

A ESTRUCTURA DAS SENTENÇAS COM
PRONOME INTERROGATIVO NO
PORTUGUÊS BRASILEIRO ATUAL

Nilmara Soares Sikansi
1994

Este exemplar é a redação final da tese
defendida por Nilmara Soares

Sikansi

e aprovada pela Comissão Julgadora em

04/11/94

PROFA. DRA. MARY AIZAWA KATO

NILMARA SOARES SIKANSI

A ESTRUTURA DAS SENTENÇAS COM PRONOME INTERROGATIVO

Dissertação apresentada à Coordenação de pós-Graduação do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas como requisito parcial para a obtenção do título de Mestre em Linguística.

Orientador:

Prof^ª. Dr^ª. Mary Aizawa Kato

Campinas
Outubro - 1994

Ao meu marido,

PAULO

Às minhas filhas,

DÉBORAH E GABRIELA

A G R A D E C I M E N T O S

À Mary Kato, pela orientação segura, precisa e amiga;

aos colegas do programa de pós-Graduação do IEL/UNICAMP, pelas discussões enriquecedoras;

às amigas Helena Britto, Marina Augusto, Célia Castilho e Wilma Rêche, pelo companherismo;

às Prof^{as}. Charlotte Galvês e Maria Luiza Braga, pelas valiosas sugestões por ocasião da minha qualificação;

ao corpo docente da pós-Graduação do IEL/UNICAMP, pela inestimável contribuição para a minha formação como pesquisadora;

aos funcionários do IEL/UNICAMP, pela colaboração;

à FAPESP, pelo auxílio financeiro através de Bolsa de Mestrado e Iniciação Científica;

aos meus familiares, pelo estímulo constante e ajuda preciosa;

ao meu marido, Paulo, pelo grande apoio que me deu;

a todos os meus amigos, por me ouvirem pacientemente;

a todas as demais pessoas ou entidades que contribuíram, direta ou indiretamente, para a realização deste trabalho,

o meu muito obrigada.

Í N D I C E

I.	INTRODUÇÃO.....	01
II.	APRESENTAÇÃO DO PROBLEMA.....	03
III.	RESENHAS.....	06
	Notas.....	29
IV.	A DEFINIÇÃO DOS FATORES.....	31
	1. Variável dependente.....	31
	2. Tipo de sujeito.....	32
	3. Tipo de verbo.....	33
	4. Tipo de voz.....	35
	5. Tipo de construção verbal.....	35
	6. Expletivo <u>é que</u>	36
	7. Tipo de elemento Q-.....	37
	8. Estatuto da oração.....	38
	9. Tipo de corpus.....	38
	9.1- Dados do Projeto NURC/SP.....	38
	9.2- Dados de Escrita.....	39
	9.3- Dados de Fala.....	40
	10. Dimensão do sujeito.....	41
V.	ANÁLISE DOS FATORES.....	44
	1. Dados do Projeto NURC/SP.....	45
	1.1 - Tipo de sujeito.....	46
	1.2 - Tipo de verbo.....	47
	1.3 - Tipo de corpus.....	50
	1.4 - Tipo de elemento Q-.....	52
	1.5 - Os demais grupos de fatores.....	55
	2. Dados de Língua Escrita.....	58
	2.1 - Tipo de verbo.....	59
	2.2 - Tipo de sujeito.....	60
	2.3 - Tipo de elemento Q-.....	61
	2.4 - Estatuto da oração.....	64
	2.5 - Os demais grupos de fatores.....	66
	3. Dados de Língua Falada.....	69
	3.1 - Tipo de verbo.....	69
	3.2 - Dimensão do sujeito.....	71
	3.3 - Estatuto da oração.....	72
	3.4 - Tipo de sujeito.....	73
	3.5 - Os demais grupos de fatores.....	75
VI.	CONCLUSÃO.....	79
	REFERÊNCIAS BIBLIOGRÁFICAS.....	87

ÍNDICE DE TABELAS

1. DADOS DO PROJETO NURC/SP:

Tabela 01: Frequência de ordem S-V e V-S em relação ao tipo de sujeito.....	46
Tabela 02: Frequência e Peso relativo de V-S segundo o tipo de sujeito.....	47
Tabela 03: Frequência de V-S e S-V conforme o tipo de verbo...	47
Tabela 04: Frequência e Peso relativo de V-S segundo o tipo de verbo.....	48
Tabela 05: Frequência de V-S a partir do cruzamento de tipo de verbo com tipo de sujeito.....	49
Tabela 06: Frequência e peso relativo de V-S segundo o tipo de corpus.....	50
Tabela 07: Frequência de V-S a partir do cruzamento de tipo de sujeito com tipo de corpus.....	51
Tabela 08: Frequência de V-S a partir do cruzamento de tipo de verbo com tipo de corpus.....	51
Tabela 09: Frequência de V-S e S-V conforme o tipo de elemento Q-.....	52
Tabela 10: Frequência e Peso relativo de V-S segundo o tipo de elemento Q-.....	53
Tabela 11: Frequência de V-S a partir do cruzamento de tipo de elemento Q- com tipo de sujeito.....	54
Tabela 12: Frequência de V-S a partir do cruzamento de tipo de verbo com tipo de elemento Q-.....	54
Tabela 13: Frequência de V-S e S-V segundo o tipo de voz verbal.....	55
Tabela 14: Frequência de V-S e S-V segundo o tipo de construção verbal.....	55
Tabela 15: Frequência de V-S e S-V conforme a presença ou ausência do expletivo <u>é que</u>	56
Tabela 16: Frequência de V-S e S-V segundo o estatuto da oração.....	57
Tabela 17: Frequência de V-S e S-V segundo a dimensão do sujeito.....	57

Tabela 18: Frequência de V-S a partir do cruzamento de tipo sujeito e dimensão do sujeito.....	58
Tabela 44: Frequência de V-S a partir do cruzamento de estatuto da oração com tipo de sujeito.....	80

2. DADOS DE LÍNGUA ESCRITA

Tabela 19: Frequência de V-S e S-V segundo o tipo de verbo....	59
Tabela 20: Frequência e peso relativo de V-S segundo o tipo de verbo.....	60
Tabela 21: Frequência e peso relativo de V-S segundo o tipo de sujeito.....	60
Tabela 22: Frequência de V-S a partir do cruzamento de tipo de sujeito com tipo de verbo.....	61
Tabela 23: Frequência e peso relativo de V-S conforme o tipo de elemento Q-.....	62
Tabela 24: Frequência de V-S a partir do cruzamento de tipo de verbo com tipo de elemento Q-.....	62
Tabela 25: Frequência de V-S a partir do cruzamento de tipo de sujeito com tipo de elemento Q-.....	63
Tabela 26: Frequência e peso relativo de V-S segundo o estatuto da oração.....	64
Tabela 27: Frequência de V-S a partir do cruzamento de tipo de verbo com estatuto da oração.....	65
Tabela 28: Frequência de V-S a partir do cruzamento de tipo de sujeito com estatuto da oração.....	65
Tabela 29: Frequência de V-S a partir do cruzamento de estatuto da oração com tipo de elemento Q-.....	66
Tabela 30: Frequência de V-S e S-V segundo o tipo de corpus....	67
Tabela 31: Frequência de V-S e S-V segundo o uso do expletivo <u>é que</u>	68
Tabela 32: Frequência de V-S e S-V segundo a dimensão do sujeito.....	69

3. DADOS DE LÍNGUA FALADA

Tabela 33: Frequência de V-S e S-V conforme o tipo de verbo...	70
Tabela 34: Frequência e peso relativo de V-S segundo o tipo de verbo.....	70
Tabela 35: Frequência e peso relativo de V-S conforme a dimensão do sujeito.....	71
Tabela 36: Frequência de V-S a partir do cruzamento de dimensão do sujeito com tipo de verbo.....	72
Tabela 37: Frequência e peso relativo de V-S segundo o estatuto da oração.....	72
Tabela 38: Frequência de V-S a partir do cruzamento de estatuto da oração com tipo de verbo.....	73
Tabela 39: Frequência e peso relativo de V-S conforme o tipo de sujeito.....	74
Tabela 40: Frequência de V-S a partir do cruzamento de tipo de sujeito com dimensão do sujeito.....	74
Tabela 41: Frequência de V-S e S-V conforme o uso do expletivo <u>é que</u>	75
Tabela 42: Frequência de V-S e S-V conforme o tipo de elemento Q-.....	76
Tabela 43: Frequência de V-S e S-V conforme o tipo de corpus..	77

I. INTRODUÇÃO

Muitos estudos teóricos recentes têm demonstrado um interesse muito grande em explicar as possibilidades de ordenação dos constituintes na sentença. Questões tais como qual seria a ordem básica nas línguas têm originado trabalhos usando os mais diversas perspectivas teóricas. Tendo esse quadro em vista, as construções interrogativas com elemento Q- é um tópico que vem sendo muito analisado, pois se trata de uma construção onde há variação, não só trans-lingüística como intra-lingüística.

