a objetos. A tendência à especialização de uso da repetição e anáfora na oralidade é clara, quando se leva em conta os dados em seu conjunto: os pesos relativos expressam a nítida preferência pela retomada referencial com o uso da anáfora para referentes animados (0,372 de usos de anáfora contra 0,096 de repetições) e da repetição para os referentes inanimados (0,415 de repetições contra 0,114 de anáforas).

Embora se note essa tendência à especialização de uso na oralidade, a característica [$\pm$ animado] do referente é apenas um dos fatores que atuam na escolha entre a anáfora e a repetição. Esta tem como característica uma maior facilidade de processamento, por ser constituída de itens lexicais plenos e não de formas pronominais. Assim, em situações em que o falante tem consciência de que o uso da anáfora cria ambigüidades em sua fala, o uso da repetição com referentes animados se coloca como uma alternativa disponível para solucionar problemas de clareza. É o que se nota no trecho a seguir:

Texto 5

Não vê *aquela moça*₁ lá no Rio também, que a polícia atirou $\emptyset$ ₁.[...] *Ela*₁ estava indo pra festa de formatura *dela*₁ na moto com *um amigo*₂. E diz que quando passaram por *um guarda*₃, *o guarda*₃, fez sinal pra *eles*_{1,2} pararem, e *o rapaz com a moto*₂ não parou, e *o guarda*₃ atirou. Ontem passou até o filme... o ... o crime reconstituído, né? no Fantástico. Daí *o guarda*₃ deu um tiro, acertou nas costas *dela*₁ e *ela*₁ caiu da moto. Mas atrás já vinha vindo outros amigos que estavam indo pra formatura de carro. Daí ainda *o rapaz*₂ diz que ainda falou: “Levem *ela*₁ pro hospital, que eu já vou indo”, mas *ele*₂ desapareceu. E agora, ontem, não, *ele*₂ apareceu, daí $\emptyset$ ₂ entregou-se. Diz que *ele*₂ fugiu de apavorado que *ele*₂ ficou. *Ele*₂ não tinha condição de... de... de prestar declaração. E ontem $\emptyset$ ₂ apareceu, *ele*₂ é chileno, até um rapaz bonito que só! O guarda diz que fez sinal pra *eles*_{1,2} pararem e *ele*₂ não parou. Agora *ele*_{3(?)} falou, como é que *ele*_{3(?)}... O *guarda*₃ falou que diz que *eles*_{1,2} quiseram atropelar *o guarda*₃,

⁹ PRPB05 - Paraná, Pato Branco, entrevista 5, linhas 658-929

¹⁰ PRIRT06 - Paraná, Irati, entrevista 6, linhas 121 a 283.

que o *guarda*₃ atirou porque *eles*_{1,2} queriam atropelar $\emptyset$ ₃.”
(PRCTB20 - linhas 636 a 672)¹¹

No texto 5, o falante toma como foco de sua narração três personagens (referentes animados). A clareza exige, naturalmente, que as referências a cada um dos três sejam distintas de alguma forma. Quando a referência é feita à personagem do sexo feminino, a questão da ambigüidade não se coloca; nesse caso, o falante recorre ao procedimento usual nos textos orais: faz inicialmente uma apresentação do referente – *aquela moça lá no Rio* – a seguir retoma-o sistematicamente com o uso de anáforas, usando eventualmente uma elipse. A questão que se apresenta para o falante é evitar que as referências aos dois personagens do sexo masculino tenham problemas de clareza. Já no início, ele escolhe a estratégia da repetição para a indicação de um dos referentes – o guarda – e a anáfora para o outro – o amigo da moça assassinada. A partir de um certo ponto, ao colocar em primeiro plano na narrativa o guarda, o falante tende a empregar a forma de retomada referencial usada correntemente para a indicação de referentes animados. Isso o faz iniciar uma frase fazendo referência ao guarda por uma anáfora, mas, ao se dar conta da ambigüidade, interromper a frase iniciada e adotar a estratégia de empregar sistematicamente a repetição como forma de referência ao personagem em questão.