Torrego (1984 - apud Âmbar 1988) postula que, para o espanhol, a inversão Verbo-Sujeito se dá obrigatoriamente quando o elemento Q- é um argumento (interno ou externo) do verbo. Já Âmbar (op. cit.) diz que, para o português europeu, a inversão é obrigatória nas interrogativas diretas desde que o elemento Q- não seja acompanhado de um N(ome) foneticamente realizado. Com relação ao português do Brasil, Kato (1987) analisa algumas sentenças interrogativas com elemento Q- e verifica que parece não haver nenhum contexto categórico que desencadeie a inversão verbo-sujeito, podendo esta regra, portanto ser vista como estilística.

O objetivo central desse trabalho é verificar qual a estrutura que ocorre nas sentenças com pronome interrogativo e observar se há algum condicionamento - interno ou externo - que esteja favorecendo o uso dessa ordenação. Dessa forma, o corpus utilizado foi obtido a partir de diversos contextos de uso - entrevista na televisão, sala de aula de adultos e crianças, redações do vestibular, peças teatrais, romances, entre outros - com o intuito de constatar se existe ou não algum tipo de condicionamento extra-lingüístico nessas sentenças. Além disso, procurarei apontar os possíveis fatores internos - tipo de verbo, tipo de sujeito, estatuto do elemento Q-, entre outros - que estejam favorecendo a aplicação da inversão Verbo-Sujeito (VS) no Português do Brasil (PB) atual.

A metodologia utilizada nessa pesquisa é a desenvolvida por Tarallo & Kato (1989) - Harmonia Trans-sistêmica - que procura unir a perspectiva paramétrica do modelo gerativo e a metodologia quantitativa do modelo variacionista, sendo que a teoria gerativista atual busca resgatar a variação inter-lingüística com o objetivo de estabelecer os princípios e os parâmetros que regem as línguas. Por outro lado, o modelo variacionista se preocupa com a variação intra-lingüística e, a partir da análise estatística de um corpus, levanta os fatores que mais favorecem ou inibem uma variação ou mudança.

O presente trabalho ficou da seguinte forma organizado: no próximo capítulo farei uma breve apresentação do fenômeno, inclusive mencionando a forma pela qual esse assunto é abordado em algumas gramáticas tradicionais do português.

Enumerarei, também, as questões que norteiam esse trabalho.

Em seguida, no capítulo três, apresentarei uma série de resenhas de textos teóricos que tratam, direta ou indiretamente, do problema da ordem dos constituintes nas sentenças interrogativas, tanto do português europeu quanto do português brasileiro.

O capítulo quatro consiste na discussão e justificação de cada um dos fatores selecionados para verificar os condicionamentos lingüísticos e extra-lingüísticos que estão atuando nesse fenômeno.

No capítulo seguinte, eu apresentarei os resultados estatísticos obtidos para cada corpus em separado. Procuro, então, interpretá-los com o intuito de compor um quadro das variáveis que estão atuando na estruturação dessas construções interrogativas.

O último capítulo reúne os aspectos convergentes e divergentes das análises descritas no capítulo cinco na tentativa de definir precisamente o conjunto de fatores que influenciam a ordem relativa entre sujeito e verbo na interrogativas com elemento Q-.

II. APRESENTAÇÃO DO PROBLEMA

Uma análise de três gramáticas da língua portuguesa - (1) Gramática Secundária da Língua Portuguesa de M. Said Ali, (2) Sintaxe Clássica Portuguesa de Cláudio Brandão e (3) Gramática do Português Contemporâneo de Celso Cunha - mostrou-me que nenhuma trata de maneira clara e sistemática o problema da ordem nas construções interrogativas com pronome. Na maioria das vezes, o leitor tem que intuir a ordem desse tipo de estrutura a partir dos exemplos que são dados. Estes são, em grande parte, de autores consagrados, porém bem antigos. Para ilustrar melhor o que eu estou falando, transcreverei abaixo a passagem onde Celso Cunha (op.cit.) aborda esse tipo de construção:

"Pronomes Interrogativos

Chamam-se INTERROGATIVOS os PRONOMES INDEFINIDOS que, quem, qual e quanto, empregados para formular uma pergunta direta ou indireta:

Que trabalho estão fazendo?
Quem disse tal coisa?
Qual dos livros preferes?
Quantos passageiros desembarcaram?
Diga-me que trabalho estão fazendo.
Ignoramos quem disse tal coisa.
Não sei qual dos livros preferes.
Pergunte quantos passageiros desembarcaram.

...

Valor e emprego dos interrogativos

QUE

1. O interrogativo que pode ser:

a) pronome substantivo, quando significa "que coisa":

Que teria havido naquela tarde? (Osman Lins)
...não sei que diria, se tivéssemos de ir conversando.
(Machado de Assis)

b) pronome adjetivo, quando significa "que espécie de", e neste caso refere-se a pessoas ou a coisas:

Que mal me havia de fazer? (Miguel Torga)
Não sei que vento mau turvou de todo o lago.
(Alphonsus de Guimaraens)

2. Em lugar de que pronome substantivo, usa-se também o que, ou, com reforço, que é que e que é o que:

O mundo? O que é o mundo, ó meu amor?
(Floribela Espanca)
Que é que o senhor está fazendo aí no Terraço a esta
hora da madrugada? (Carlos Drummond de Andrade)

QUEM

1. O interrogativo quem é pronome substantivo e refere-se apenas a pessoas ou a algo personificado:

Quem não canta? Quem? Quem não a canta e sente?
(Jorge de Lima)
Ninguém sabe quem sou... (Antero de Quental)

2. Em orações com o verbo ser, pode servir de predicativo a um sujeito no plural:

Quem sois vós, meus irmãos e meus algozes?
(Antero de Quental)

QUAL

1. O interrogativo qual tem valor seletivo e pode referir-se tanto a pessoas como a coisas. Usa-se geralmente como pronome adjetivo, mas nem sempre com o substantivo contíguo. Nas perguntas feitas com o verbo ser, costuma-se empregar o verbo depois de qual:

Qual foi o entendimento que não chegamos a ter?
(= Qual entendimento foi o que não chegamos a ter?)
(Álvaro de Campos)

2. A idéia seletiva pode ser reforçada pelo emprego da expressão qual dos (ou das) anteposta ao substantivo ou ao pronome no plural:

Qual dos senhores é pai dum menino que está de cócoras
no jardim há mais de meia hora? (Aníbal M. Machado)
Qual de nós ficará codilhado? (Oliveira Martins)

QUANTO

O interrogativo quanto é um quantitativo indefinido. Refere-se a pessoas e a coisas e usa-se quer como pronome substantivo, quer como pronome adjetivo:

Quanto é que o senhor oferece? (Rubem Braga)
Quantos livros já publicou? (Carlos Drummond de Andrade)

”(CUNHA, C. pag. 247,248)

Como pode ser visto no trecho transcrito acima, o autor nada fala com relação a ordem dos constituintes nesse tipo de sentença - exceto brevemente quando aborda o pronome interrogativo qual. Além disso, na maioria dos exemplos, o sujeito não é preenchido lexicalmente, o que não coloca em questão sua posição em relação ao verbo. Resumindo, pode-se dizer que as gramáticas pouco ou nada esclarecem sobre a estruturação das sentenças com elemento Q-.

A questão que embasou essa pesquisa surgiu, basicamente, a partir de um projeto de Iniciação Científica (Processo 90/2556-2 - FAPESP) que procurou verificar a ordem das sentenças com pronome interrogativo no português do Brasil atual. Foram coletados exemplos no corpus do projeto NURC/SP (Norma Urbana Culta), em entrevista na televisão e em alguns peças teatrais e romances da atualidade. Uma primeira análise dessas estruturas mostrou que as gramáticas da língua portuguesa não estavam refletindo a ordem que estava sendo usada pelos falantes. Posteriormente, ampliei o corpus e, procurando diversificá-lo ainda, coletei dados em sala de aula e redação do vestibular, entre outros contextos.

A questão básica que norteou esse trabalho é a seguinte:

Quais os fatores, externos e internos, que mais influenciam a ordem relativa entre sujeito e verbo nas interrogativas com elemento Q-?

Além disso, procurarei ainda verificar se fatos e conclusões de outros trabalhos sobre o assunto se confirmam ou não, como por exemplo:

- a) O expletivo "é que", ou sua versão reduzida "que", realmente favorecem a não inversão sujeito-verbo nessas construções?
- b) Qual é o comportamento das sentenças que contém um elemento Q- seguido por um nome (Q- NP)? Há diferenças entre essas construções e aquelas em que o elemento Q- não é seguido por um nome?

III. R E S E N H A S

Esse capítulo será dedicado à apresentação de alguns textos que abordam a questão da inversão Verbo-Sujeito no português europeu (PE) e brasileiro (PB) na atualidade. Primeiramente tratarei do PE, mostrando a interpretação dada por Âmbar, em sua tese de doutorado, para esse fenômeno. Uma outra opção de explicação é dada por Raposo, que será também apresentada em linhas gerais, no final do capítulo. Para o PB mostrarei algumas possibilidades de explicação para o fenômeno de inversão V-S nas interrogativas com elemento *qu*. Dado que, aparentemente, essa inversão não tem um caráter obrigatório no PB, em oposição ao PE, muitos dos textos que aqui serão apresentados atribuem a essa inversão um caráter estilístico. Dessa forma, antes de passar para os textos propriamente ditos, farei uma rápida apresentação do que vários autores falaram sobre variação estilística.

Aparentemente o conceito de variação estilística mais difundido na Linguística é o apresentado por Lyons (1987) que coloca que o sistema linguístico frequentemente fornece aos seus usuários meios alternativos de dizerem a mesma coisa. Entretanto, duas palavras ou sintagmas podem ser equivalentes do ponto de vista descritivo e, no entanto, diferir em termos de significação social e expressiva. Tais expressões sinônimas de maneira incompleta podem ser denominadas de variantes estilísticas: mais precisamente, de variantes não equivalentes estilisticamente.

Kato (1992), em seu artigo sobre "Variação sintática e Estilo", levanta uma questão que tem criado polêmica tanto na área dos estudos sobre variação - na linha laboviana - quanto na sintaxe formal chomskyana que é a existência ou não de variantes de ordem dos constituintes que possam ser efetivamente consideradas estilísticas.

Kato (1992) apresenta a noção de estilo presente em autores da Sociolinguística, contrapondo-as com as da linha gerativista. Sendo assim, temos que, "para Labov (1972), o próprio conceito de variante está ligado a estilo, pois as variantes são necessariamente idênticas em referência e valor verdade, mas são diferentes em significado social e/ou estilístico.

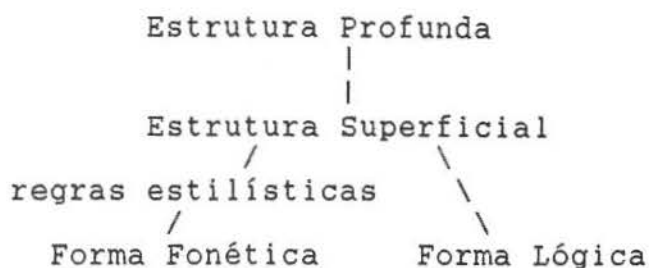
Por outro lado, Cedergren (1983) coloca que "é essencialmente por meio da variação que se manifestam os parâmetros de diferenciação social, os processos dinâmicos de variação estilística e a interação de fatores do sistema linguístico".

Já para Lavandera (1978), "se cada construção sintática tem seu próprio significado, como é possível que haja variação, se por variação entendemos duas (ou mais) maneiras de dizer a mesma "coisa"?"

Bentivoglio (1987), por sua vez, diz ser difícil encontrar um fenômeno real de variação sintática, dado que, ao contrário da fonologia, encontrar contextos iguais em sintaxe é muito mais difícil.

Para Chomsky e Lasnik (1977), uma regra é estilística se ela acontece entre a Estrutura-S (ES) e a Forma Fonética (FF) - veja Figura 1 abaixo. Ou seja, uma regra é estilística se sua aplicação não afetar a interpretação em Forma Lógica, e for condicionada por fatores fonéticos.

Figura 1:



Por outro lado, e seguindo o mesmo raciocínio, uma regra é sintática se ela ocorre entre a Estrutura Profunda e a Estrutura Superficial ou entre esta última e a Forma Lógica."

Após essa breve apresentação do ponto de vista de vários autores sobre o problema da variação estilística, comecemos, agora, com as análises propostas para o português europeu, no qual, em certos contextos o falante não tem escolha.

Âmbar (1988), analisando as inversões Sujeito-Verbo (ISV) numa variante do Português Europeu, coloca como sendo os seguintes os contextos de ocorrência dessa inversão nas estruturas interrogativas:

(a) a inversão em interrogativas Q- diretas é obrigatória sempre que o constituinte interrogativo, em posição inicial de sentença, não integra um N foneticamente realizado; porém, quando isto ocorre a inversão se torna facultativa e deve recair sobre o sujeito ou sobre o constituinte Q- um acento de intensidade, conforme informa a autora:

01. (6) (O) Que ofereceu o Pedro à Joana?¹
*(O) Que o Pedro ofereceu à Joana?

02. (7) Quem encontrou o João no cinema?
*Quem o João encontrou no cinema?

03. (21) ?Que disco o Pedro ofereceu à Joana?
Que disco O PEDRO ofereceu à Joana?
QUE DISCO o Pedro ofereceu à Joana?

(b) com exceção do que acontece com os elementos Q-interrogativos que e porque, a inversão em interrogativas encaixadas é facultativa, integrem ou não aqueles constituintes um N foneticamente realizado:

04. (26) Não sei que ofereceu o Pedro à Joana.
*Não sei que o Pedro ofereceu à Joana.
05. (27) Não sei o que ofereceu o Pedro à Joana.
Não sei o que o Pedro ofereceu à Joana.
06. (28) Não sei quem encontrou o João no cinema.
Não sei quem o João encontrou no cinema.
07. (31) Não sei porque saiu a Rita.
?*Não sei porque a Rita saiu.

(c) a ocorrência de é que nestas estruturas permite em todos os casos a ordem SV; da coexistência de é que com inversão resultam estruturas marcadas em termos de aceitabilidade:

08. (14) (O) Que é que o Pedro ofereceu à Joana?
Quem é que o João encontrou no cinema?
09. (15) ?(O) Que é que ofereceu o Pedro à Joana?
?Quem é que encontrou o João no cinema?

Âmbar observa ainda que, em estruturas de inversão, o sujeito ocorre em posição pós-verbal ou em posição final de frase. Tratando-se de um pronome, a posição final é marginal. Na presença de um auxiliar e de uma inversão, o sujeito pode ocorrer em posição pós-verbo auxiliar, pós-verbo principal e em posição final. Há, contudo, segundo ela, um decrescimento de aceitabilidade à medida que se passa da primeira à última dessas posições.

Depois de enumerar os contextos que apresentam ISV, a autora passa a propor uma explicação para essas estruturas. Primeiramente ela postula que, em qualquer língua, a inversão Sujeito-Verbo é provocada por um dado mecanismo. Nesse sentido, a chamada "inversão livre" não existiria nas línguas.