A alternância entre o uso da anáfora e da repetição é uma estratégia bastante utilizada para evitar a ambigüidade na retomada de referentes animados. Havendo dois referentes animados em foco simultaneamente, é comum o falante fazer escolhas diferenciadas na forma de retomada de cada um deles: enquanto um é indicado sistematicamente pela repetição, para o outro a anáfora é a forma escolhida.

No entanto, a resolução das ambigüidades não é tão crucial no discurso oral não planejado, que apresenta muitos casos de ambigüidades não resolvidas com recursos lingüísticos, situações em que fica a cargo do ouvinte identificar os referentes das retomadas anafóricas. Isso pode ser observado no Texto 6, em que o pronome *ele* ora é utilizado para a referência ao noivo, ora ao pai da informante, sem que a resolução dos casos potencialmente ambíguos seja um problema para o interlocutor.

Texto 6

Aí depois *ele*₁ voltou, aí a gente se encontrou de novo. Aí nós começamos a namorar de novo. Mas *o meu pai*₂ e a minha mãe... mas a minha mãe era muito contra. Ah, minha mãe brigava muito, muito, muito. Então uma vez *o meu pai*₂ estava bêbado. Era uma meia-noite. *Ele*_{1(?)} estava lá em casa. E nós éramos noivos, dois anos noivos. E a minha mãe começou a falar, falar. *O meu pai*₂ quase matou *ele*₁, né? *o meu noivo*₁. *O meu pai*₂

¹¹ PRCTB20 - Paraná, Curitiba, entrevista 20.

pegou *ele*₁, porque *meu pai*₂ era... *ele*₂ não era alto, mas *ele*₂ era forte, o *meu pai*₂ é gordo assim, né? *Ele*_{2(?)} pegou *ele*_{1(?)} e queria estrangular *ele*_{1(?)}. (SCBLU06 - linhas 457 a 470)¹²

As possíveis ambigüidades assinaladas com interrogações no texto se resolvem pelas demais informações contextuais. No primeiro caso indicado (*Ele*_{1(?)} *estava lá em casa.*), a anáfora poderia retomar tanto o pai quanto o noivo. A coerência da narrativa aponta para a referência ao noivo, embora o sintagma nominal mais próximo seja *o meu pai*. O segundo caso de ambigüidade apontado (*Ele*_{2(?)} *pegou ele*_{1(?)} *e queria estrangular ele*₁.) se resolve pela sua associação a uma frase que se encontra pouco acima no texto: *O meu pai*₂ *quase matou ele, né? o meu noivo*₁.

Se a maior facilidade de processamento pode ser uma arma a favor do uso da repetição como forma de retomada referencial em todos os contextos, há, por outro lado, uma força contrária ao seu uso generalizado, que vem da pressão da escola, que insiste em corrigir a repetição, considerada um vício de linguagem.

Mas deve haver algo mais na percepção que os falantes têm dos pronomes que faz com que o uso das anáforas na oralidade tenha essa tendência a se especializar como a forma preferencial de retomada de referentes animados. Benveniste (1976, p.282-3) aponta que todas as línguas têm uma diferenciação no conjunto de pronomes pessoais: há, de um lado, as formas pronominais de primeira e segunda pessoas, que são propriamente pessoas, passíveis de se alternarem no papel de enunciador; de outro, a terceira pessoa, mais adequadamente caracterizada como a não-pessoa, já que não pode compartilhar com a primeira e a segunda pessoa da mesma alternância de papéis.

Apesar da pertinência da formulação desse autor, não se pode negar que para os falantes do português é relevante a inserção da forma pronominal *ele* no mesmo paradigma que *eu/tu* (ou *eu/você*). Nos textos orais, os referentes retomados preferencialmente pelo pronome *ele* (-a,-s,-as) são aqueles que potencialmente poderiam estar na posição de primeira e segunda pessoas: seres animados que poderiam eventualmente estar nas posições de falante ou ouvinte. Assim, é possível supor que na fala oral não planejada, o falante utilize um princípio geral do tipo: *é adequado fazer as remissões textuais a pessoas pela utilização de pronomes, mas para as não-pessoas deve-se evitar o uso pronominal*. Esse parece ser o princípio mais forte a direcionar o processo de especialização que se percebe no uso da anáfora.

Especificidade e definição formal dos referentes

A distinção entre referentes genéricos e específicos feita aqui tem como ponto de partida o estudo de Oliveira e Silva (1996) sobre o uso de possessivos de

¹² SCBLU06 - Santa Catarina, Blumenau, entrevista 6.

terceira pessoa. Ao classificar os referentes aos quais os possessivos remetem (os possuidores), a autora distingue entre seis graus de especificidade:

1. Totalmente geral: *todos, cada, alguém*;
2. Indefinido, não referencial: *um cara, mulher*;
3. Definido, não referencial: *o cara, ser humano*;
4. Grupo definido, referencial, de tamanho ilimitado: *os padres, as crianças*;
5. Grupo de tamanho limitado: *meus netos, meus alunos*;
6. Totalmente definido: nome próprio, pronome *ele*. (OLIVEIRA e SILVA, 1996, p.178)

A gradação utilizada pela autora combina dois critérios: a distinção semântica entre o uso atributivo e referencial do sintagma nominal e a presença de marcas de definição no sintagma nominal. O primeiro critério apontado acima pode ser explicitado a partir da caracterização de descrições definidas atributivas e referenciais feita por Borges Neto (1986, p.145):

Dito rapidamente, fazemos um uso referencial de uma descrição definida quando a usamos para chamar a atenção do ouvinte para uma determinada coisa ou pessoa (a descrição é um entre vários modos que a língua tem para identificar indivíduos e pode ser substituída por um nome próprio, por um pronome, por outra descrição definida, etc., sem prejuízo da comunicação). Nos casos em que o uso é atributivo, por outro lado, a descrição definida é essencial e não pode ser substituída sem prejuízo da comunicação. Sua função não é mais identificar um indivíduo, mas permitir que se afirme algo sobre quem quer que satisfaça a descrição. Enquanto uma descrição definida, usada referencialmente, envolve uma pressuposição de que existe algo ou alguém *em particular* que a satisfaz, a mesma descrição, usada atributivamente, dispensa tal pressuposição.

Enquanto Oliveira e Silva constitui uma escala de especificidade mediante a combinação de dois tipos de características, no presente estudo, optou-se por examinar separadamente a especificidade e presença de marcas formais de definição do referente, a partir da observação de algumas pistas no *corpus* e na bibliografia consultada, que apontam para uma diferenciação no peso que essas duas características têm na escolha da forma de retomada referencial.

O uso do fator especificidade permitiu fazer uma distinção entre dois tipos de referentes: específicos e não específicos. O grupo classificado como *não específico* reúne os quatro primeiros níveis da escala de Oliveira e Silva, ou seja, os sintagmas nominais não referenciais e os referenciais que designam grupos de tamanho ilimitado. A classificação *específico* recobre os níveis cinco e seis da escala: sintagmas nominais referenciais, que remetem a conjuntos de tamanho limitado ou a seres individualizados.

Para a classificação dos dados em relação às características formais do sintagma nominal retomado, foram estabelecidas as distinções abaixo, a partir do estudo de Oliveira e Silva (1991, p. 97):

1. Formalmente definido: nome precedido por artigo definido, nomes próprios, pronomes pessoais, nomes precedidos por numerais;
2. Formalmente indefinido: nomes contáveis plurais sem artigo, nomes precedidos de artigo indefinido ou pronome indeterminado.