Tendo isso em vista, Âmbar procura explicar o contraste entre as sentenças (23a) e (23b) abaixo - mais precisamente a ausência vs presença de um N foneticamente realizado no constituinte Q- interrogativo movido para Espec -, propondo que

a ausência desse N corresponde efetivamente à presença de uma categoria vazia. Os constituintes Q- interrogativos teriam, portanto, a representação estrutural de (24):

10. (23) a. *Que o Pedro ofereceu à Joana?
 b. ?Que disco o Pedro ofereceu à Joana?²
11. (24) [QUE [e]]_{+Q-}

onde QUE seria um determinante interrogativo e [e] uma instância qualquer de N'.

Se [e] é uma categoria vazia e se há um princípio da gramática que estipula que uma categoria vazia deve ser apropriadamente regida (PCV)³, então [e] deve ser apropriadamente regida. Nas interrogativas diretas, porém, [e] não tem como satisfazer este princípio - não existe elemento regente capaz de desempenhar a tarefa, daí a agramaticalidade de (23a). Porém, se o verbo for deslocado para a posição de COMP, através da aplicação de mova-a, ele poderá reger a categoria vazia na posição de Espec.

Com esta análise, Âmbar propõe explicar o porquê do caráter obrigatório deste tipo de inversão: o verbo sobe, ou seja, há ISV obrigatória, porque existe, em Espec de Comp, uma categoria vazia que precisa ser regida. Além disso, para a autora, a hipótese de uma categoria vazia no constituinte Q- interrogativo é válida para qualquer elemento Q- que não integre um N foneticamente realizado. Assim, pronomes interrogativos como quem, o que, quando, onde, como e porque recebem, do ponto de vista da existência de uma categoria vazia, um tratamento uniforme. É, no entanto, desejável, como bem coloca a autora, que alguma distinção seja feita entre estes últimos interrogativos e que, pois nas estruturas encaixadas verifica-se um comportamento diversificado. Se nas interrogativas diretas a inversão é sempre obrigatória, nas encaixadas isso só ocorre quando o constituinte em Espec de Comp é que ou porque. Âmbar explica essa diferença de comportamento considerando que a categoria vazia em questão tem traços [+humano], [+específico], [+tempo], [+lugar], [+modo], [+causal] que restringem o seu domínio de referência. Dessa forma, os constituintes interrogativos teriam, em estrutura-S, a forma dada à esquerda e, em FF, a forma à direita:

12. (29) [QUE [e]_{-r}][que]
 [QUE [e]_{+humano}][quem]
 [QUE [e]_{+específico}][o que]
 [QUE [e]_{+tempo}][quando]
 [QUE [e]_{+lugar}][onde]
 [QUE [e]_{+modo}][como]
 [POR [QUE [e]_{-r}][porque]

Assim, os constituintes Q- interrogativos quem, o que, quando, onde e como seriam marcados [+r], enquanto que que seria [-r]. Segundo a autora, intuitivamente, que tem, nas frases em que ocorre, escopo sobre um domínio de referência mais genérico que os outros constituintes. Quanto ao porque, ela afirma que há uma certa variação na aceitabilidade de frases encaixadas introduzidas por este interrogativo - alguns falantes aceitam marginalmente a ausência de inversão enquanto que outros a exigem. Portanto, ela atribui ao porque a estrutura dada acima. Âmbar ainda levanta a hipótese de que a oscilação de gramaticalidade pode se dever ao fato de haver, neste constituinte, um elemento regente (por) que faz da categoria vazia uma categoria regida, mas não suficientemente regida.

Nas construções encaixadas, o verbo da matriz rege, mas não rege apropriadamente $[el]_{+/-r}$, dessa forma o verbo da encaixada deve subir nas interrogativas introduzidas por que. Sem a subida do verbo, apenas uma parte das condições de legitimação da categoria vazia estará satisfeita - licenciamento formal. A identificação do conteúdo só será realizada mediante a subida do complexo V-Flex que, uma vez em Comp, pela estratégia de concordância Espec-Núcleo se qualifica como um antecedente para a categoria vazia.

Nos casos em que a categoria vazia é [+r] a subida do verbo não é obrigatória pois o seu conteúdo é identificado pelos traços [+r] que lhe são inerentes e o licenciamento formal é feito pelo verbo da sentença matriz.

Essa é a forma pela qual Âmbar procura explicar o caráter obrigatório da inversão nas diretas vs o seu caráter facultativo nas encaixadas e, também, a obrigatoriedade da subida do verbo no contexto de que interrogativo em encaixadas, bem como a não obrigatoriedade de inversão nos casos em que o constituinte Q- interrogativo integra um N foneticamente realizado - onde não há qualquer categoria vazia que tenha de ser legitimada.

Resta à autora ainda explicar a razão pela qual a estratégia é que surge como alternativa ao mecanismo de inversão. Para ela, no entanto, esta estratégia é apenas aparente. De acordo com sua análise, se é que pode desempenhar o mesmo papel que a subida do verbo, e se esta é exigida para a legitimação da categoria vazia, logo, tem-se de admitir que é que permite também a identificação dessa categoria.

Âmbar propõe que a estrutura resultante do movimento do elemento Q- interrogativo numa frase como (42) é a representada em (43):

13. (42) (O) que é que o Pedro ofereceu à Joana?

14. (43) [_{Comp} [QUE [e]_{+/-r}]_i [_{Comp} [_{Flex} pro Flex e- [_{Comp} v'_i [_{Comp} que
[_{Flex} o Pedro Flex oferece- v_i à Joana]]]]]]

onde a categoria vazia no especificador mais alto tem de ser regida. Para isso, Flex, amalgamada com ser sobe para Comp, de onde identifica a categoria vazia em Espec, da forma prevista para os restantes casos. A estrutura-S resultante será então como em (44) abaixo:

15. (44) [_{Comp} [QUE [e]_{+/-r}]_i [_{Comp} é_{jk} [_{Flex} pro v_{jk} v_k [_{Comp} v'_i [_{Comp}
que [_{Flex} o Pedro Flex oferece- v_i à Joana]]]]]]

Como foi descrito anteriormente, a co-ocorrência de é que com a inversão é marginalmente possível como o exemplo abaixo mostra:

16. (45) ?(O) que é que ofereceu o Pedro à Joana?

Dada a posição do sujeito em (45), temos que a ordem VS resulta de uma subida do verbo. Por imposições de preservação de estrutura, a posição de destino do verbo só pode ser a posição em Comp. Esta, porém, está preenchida pelo complementador que. Âmbar admite que o verbo pode ocupar um Comp preenchido pelo complementador que. A representação da estrutura-S de (45) será a de (47):

17. (47) [_{Comp} [QUE [e]_{+/-r}]_i [_{Comp} é_{jk} [_{Flex} pro v_{jk} v_k [_{Comp} v'_i [_{Comp}
que ofereceu_{lm} [_{Flex} o Pedro v_i v_m v_l à Joana]]]]]]

A estratégia é que não constitui assim uma exceção a ISV no contexto Q- interrogativo, segundo a autora. É, na verdade, uma outra manifestação do mecanismo de subida do verbo, obrigatória sempre que é exigida a identificação da categoria vazia postulada nos constituintes Q- interrogativos.

A explicação dada por Âmbar para o fenómeno de ISV nas interrogativas com constituinte Q-, a meu ver, apresenta, pelo menos, um problema básico - o Comp duplamente preenchido que ocorre nas estruturas com a partícula é que e com ISV.

Apesar dos argumentos apresentados pela autora a favor da hipótese de que o Comp poderia ter dois elementos ocupando uma mesma posição, isso me parece uma solução ad hoc que não tem fundamentação teórica no modelo por ela adotado.

Passando agora para o Português do Brasil (PB), mostrarei, a seguir, como o fenómeno da inversão sujeito-verbo tem sido tratado por alguns autores.

Pontes (1982) faz uma análise da ordem VS em português usando dados tanto da língua escrita quanto da oral. Seu objetivo é verificar a extensão desse fenômeno e dar a ele uma explicação com base na Análise do Discurso.

No início do seu artigo ela menciona que alguns autores, buscando precisar se a ordem básica do português era VSO ou SVO, utilizavam, basicamente, sentenças intransitivas para defender a primeira hipótese (VSO - Perlmutter 1976) e orações transitivas a favor da segunda (SVO - Berman 1974). Analisando diversas gramáticas do português, a autora enumera os vários contextos de ocorrência da ordem VS por elas apontados. As interrogativas com pronome, as parentéticas, as sentenças com verbos existenciais e com o verbo ser são alguns dos casos onde a inversão mais se verifica.

Após esse levantamento, Pontes passa a analisar alguns textos escritos - Galo das Trevas de Pedro Nava (1981), Discurso de Primavera e Algumas Sombras de Carlos Drummond de Andrade (1978) e revista Isto É (10/03/1982) - para ver se no uso da língua escrita se confirmam as regras encontradas nos gramáticos. Ela realmente encontra, nos textos estudados, praticamente todas as situações de inversão descritas. Porém, a meu ver, alguns exemplos por ela fornecidos me parecem duvidosos quanto ao fato de ali ocorrer a ordem VS. Coloco-os a seguir:

18. Quem seria a segunda esposa de Pedro I? (I.É. 20)

19. Onde estão? meus convivas e as flores d'antanho, onde estão?
(PN 37)

Em (18) me parece que o pronome interrogativo quem está substituindo o sujeito, portanto não haveria VS pois o SN pós-verbal seria o predicativo do sujeito e, em (19), o sujeito está numa posição deslocada, aparecendo anteposto ao elemento interrogativo onde.

O que se observa nos exemplos listados pela autora é que, em sua maioria, predominam os verbos existenciais e o próprio verbo ser, construções essas onde não me parece que houve inversão do sujeito.

Para analisar a linguagem coloquial, a autora examina 05 gravações (duração de 15 a 30 minutos) mais alguns dados avulsos que foram colhidos por ela. Abaixo, transcrevo uma passagem do seu texto onde ela apresenta sua conclusão sobre esses dados.

O que se observa a respeito da língua oral é que há uma incidência de certos verbos que ocorrem mais em posição inicial. Em todos os tipos de oração, tanto afirmativas, como

negativas, como exclamativas, etc., os verbos que aparecem são daqueles que Ephiphanyo Dias coloca na regra c): existir, ocorrer, aparecer e seus sinônimos. Contreras (1976) diz que verbos que denotam começo, surgimento, existência, etc., parecem ter a função apresentacional e Givón (1979) chama-os de "existenciais-apresentativos": be, exist, stand-sit-lie down, live, appear, often remain, be left and sometimes verbs of entrance into the scene such as come or enter (p. 28). Segundo Givón, esta tendência para estes verbos aparecerem em posição inicial é universal nas línguas SVO. No Português oral coloquial, os casos de VS mais frequentes são, de longe, aqueles em que um verbo deste tipo aparece: vir, entrar, chegar, passar, estar, existir, aparecer, sumir, sobrar, faltar, cair, são verbos que ocorrem a todo momento, em todo tipo de frase. Pode-se dizer que, do ponto de vista da frequência de ocorrência, os casos de VS em Português coloquial mais significativos se resumem àqueles em que um destes verbos ocorre.

Com relação às interrogativas com pronomes, a autora verifica a predominância de VS, mesmo quando aparece é que. Nota, ainda, que em todos os exemplos o verbo é intransitivo.

20. Cumé que chama isso, é um...microfone.

21. Onde fica o correio?

22. Onde está aquela gramática?

23. Por que naquele lugar não nasce mato?

Após esse levantamento, Pontes passa a analisar a frequência de VS tanto na língua escrita quanto na oral. Para isso, ela examina as primeiras 100 estruturas em que ocorre algum verbo na revista Isto É. Ela encontra apenas 07 casos de VS. Repetindo o processo com as 100 sentenças seguintes, obteve-se 08 ocorrências o que mostra que o primeiro resultado não foi aleatório. Para ter uma idéia mais completa do quadro, Pontes observa que encontrou uma alta porcentagem de verbos sem sujeito expresse (39 casos nas primeiras 100 estruturas). A autora cita Givón (1981) que considera como sendo este o caso mais alto numa escala de estratégias para manter a continuidade do tópico no discurso. Considerando também as demais estratégias, ela conclui que, no texto analisado, predominam os casos de continuidade (93 ocorrências). Sobram 07 exemplos que, segundo Givón, são estratégias de descontinuidade: construções de

ordem VS e existenciais.

Em virtude da alta incidência de orações sem sujeito explícito, por causa da anáfora por elipse, não se tem uma idéia, por essas estatísticas, da relação existente entre ordem VS e SV, uma vez que quando o sujeito está elíptico, não se pode saber se ele viria antes ou depois de V, como bem coloca Pontes. Para ter idéia desta correlação, a autora contou 100 estruturas em que o verbo aparecia com sujeito expresso e encontrou o seguinte resultado: 85 sujeitos antepostos para 15 pospostos (I.E. 10/03/82, p. 3-9). Confirmou-se, assim, a superioridade da anteposição do sujeito sobre a posposição e o caráter não-marcado da primeira em relação à segunda.

Para comparar a língua escrita com a oral, examinou-se, primeiramente, 100 frases em que ocorria algum verbo, num diálogo informal entre duas pessoas jovens, de Belo Horizonte, de nível universitário e encontrou-se apenas dois casos claros de VS:

24. Não, mas tá faltando um, tem que jogar.

25. E o Cristo fui eu.

A porcentagem de VS na língua oral é bastante reduzida, portanto. Isso foi confirmado por um estudo de outro diálogo - nas mesmas condições do anterior - em que não se encontrou nenhum caso de VS em 100 estruturas com verbo.

Esta diferença entre língua escrita e oral é explicável considerando-se que vários dos casos que a gramática arrola como de VS são estruturas típicas de língua escrita: particípio absoluto, construções com se, construções com infinitivo, orações subjetivas. Além disso, como a ordem VS é mais conservadora, a menor frequência no diálogo é explicável, pois sabe-se que a língua escrita tende a ser mais conservadora do que a oral.

Repetindo o processo anterior para verificar a frequência relativa SV versus VS de maneira mais clara, a autora contou 100 estruturas em que havia sujeito expresso claramente. Além dos dois exemplos citados acima, ela encontra apenas mais três ocorrências. Obteve-se, portanto, uma proporção de 5 VS para 95 SV, o que demonstra o caráter marcado, em termos de frequência, da ordem VS. A ordem normal, não-marcada, é SV, confirmando-se, também, a diferença entre oral e escrita, onde foram encontrados 15 exemplos de VS em 100 estruturas com sujeito expresso. Na língua oral informal, em relação à escrita formal, os casos de VS são ainda mais marcados, mais inusitados e a ordem SV é a norma.

Em seu texto, Pontes não lida diretamente com as sentenças interrogativas mas, segundo a autora, tanto na escrita

quanto na fala a ordem VS predomina nessa construção em particular. Esse resultado não coincide com o que eu obtive na minha pesquisa preliminar - onde predominou a ordem SV. Uma hipótese é de que, talvez essa discrepância nos resultados - ela encontrando mais a ordem VS e eu a SV - pode estar relacionada ao tipo de elemento Q-que ocorreu em suas sentenças. No entanto, parece-me que é interessante examinar mais de perto a diferença entre a escrita e a oralidade e, também, o contraste entre sentenças produzidas num contexto mais formal ou informal.

Uma hipótese semelhante a essa é levantada por Kato (1987). Nesse artigo, a autora propõe que a ordem V + SN ocorre em três contextos distintos no PB:

- a) construções com verbos ergativos;
- b) construções em que o verbo é anteposto ao sujeito (V-front);
- c) construções em que o SN sujeito é deslocado para a direita.

Essas construções têm em comum o fato de ocorrerem com verbos que nesses contextos só admitem um argumento foneticamente realizado.

Kato, analisando basicamente as construções em que o verbo é anteposto ao sujeito (item b acima), menciona a análise de Âmbar (1985 - a proposta apresentada pela autora nesse artigo é bastante semelhante àquela que foi mostrada anteriormente e suas diferenças são irrelevantes no momento) e coloca que a regra de V-front parece constituir um caso típico de regra variável no português, sendo que na variante européia há mais contextos categóricos do que na variante brasileira, onde o único contexto em que a regra é obrigatória é naquele em que há um advérbio dêitico em Comp.