A separação entre a especificidade e definição formal do referente como fatores distintos foi decidida a partir da observação empírica de que nem sempre a forma pela qual o referente é introduzido é relevante para as retomadas anafóricas. Observe-se no texto 7 a ausência de concordância entre as anáforas e o sintagma nominal utilizado para a introdução do referente no discurso:

Texto 7

Que ele morava em Portugal, eles eram cinco irmãos. E *a mãe*₁, quando iam trabalhar na lavoura, *eles*₁ amarravam todos eles, um com o outro. Um no outro *eles*₁ amarravam eles e $\emptyset$ ₁ deixavam uma sardinha e um pãozinho pra cada um, e eles passavam todo o dia assim. (RSFLC04- linhas 1031 a 1037)¹³

Esse exemplo sugere que o falante, ao fazer as retomadas referenciais na oralidade, leva em conta as características semânticas do referente, mas não necessariamente a forma do sintagma nominal utilizado para colocá-lo em foco. Como bem observam Brown & Yule (1983, p.200):

Vamos sugerir que parece mais provável que o processador¹⁴ estabelece um referente em sua representação mental do discurso e relaciona as referências subsequentes a essa representação mental e não à expressão verbal original empregada no texto.

Para a análise dos fatores especificidade e presença de marcas formais de definição, foi observado sempre o sintagma nominal que coloca o referente em foco no discurso. A análise estatística da presença x ausência de marcas formais de definição nesse sintagma confirmou a hipótese formulada a partir de casos semelhantes ao Texto 7: esse não é um fator relevante na escolha da forma de retomada referencial. O abandono da hipótese de que esse possa ser um fator relevante para a forma das retomadas referenciais confirma a hipótese de Brown & Yule.

A Fig. 2 mostra a influência da especificidade do referente na forma da retomada referencial:

¹³ RSFLC04, Rio Grande do Sul, Flores da Cunha, entrevista 4.

¹⁴ BROWN & YULE referem-se nesse trecho ao processo de interpretação do discurso oral ou escrito e usam a expressão *processador* (processor) para designar o ouvinte/leitor.

Influência da especificidade do referente na forma das retomadas referenciais (VARSUL)

	Repetição		Anáfora		Expressão descritiva		Zero	
	%	Peso relativo	%	Peso relativo	%	Peso relativo	%	Peso relativo
Específico	369/1589 23%	0,225	1010/1589 64%	0,294	82/1589 5%	0,244	128/1589 8%	0,236
Não específico	215/592 36%	0,275	252/592 43%	0,210	39/592 7%	0,253	86/592 15%	0,262

Fig.2

Há, possivelmente, uma influência do fator especificidade do referente na escolha da forma de retomada referencial, mas não com a mesma força da distinção [$\pm$ animado]. Quando o referente é específico, há uma nítida preferência pela anáfora para retomá-lo ao longo do texto: 64% contra 23% para os casos de repetição e 8% de elipses. Por outro lado, se for não específico, as retomadas referenciais se distribuem entre a anáfora (43%), a repetição (36%) e a elipse (15%). A única generalização a que se pode chegar com segurança a partir dos dados relativos ao fator especificidade do referente contidos na Fig.3 é que há uma tendência ao uso da anáfora como forma de retomada de referentes específicos.

Isso reforça as afirmações feitas em *A retomada de referentes animados e inanimados* sobre a tendência à especialização de funções entre a repetição e anáfora na oralidade, mas reitera também que o critério nuclear para essa especialização é a distinção entre referentes animados e inanimados. A adição da característica semântica de especificidade aumenta a probabilidade de um referente ser retomado pela anáfora.

O paralelismo nas cadeias referenciais

Halliday & Hasan (1976, p.15), quando formulam o conceito de coesão textual, afirmam que as retomadas referenciais levam à formação de cadeias coesivas no texto. Nesse caso, um referente posto em foco pode ser sucessivamente retomado, de forma que a interpretação de uma anáfora ou elipse, por exemplo, dependa do reconhecimento de sucessivas ligações a outras elipses ou anáforas, culminando, necessariamente na sua associação ao sintagma nominal que contém as informações lexicais necessárias à interpretação.

Qualquer texto tomado como objeto de observação apresenta cadeias de forma sistemática. Mas, ao investigar os fatores que podem atuar na escolha da forma de expressão da correferencialidade ao longo do texto, é relevante observar se a forma de uma retomada referencial na cadeia pode afetar as escolhas subsequentes.