A questão que a autora levanta nesse texto é se a regra de V-front no PB não seria uma regra estilística pelo seu caráter quase totalmente facultativo. Kato inclusive menciona o texto de Pontes (1982 - acima apresentado) e coloca que este parece apontar para um caráter estilístico dessa regra.

Lobato (1988), em resposta à questão levantada por Kato (1987 - anteriormente mencionado), propõe que a regra de V-front no PB é estritamente sintática por razões independentes do seu caráter facultativo.

A autora começa o seu artigo se perguntando em que nível - sintaxe ou Forma Lógica (FL) - se dá a aplicação dessa regra⁵. Após discutir brevemente as possibilidades descritas em Chomsky & Lasnik (1977), Lobato coloca que a regra em questão se

aplica exclusivamente na sintaxe no PB. Mais ainda, para ela, "o deslocamento de sintagma Q- é que é, pelo menos em certos casos, favorecido por V-front" (pag. 127). Tal colocação inverte a proposta de Kato (op.cit.) que postula que o alçamento do verbo é desencadeado por algum elemento em Comp. Para justificar o seu raciocínio, Lobato fornece alguns exemplos onde, segundo ela, é a presença do verbo em Comp que obriga a subida do sintagma Q- para sua posição de Espec.

26. (1) a. Onde moram os meninos?
(1')a. Onde moram vocês?

27. (9) a. Onde os meninos moram?
(9')a. Onde vocês moram?

28. (9) b. Os meninos moram onde?
(9')b. Vocês moram onde?

29. (9) c. Moram os meninos onde?
(9')c. Moram vocês onde?

Segundo a autora, 9.c. é uma sequência de uso marcado - normalmente usada como pergunta eco ou em contextos em que já se tivesse falado sobre a moradia das pessoas em questão. Sua conclusão é que "desde que escolhida a anteposição do verbo, parece que, pelo menos nos casos considerados, o natural é também deslocar o sintagma Q-" (pag. 128). Além disso, Lobato coloca que com o fronteamento do auxiliar pode-se ter uma pergunta não-eco e também não retórica sem o deslocamento de Q-.

30. (10) a. Vai você descer aonde?
b. Estava você hospedado onde no Rio?

Um problema que encontro no texto acima é que os exemplos de v-front fornecidos pela autora me parecem excessivamente marcados. Uma análise rápida do corpus por mim coletado me mostra que exemplos como os dados acima (exs. 28, 29 e 30) são inexistentes. Porém, continuemos apresentando a proposta de Lobato.

Em seguida, a autora postula que a operação de v-front ou aux-front ocorre em dois contextos básicos: (1) contextos em que elas parecem não interagir com nenhum fator gramatical e (2) contextos em que interagem com fator gramatical. São exemplos de (1):

31. (11) a. Complicaram vocês a questão.
b. Querem eles uma recompensa.
c. Procuram eles a paz.

- d. Fez ele uma safadeza com ela.
- e. Faz ele questão de voltar.
- f. Noticiaram os jornais a queda do dólar.

Em (2) encontram-se três categorias de exemplos: (a) exemplos com operador em Espec de CP; (b) exemplos com verbos epistêmicos e dicendi na matriz e (c) exemplos na forma interrogativa e imperativa.

A regra, no entanto, é variável, como a própria autora afirma, sendo que a anteposição do verbo parece não obrigatória na fala coloquial, ao contrário da formal. Dessa forma, Lobato postula a possibilidade de se gerar livremente seqüências com anteposição do verbo na sintaxe, e que só em nível posterior essas seqüências seriam checadas, a fim de ser verificado o atendimento de outras exigências gramaticais.

A meu ver, o fato de ocorrerem naturalmente sentenças onde o sintagma Q- se encontra deslocado e o verbo não foi anteposto são indícios fortes de que talvez a proposta de Lobato não seja a mais adequada.

Duarte (1992) procura identificar os fatores que mais fortemente teriam condicionado a mudança da ordem VS para SV nas construções interrogativas Q- no PB, sob a hipótese de que ela coincide com o aparecimento do expletivo é que, para depois estender-se pelas estruturas em que tal inserção não ocorre, transformando os poucos casos de inversão em exceções que dependem de outros fatores, como o tipo de elemento Q- e o número de argumentos foneticamente realizados ou ainda a estrutura argumental projetada pelo verbo. A autora trabalha tanto com dados diacrônicos (peças teatrais dos séculos XIX e XX) quanto com dados sincrônicos (linguagem das novelas de televisão de 1989), porém foram computadas apenas as estruturas com sujeito expreso e não representado pelo elemento interrogativo em Comp. Além disso foram excluídas as frases existenciais e temporais que continham sujeito elíptico, indeterminado ou topicalizado.

Os grupos de fatores por ela analisados foram:

- (a) tipo de interrogativa (direta ou indireta);
- (b) estrutura do sintagma Q- (palavra Q- ou palavra Q- + N');
- (c) palavra Q- (que ou o que);
- (d) função do sintagma Q- (argumento interno ou expressão adverbial);

(e) número de argumentos foneticamente realizados (um ou mais de um);

(f) presença do expletivo é que;

(g) estatuto lexical do sujeito (SN ou pronome);

(h) período de tempo.

A análise dos dados considerando-se os diferentes grupos de fatores em relação à ordem SV não trouxe resultados relevantes. Apresento, a seguir, os resultados obtidos a partir do cruzamento dos fatores:

- Em 1734, todas as interrogativas diretas exibem a ordem VS. O mesmo acontece 100 anos depois;

- Em 1882 e 1918, percebe-se um aumento sensível no uso da ordem SV, que a partir de então cresce de maneira acentuada, chegando a superar a anteposição do verbo;

- Nas interrogativas indiretas, a ordem SV sempre teve preferência, sendo que a partir de 1955, os índices superam 90%.

- O uso da ordem SV coincide com o aparecimento do expletivo é que em 1882.

- Em 1937, SV implementa-se no sistema e a presença do expletivo é determinante, porém, a partir daí, a ordem SV não pára de crescer, mas o expletivo deixa de ser seu único condicionador;

- Em 1975, permanece a variação SV/expletivo, porém aparecem casos de inversão que exibem o expletivo, caracterizando-se a perda de seu caráter detonador ou facilitador da ordem SV;

- Com relação às construções Q- + N, apesar de sua reduzida ocorrência, observa-se que tais estruturas apresentam a ordem VS até 1937 e, a partir daí, a ordem SV;

- Desde que começa a ser usada a ordem SV privilegia as estruturas com sujeito pronominal, tendência que cresce a partir de 1937.

A conclusão de Duarte é que a ordem VS, no PB atual, está restrita (embora não obrigatoriamente) a interrogativas que se constroem com os verbos ser, estar, e com os chamados verbos apresentativos, que, nas sentenças declarativas, normalmente apresentam o SN posposto o que deixa dúvidas quanto ao fato de se tratar realmente de inversão. A tal conjunto de verbos, soma-se a representação do sujeito por um SN lexical e o fato de que apenas este argumento se faz representar foneticamente.

A autora retoma ainda a questão levantada por Kato (1987) - "Seria a ordem VS em português resultante de uma regra sintática (como em PE e espanhol) ou seria uma regra estilística, no conceito ortodoxo do termo?"⁶ - e diz que, sem dúvida, os dados de textos recentes por ela analisados confirmam o caráter facultativo da inversão e conduzem a uma clara correlação entre 'ordem SV' e 'informalidade de estilo'. Mais ainda, Duarte considerando os dados apresentados por Pontes (1982), diz que a análise dessa autora não contradiz o que foi apontado nesse artigo, pois os dados de Pontes foram extraídos de um romance, um livro de poesias e de um exemplar da revista Isto é, que pertencem a um estilo mais formal de escrita. Dessa forma, ela acredita ter elementos para supor que a ordem VS no PB esteja relacionada não à distinção entre língua escrita e língua falada, mas à distinção entre estilo formal e informal.

Como já foi apontado no início desse capítulo, Kato (1992) coloca que, para Labov (1972), o conceito de variante está ligado a estilo, pois as variantes são necessariamente idênticas em referência e valor verdade, mas diferentes em significado social e/ou estilístico. Sendo que, para ele, a própria distinção entre discurso cuidado e discurso informal é estilística. Já para os gerativistas, uma regra é estilística se ela acontece entre a Estrutura-S e a Forma Fonética (FF).