O princípio geral de que o paralelismo formal atua de forma significativa nos processos de variação lingüística está bem estabelecido. O princípio geral de que “marcas conduzem a marcas e zeros conduzem a zeros” foi formulado

inicialmente por Poplack em estudos sobre a concordância nominal no espanhol de Porto Rico¹⁵. Esse princípio foi a seguir utilizado em vários estudos sobre o português, entre os quais deve-se, sem dúvida, destacar a tese de Scherre (1988, p.378-415) sobre a concordância nominal, que além de mostrar que esse fator atua de forma marcante na concordância nominal no português, aponta sua relevância em outros processos de variação. O princípio do paralelismo formal é proposto como um processo universal de processamento linguístico nos textos de SCHERRE & NARO (1991, 1992).

A afirmação de que o princípio do paralelismo formal seja relevante para um determinado processo de variação equivale a dizer que havendo cadeias em que um determinado processo linguístico variável possa atuar, a escolha feita em um determinado ponto dessa cadeia possa ter consequências nas escolhas seguintes.

Para o estudo da atuação do princípio do paralelismo formal na escolha das formas de retomada referencial na oralidade, cada ocorrência encontrada no *corpus* foi observada na sua relação com as formas que a precederam. Foram observadas especialmente duas características contextuais que podem ser relevantes para a escolha da forma de retomada referencial: sua distância em relação à presença do sintagma nominal que introduz o referente no discurso, medida pelo número de vezes em que o referente é retomado a partir de sua introdução como tópico no texto; sua semelhança ou diferença em relação à forma que a precede no texto. A Fig. 3 mostra os resultados do levantamento feito.

Influência do paralelismo formal nas retomadas referenciais (VARSUL)

	Posição na cadeia referencial	Repetição	Anáfora	Expressão descritiva	Zero
Ponto de Partida	1	217/513 42%	226/513 44%	33/513 6%	37/513 7%
Retomadas com paralelismo	2	81/223 36%	121/223 54%	6/233 3%	15/233 7%
	3 ou mais	176/939 19%	691/939 74%	25/939 3%	47/939 5%
Retomadas sem paralelismo	2	24/98 24%	39/98 40%	5/98 5%	30/98 31%
	3 ou mais	86/408 21%	185/408 45%	52/408 13%	85/408 21%

Fig.3

Inicialmente são necessários alguns esclarecimentos sobre a forma como os dados foram organizados no quadro acima. Optou-se por incluir no quadro somente os percentuais e não os pesos relativos, porque a codificação escolhida dá um tratamento probabilístico conjunto aos contextos em que a atuação do paralelismo é possível (segunda retomada em diante) e contextos em que a possibilidade de paralelismo não se coloca (primeira retomada). Por outro lado, só se pode falar na atuação do paralelismo tendo-se como ponto de partida um índice

¹⁵ *apud* SCHERRE (1988, p. 381-2)

básico, que seja revelador da tendência de uso de cada forma nos contextos em que o paralelismo não esteja em questão. Assim, os índices da posição 1 estabelecem o ponto de partida para a comparação com as demais retomadas. Cada conjunto de dados foi analisado na sua relação com os índices dessa posição. Tomando-se como exemplo o início do texto 3 citado acima:

Porque um... *um cara*₁ que é, vamos supor, entende? *o cara*₁ que é... *ele*₁ é um (inint). Então, *ele*... *ele*₁ pega... *ele*₁ empreita a empreiteira pra fazer alguma coisa (inint).

tem-se a introdução do referente com o sintagma nominal *um cara* e a primeira retomada mediante sua repetição: *o cara*. A seguir, na segunda, terceira (desconsiderando-se a hesitação) e quarta retomadas, tem-se o uso de anáforas.

A repetição tem um índice de ocorrência de 42% na posição 1, contexto em que a atuação do paralelismo é impossível. Esse índice estabelece o ponto de referência para a leitura das ocorrências da repetição nas demais posições. Na posição 2, havendo uma retomada precedente com o uso da repetição (casos de paralelismo formal), o índice de utilização da mesma é mais baixo (36%), sendo ainda mais reduzido da terceira posição em diante. Se a retomada precedente não for de uma repetição (casos sem paralelismo formal), os índices de uso da repetição também são baixos: 24% na posição 2 e 21% da posição 3 em diante. Em outras palavras, a ocorrência de uma retomada referencial por meio da repetição não favorece a ocorrência de outras retomadas com o uso do mesmo recurso formal na sua seqüência imediata. Observe-se ainda que o uso da repetição a partir da posição 2, quando não for antecedida por outra forma de repetição (retomadas sem paralelismo formal) tem índices ainda mais baixos do que a posição 1, onde atuam outros fatores, e não o paralelismo. Esses dados permitem duas conclusões: em primeiro lugar, que o princípio do paralelismo formal não atua na escolha da repetição como forma de retomada referencial no texto oral; além disso, se houver uma retomada referencial anterior, a tendência é que não haja a utilização da repetição nas retomadas subsequêntes, ou seja, no português falado evita-se a constituição de cadeias coesivas com a utilização da repetição como forma de expressão das retomadas referenciais.

A coluna correspondente aos casos de anáfora, por sua vez, mostra um comportamento muito diferente. O índice tomado como ponto de partida é 44% na primeira posição, que corresponde ao uso dessa forma de retomada referencial em um contexto em que o paralelismo formal não está em questão. Na segunda posição, antecedido de outra anáfora (casos de paralelismo formal), há um aumento significativo do uso dessa mesma forma (54%), e, a partir da terceira posição, o índice é ainda maior: 74%. Nos casos de retomadas referenciais não antecedidas, na posição imediatamente anterior, de outro caso de anáfora (sem paralelismo formal), os índices de ocorrência ficam próximos da posição 1. Como se vê, nesse caso a atuação do princípio do paralelismo formal é nítida.

A Fig. 3 mostra também que o princípio do paralelismo formal não atua no uso de expressões descritivas e zeros: as ocorrências dessas formas de

retomadas referenciais precedidas de uma retomada com o mesmo tipo de realização formal mantêm os mesmos índices de sua ocorrência na posição 1. O quadro permite ainda uma observação curiosa em relação ao uso de expressões descritivas: essas têm um aumento no índice de uso quando há um distanciamento maior em relação ao sintagma nominal que põe em foco o referente retomado. As ocorrências de expressões descritivas são encontradas especialmente a partir da terceira retomada do referente, provavelmente como uma estratégia para reavivá-lo mediante nova forma de apresentação (nova descrição) para o interlocutor.

Outra observação curiosa que se pode fazer a partir da organização de dados da Fig. 3 diz respeito aos contextos em que os índices de retomada referencial com o uso de um zero ocorrem de forma significativa. Trata-se dos contextos em que a retomada imediatamente anterior foi feita com o uso de outra forma (ausência de paralelismo formal), e que se distanciam do sintagma nominal que coloca em foco o referente. Se a escolha de uma expressão descritiva pode ter sua motivação na reapresentação do referente, em sua recolocação em foco, o uso da elipse certamente tem motivação oposta: a pressuposição de que o foco sobre o referente não precisa ser reafirmado, de que sua omissão não causa problemas.

Esses dados fornecem indícios de que o princípio do paralelismo formal atua de forma seletiva na constituição de redes de retomadas referenciais. A ocorrência de uma retomada mediante o uso da anáfora leva a outras retomadas com essa mesma forma, o que não acontece com as demais formas. Entre as quatro formas de retomada referencial incluídas no levantamento, a polarização do uso se dá realmente entre duas: a anáfora e a repetição. Essas duas formas comportam-se diferentemente em relação ao princípio do paralelismo formal: nas retomadas pela anáfora, o paralelismo atua de forma clara, possibilitando a formulação de um princípio geral: em uma cadeia referencial, a inserção de uma anáfora em um determinado ponto favorece sensivelmente o uso de outras anáforas na sua sequência. Esta questão merece investigações posteriores.

Hipóteses descartadas

O quarto grupo de fatores incluído como possível motivador entre a escolha das formas diferenciadas de expressão das retomadas referenciais foi a função sintática desempenhada pelo elemento que faz essa retomada. A razão desse levantamento está em vários estudos que apontam uma diferença entre o preenchimento da posição de sujeito e de objeto. Enquanto no português, há uma tendência a preencher sistematicamente a posição de sujeito, o mesmo não ocorre na posição de objeto, em que os zeros são mais frequentes. Assim, foi incluída no levantamento estatístico a diferenciação entre as retomadas referenciais em que o elemento responsável por essa retomada está na posição de sujeito ou de objeto, abrindo-se uma terceira classificação para as demais posições.