Tendo isso em vista, Kato quer mostrar que as inversões em interrogativas Q- no PB são candidatas perfeitas tanto para a regra estilística, na concepção chomskyana, quanto para o conceito de variação sintática de Labov, uma vez que as sentenças abaixo são plenamente possíveis:

32. (7) a. Quando telefonaram os meninos?
b. Quando os meninos telefonaram?

Kato mostra que Âmbar (1985), comparando o fenômeno da inversão com pronomes interrogativos no espanhol (Torrego, 1984) e no português de Portugal, observa que, enquanto no espanhol apenas Q- argumentos obrigam o fronteamento do verbo, no PE é o traço (- referencial) da palavra Q- que obriga o verbo a subir para Comp, deixando de ser o movimento obrigatório quando o pronome interrogativo é um sintagma do tipo qu + N. O argumento sintático de Âmbar, porém, se enfraquece empiricamente ao se constatar a possibilidade de subida do verbo com a ocorrência concomitante de um complementizador - conforme já apresentado anteriormente nesse texto. Diferentemente do espanhol e do PE, o PB apresenta apenas condições pragmáticas para essa inversão, uma vez que ela é quase categórica na escrita (Pontes, 1982) e quase que restrita aos verbos inacusativos na língua oral (Duarte, 1992). Kato, inclusive, retoma Lobato (1988) quando esta afirma que a subida do verbo para Comp é que desencadearia a anteposição dos pronomes interrogativos,

o que caracterizaria essa inversão como uma regra sintática. Considera, ainda, que o fato de Lobato não diferenciar o português literário do português brasileiro coloquial enfraquece a teoria desta autora.

Mais adiante, Kato coloca que a mudança que se operou na teoria, de uma gramática de regras para uma gramática de princípios e parâmetros de variação, tornou sem sentido a discussão do caráter estilístico ou não de uma regra. Estilo sintático passa a ser entendido "como diferentes definições internas para os dispositivos que acionam o valor positivo ou negativo dos parâmetros, mas cujo 'output' em termos de língua-E deve se adequar a um determinado contexto de uso" (pag. 133).

A autora apresenta o seguinte conjunto de sentenças:

33. (10) a. Que bebe João?
b.*Que o João bebe?
c. Que que bebe o João?
d. Que que o João bebe?

Para ela (10.d) representa o padrão canônico da interrogativa (Q- SV) e, em sentenças como (10.c), têm-se apenas uma aparente inversão, cuja representação seria a de uma construção de deslocamento à direita com o pronome co-referente nulo.

Dessa forma, o quadro de possibilidades de variantes dentro do parâmetro marcado negativamente (V- para Comp) seria o seguinte:

34. (16) a. Que que ele bebe o João?
b. Que que pro bebe o João?
c. Que que o João bebe?

A conclusão de Kato é que "o que possibilita as variantes não seria o movimento do Verbo, mas a opção pelo preenchimento ou não do pronome sujeito. Isso nos leva a especular se as chamadas inversões interrogativas encontradas na língua escrita não seriam inversões aparentes do tipo deslocamento à direita com sujeito co-referente nulo. Se assim for, e supondo que o parâmetro do sujeito nulo admita, nesse contexto, a variação livre entre preenchimento ou não de pronome, as opções estilísticas estariam dentro de uma mesma gramática nuclear, (...)". Sendo que, do ponto de vista chomskiano, a regra seria estilística, pois a seleção entre ele e pro seria apenas um fenômeno de realização fonética. E, do ponto de vista laboviano, seriam variantes, uma vez que possuem o mesmo valor verdade.

Num interessante artigo, Lopes Rossi (1993) propõe uma explicação sintática para a baixa incidência de inversão Sujeito-Verbo no PB atual. Seu objetivo no texto é fazer um estudo sobre a mudança diacrônica das interrogativas do PB segundo o modelo de análise apresentado em Roberts (1990) e levantar hipóteses sobre a reanálise diacrônica e conseqüente mudança paramétrica pelas quais as interrogativas passaram. Seu corpus diacrônico abrange cinco séculos (XVI-XX) e é extraído de peças teatrais e , sincronicamente, é obtido em programas de televisão.

Lopes Rossi compara a descrição de Âmbar sobre o PE atual com os dados do português europeu clássico (PEC) e observa que, no que se refere às interrogativas, o PEC tinha características semelhantes ao PE atual no sentido de também realizar o critério-WH na sintaxe e de não obrigar o movimento de V apenas nas interrogativas com um N acompanhando o elemento Q-. Outra característica sintática importante nas interrogativas do PEC e PE atual é que, devido ao movimento de V, a atribuição de caso nominativo se dá por regência, conforme Roberts (1990), e a maior parte dos sujeitos é não lexical (pro).

Em contrapartida, analisando-se diacronicamente o PB percebe-se a perda do movimento de V, a não obrigatoriedade de movimento de elemento Q- na sintaxe e a perda de marcação de caso nominativo por regência.

Uma das hipóteses levantada por Lopes Rossi se refere às construções onde a partícula é que ocorre. Segundo ela, esse tipo de interrogativa é derivada da sentença clivada do tipo "it cleft sentence" e apresenta, portanto, uma sentença matriz e uma encaixada, e não apenas uma sentença como tem sido comumente considerado. Fazem parte da sentença matriz do PEC Que é e do PE atual e PB atual, Qu é, pois nesses últimos não apenas que, mas qualquer um dos elementos interrogativos pode ocorrer nesse tipo de interrogativa. Da sentença encaixada do PEC fazem parte o que NP V e da encaixada do PE e PB, que NP V.

Os pares abaixo, possíveis no PB e PE atuais, mostram sentenças clivadas com vários tipos de elementos focalizados e sentenças interrogativas com é que incidindo sobre esses elementos focalizados.

35. (12) Foi você que leu o livro.
Quem foi que leu o livro?
36. (13) É um livro que você leu.
O que é que você leu?
37. (14) É na escola que você vai.
Onde é que você vai?

De acordo com as características sintáticas das interrogativas do PEC e também do PE atual pode-se dizer que o V da encaixada pode ser pré ou pós-verbal porque em oração encaixada no português não há exigência de movimento de V para Comp. Com essa análise as interrogativas com é que não precisam ser consideradas como uma exceção à regra do movimento de V no português europeu com em Âmbar (op. cit.).

A autora utiliza dados de interrogativas sim/não para justificar a afirmação de que o PEC possuía um padrão sintático de interrogativas Q- de língua V2, com atribuição de nominativo por regência. O PE atual manteve o padrão sintático de interrogativa com a possibilidade de atribuir nominativo por regência e por concordância. Esse fator (atribuição de caso nominativo por regência) explica o caráter obrigatório da inversão V-S nas interrogativas diretas do PE atual, exceção feita à construção que contém um N acompanhando o elemento Q- (nesse caso, a hipótese é que esse tipo de elemento interrogativo não é um quantificador da mesma natureza que os demais).

Com relação ao PB, os dados das interrogativas sim/não mostram que houve uma mudança paramétrica no sentido da perda da estrutura CP e da atribuição de caso nominativo, que passou a ser atribuído unicamente por concordância aos sujeitos lexicais. A baixa porcentagem de sujeito não preenchido nos dados de língua falada evidencia o enfraquecimento da concordância no PB. Interrogativas do tipo Qu NP V e NP V Qu, não encontradas no PE atual, mas frequentes no PB, confirmam, segundo a autora, que houve uma mudança no parâmetro de marcação de caso nominativo e no parâmetro do movimento de Qu no PB.

O último texto sobre o PB a ser apresentado é o de Kato (1993) onde a autora propõe uma explicação para a ordem dos constituintes nas sentenças com pronome interrogativo. Nesse artigo, a autora desenvolve mais suas idéias apresentadas em Kato (1992 - op. cit.) e utiliza basicamente todos os textos resumidos acima.