A hipótese inicial que motivou a inclusão desse fator na análise estatística, a indicação de uma tendência a um maior preenchimento da posição de sujeito que

de objeto, não se confirmou: não se observou uma concentração significativa de zeros na posição de objeto, talvez em consequência do cuidado na seleção das ocorrências de zeros incluídas no levantamento estatístico.

Na organização inicial do *corpus*, buscou-se um equilíbrio entre textos narrativos e argumentativos, para se ter uma diversificação na forma de articulação textual. O levantamento estatístico feito aponta para a irrelevância do tipo de texto na forma das retomadas referenciais: há, sem dúvida, uma diferenciação entre o uso da repetição e da anáfora nos dois tipos de texto, em consequência: 1) de uma maior concentração de referentes animados e específicos nas narrativas, favorecendo uma maior ocorrência de anáforas; 2) de um ligeiro aumento de referentes inanimados e genéricos nos textos argumentativos, o que resulta em um pequeno favorecimento do uso da repetição. As características semânticas do referente atuam como mediadoras entre o tipo de texto e a escolha da forma das retomadas referenciais. Essa não decorre diretamente do tipo de texto, mas da diferenciação entre os referentes que são tomados como foco de um e outro tipo de texto.

A inclusão da escolaridade como um possível fator no uso das formas de retomadas referenciais foi definida a partir do estudo de Possenti (1988), que, ao associar a escolha entre a anáfora e a repetição às imagens que o locutor faz do seu interlocutor, opera com uma atribuição de valores diferenciados às formas de retomada referencial. É óbvio que não se pode transpor mecanicamente as observações sobre a imagem do interlocutor feitas por Possenti para a escolha entre essas formas de realização das retomadas referenciais no texto oral não planejado, ainda mais levando-se em conta que as entrevistas analisadas têm sempre um interlocutor com nível de escolaridade superior ao do falante. Mas a partir do estudo de Possenti, pode-se pensar em uma escala de valorização das formas de retomada referencial, que a educação formal se encarregaria de reforçar. Essa escala teria, nos seus níveis mais elevados, as formas mais características da escrita: elipse e expressões descritivas. Num segundo nível, a anáfora e, por fim, a repetição, associada, segundo Possenti, à baixa capacidade de interpretação do discurso. A maior permanência do falante na escola poderia resultar em uma preferência pelas formas mais valorizadas de expressão da coesão. Se a escolarização formal tivesse uma atuação significativa nas seleções feitas pelo falante na linguagem oral, esperaríamos encontrar índices mais altos de uso das formas mais valorizadas entre falantes com maior grau de escolaridade e, paralelamente, índices mais altos da repetição, por exemplo, entre falantes com escolaridade primária.

Essa hipótese não se confirmou: os dados estatísticos não nos permitem perceber uma correspondência entre o uso das formas de retomadas referenciais e grau de escolaridade que seja consistente com as hipóteses apontadas acima. As formas de retomada referencial na oralidade são escolhidas pelo falante a partir de princípios próprios de organização do discurso oral. A maior familiaridade com a escrita que se verifica ao longo do processo de escolarização não afeta os

princípios mobilizados pelo falante para a escolha da forma das retomadas referenciais na fala.

CONCLUSÕES

Os trabalhos sobre a coesão consultados para a elaboração deste estudo têm como objeto o texto escrito, mesmo quando procuram formular princípios válidos também para a modalidade oral. Procurou-se neste estudo identificar os princípios atuantes na oralidade, mediante o controle estatístico de vários fatores que poderiam ser relevantes na escolha entre as formas alternativas de expressão das retomadas referenciais. Assim, foi possível demonstrar, primeiramente, que no texto oral as retomadas referenciais se fazem basicamente mediante o uso da anáfora e da repetição, sendo muito baixos os índices de uso de expressões descritivas e de zeros. A escolha entre a anáfora e a repetição leva em conta basicamente duas características semânticas do referente: a distinção entre referentes animados e inanimados e a sua especificidade, tendo a primeira característica um peso bem maior do que a segunda. Na constituição de cadeias coesivas, o princípio do paralelismo formal tem uma atuação restrita: só é significativo no uso das anáforas, mas não nas demais formas de realização das retomadas referenciais.

Cumpre destacar, por fim, a provisoriidade destas conclusões, que poderão ser reformuladas a partir do estudo de *corpora* mais extensos ou do refinamento dos critérios ora adotados.

BIBLIOGRAFIA

- BENVENISTE, Émile. *Problemas de lingüística geral*. São Paulo, Nacional/Ed. da Universidade de São Paulo, 1976.
- BORGES NETO, José. Os possessivos como indicadores de referência e atribuição. In: D.E.L.T.A. São Paulo, EDUC, 2(1), p. 145-9, 1996.
- BROWN, Gillian; YULE, George. *Discourse analysis*. Cambridge University Press, 1983.
- CHAROLLES, Michel. Introdução aos problemas da coerência dos textos (abordagem teórica e estudo das práticas pedagógicas). In: GALVES, C; ORLANDI, E.; OTONI, P. (orgs) . *O texto: escrita e leitura*. Campinas, Pontes, 1988.
- DE BEAUGRANDE, Robert-Alain; DRESSLER, Wolfgang Ulrich. *Introduction to text linguistics*. Londres, Nova York, Longmann, 1981.
- FÁVERO, Leonor Lopes. *Coesão e coerência textuais*. São Paulo, Ática, 1991.
- ____; KOCH, Ingedore G. Villaça. *Lingüística textual: introdução*. São Paulo, Cortez, 1994.
- HALLIDAY, M.A.K; HASAN, R. *Cohesion in english*. Londres, Longman, 1976.

- KNIES, Clarice Bohn; COSTA, Iara Bemquerer. *Banco de dados lingüísticos VARSUL: Manual do usuário*. Curitiba, Florianópolis, Porto Alegre, UFPR, UFSC, UFRGS, PUC/RS, 1996.
- KOCH, Ingedore G. Villaça. *A coesão textual*. São Paulo, Contexto, 1990.
- _____. Principais mecanismos da coesão textual em português. *Cadernos de estudos lingüísticos*. Campinas, UNICAMP, v. 15, p. 73-80, 1988.
- LEMO, Maria Tereza G. de. Sobre o que faz texto: uma leitura de *Cohesion in English*. *D.E.L.T.A.*. São Paulo, EDUC, v. 8, n. 1, p. 21-42, 1992.
- OLIVEIRA e SILVA, Gisele Machline de. Um caso de definitude. *Organon*, Porto Alegre, v. 18, p. 90-108, 1991.
- _____. Estereótipos da forma *seu* na língua oral. In: OLIVEIRA E SILVA, G. M. ; SCHERRE, M. Marta (orgs) . *Padrões sociolingüísticos*. Rio de Janeiro, Tempo Brasileiro/ UFRJ, 1996.
- PÊCHEUX, Michel. *Analyse automatique du discours*. Paris, Dunod, 1969.
- PINTZUK, Susan. VARBRUL Programs. 1988. Inédito.
- POSSENTI, Sírio. *Discurso, estilo e subjetividade*. São Paulo, Martins Fontes, 1988.
- SCHERRE, Maria Marta Pereira. *Reanálise da concordância nominal em português*. Universidade Federal do Rio de Janeiro, 1988. Tese de Doutorado.
- _____; NARO, Anthony J. Marking in discourse: Birds of a feather. In: *Language Variation and Change*, New York, v. 3, n. 1, p. 23-32, 1991
- _____. The serial effect on internal and external variables. In: *Language Variation and Change*, New York, v. 4, n. 1, p. 1-13, 1992.
- VAN DIJK, Teun. *Estructuras y funciones del discurso*. México, Siglo Veintiuno, 1980.