Primeiramente Kato apresenta um breve resumo de algumas análises que trataram da mudança de ordem dos constituintes no português e também da representação das estruturas com Q-. Entre outros textos, ela cita Ribeiro (1991) que propõe que o português antigo era uma língua V2. Torres Morais (1993) também encontra um número significativo de inversões típicas de línguas V2 em cartas do século XVI e XVII. Com relação às interrogativas Q-, Duarte (1993) coloca que a estrutura Q- V S era a regra no século XVIII, tornando-se residual no século XX e relaciona a mudança de VS para SV com o aumento no uso do expletivo é que (que). Lopes Rossi (1993) aponta para o fato de que, no PEC, a ordem VS era obrigatória em sentenças matriz (exceto quando o elemento Q- era

seguido de um N foneticamente realizado) e muito frequente com sujeitos pronominais, o que a leva a propor uma análise de V2 para essas estruturas.

A princípio, portanto, cogitou-se que, quando ocorria a inversão VS nas interrogativas Q- no português atual, o verbo ocupava a posição de núcleo de CP, porém, como essa ordem também é possível com o expletivo é que, eles estariam em competição pela mesma posição. Aqui entra a proposta de Lopes Rossi (op. cit.) segundo a qual, nessas construções, a extração se faz da sentença clivada. Assim estaria explicada tanto a ordem VS quanto SV que ocorre com é que, pois esta última seria, na verdade, a mais comum, mas a primeira também poderia ocorrer.

Falta ainda, no entanto, explicar porque a inversão VS nas interrogativas Q- que, possivelmente, não são derivadas de clivadas também se tornou um fenômeno residual e restrito no PB.

Para explicar essa questão, Kato menciona o artigo de Lobato (1988) onde a autora, usando dados de uma modalidade mais conservadora do PB, aponta para a possibilidade de ocorrência recursiva de auxiliares antepostos ao sujeito. Essa observação da autora mais a proposta de Âmbar de adjunção de V a Comp nos casos em que há variação livre entre SV/VS nas interrogativas Q- é crucial para Kato, pois coloca sérias dúvidas quanto ao tratamento do fenômeno de inversão VS como um movimento de V para Comp.

A autora primeiramente discute as possibilidades de atribuição de caso nominativo e postula que, no PE atual e na modalidade conservadora do PB, esse caso pode ser atribuído por regência de I para o constituinte em Spec de VP, liberando a posição de Spec de IP para outro constituinte (ver proposta semelhante de Raposo a ser apresentada a seguir). Dessa forma, pode-se explicar a ordem VS nas interrogativas com palavras Q- que não sejam derivadas da sentença clivada nesses dois dialetos do português.

Porém, o PB parece ter perdido a capacidade de atribuir caso nominativo por regência (Lopes Rossi, 1993), restando-lhe apenas a possibilidade de sua atribuição por concordância de I com seu Spec. Como explicar, então, a ocorrência de interrogativas Q- com inversão VS?

Kato propõe que, nessas construções, a extração se faz de estruturas que apresentam um deslocamento à direita com pronome nulo. Ou seja, o que ocorre na verdade é uma falsa inversão. A autora postula que as sentenças abaixo seriam a base da extração e o sujeito se encontraria em posição final, sendo que o pro é que receberia o caso por concordância - caso o pronome se realizasse lexicalmente, este é que receberia o caso.

38. (32) a. (Eles) Estão aqui os meninos.
b. (Ele) Vem aí o Pedro.
c. (Ela) Faturou vinte mil cruzeiros a nossa barraca.

Como os sujeitos nulos parecem estar desaparecendo do PB (Duarte, 1993), acabará por se revelar a falsa inversão que ocorre nessas estruturas. Veja os exemplos abaixo:

39. (29) a. Onde eles estão os meninos?
b. Quanto ela faturou a sua barraca?

Finalmente, para as sentenças interrogativas com ordem SV, Kato analisa essas estruturas como sendo uma consequência da extração das sentenças clivadas, porém, aí houve o apagamento tanto da cópula quanto do complementizador que.

40. (38) d'. (Foi) [O PEDRO [(que) saiu]].

41. (39) b. (Foi) [QUEM [(que) saiu]].

A autora conclui, portanto, que, no PB, ocorreram três mudanças nos últimos séculos:

- a) o caso nominativo passou a ser atribuído somente para a posição de Spec de IP;
b) houve um aumento considerável de extrações de Q- a partir de sentenças clivadas;
c) aconteceu uma mudança no paradigma pronominal com uma diminuição no uso de construções com pro.

Esse artigo parece dar conta de todas ordens que ocorrem com pronomes interrogativos no PB. Vejamos agora qual é a proposta de Raposo para o PE.

A proposta de Raposo⁷ é de que também no PE o verbo não sobe para a posição de núcleo de CP nas interrogativas com elemento Q- e apresenta os seguintes argumentos.

Em primeiro lugar, é possível a ocorrência de advérbios que precedem obrigatoriamente, na maioria das vezes, o verbo. É o caso dos exemplos abaixo:

42. (21) a. Pergunto-me com quem ainda conversou o Luis.
??b. Pergunto-me com quem conversou o Luis ainda.

43. (30) a. Quem já encontrou o Luis?
*b. Quem encontrou o Luis já?

Se o verbo, nos exemplos acima, subisse para o núcleo de CP, o advérbio deveria ficar na sua posição original logo após o sujeito. Porém, quando isso ocorre as sentenças são sentidas como, no mínimo, estranhas.

Além disso, o verbo não precisa, necessariamente, estar na segunda posição da sentença nas interrogativas - ao contrário do que ocorre nas línguas V2 canônicas onde o verbo deve, obrigatoriamente, ocupar a segunda posição (que no caso é C) para que a sentença seja gramatical.

44. (87) a. Ao Luis, quantos livros recomendou o professor? (V3)
b. Sobre esse assunto, ao professor, que resultados apresentou o Luis? (V4)

O terceiro argumento de Raposo vem das orações subordinadas. Nelas o núcleo de Comp está preenchido pelo complementizador que e mesmo assim a inversão V-S é possível. Além disso, também podem ocorrer advérbios antes do verbo.

45. (107) a. Disseste que ainda/já/também tinha quem comprado os bilhetes?
b. *Disseste que tinha quem ainda/já/também comprado os bilhetes?
46. (108) a. A Ana diz que ainda/já/também comprou a Maria os bilhetes.
b. *A Ana diz que comprou a Maria ainda/já/também os bilhetes.

Após mostrar seus argumentos a favor da não subida do verbo para Comp, Raposo apresenta a teoria da verificação dos traços verbais (traços-V) de Chomsky (1992) que coloca o seguinte:

i. Os traços V (Agr, T) encontram-se associados a V no léxico(para além da sua posição canônica na estrutura sintática).

ii. Os traços morfológicos de T e Agr têm duas funções: verificar propriedades do V (V-relacionados) e verificar propriedades do NP (N-relacionados) que ocupa a sua posição de spec (no caso de T, verificar o tempo do V e verificar o Caso Nominativo do NP, segundo Chomsky).

iii. V sobe para Agr e T (o complexo (Agr,T) segundo Chomsky) para que os seus traços sejam verificados.

iv. NP sobe para a posição de spec de T e Agr para que os seus traços sejam verificados.

v. Qualquer um dos sub-conjuntos V ou N relacionados de Agr e T são parametrizáveis em "fortes"(morfológicos num sentido abstrato ou visíveis em Forma Fonética - FF) ou "fracos"

(invisíveis em FF);

a. se são V-fortes a subida "visível" em (Sx) do V é obrigatória; se são V-fracos, a subida de V dá-se em FL.

b. se são N-fortes, a subida "visível" do NP é obrigatória (em Sx); se são N-fracos, a subida do NP dá-se em FL (ou seja, a sintaxe "visível" é determinada por propriedades "morfológicas" dos itens lexicais que possam ter um reflexo em FF).

Tendo isso em vista, Raposo propõe que, em PE, Agr é V-forte e N-forte. Já T é, pelo menos, V-forte, deixando em aberto a possibilidade de ser N-forte ou N-fraco. Dessa forma, a atribuição do caso nominativo ficou como se segue:

i. Atribuído por T (+ fin) ao seu Spec ou a Agr regido por T (Raposo 1987).

ii. Agr N-forte em português permite a atribuição/realização de caso nominativo no seu spec.

47. (11) A que amiga a Maria ofereceu o disco? (i)

48. (12) A quem ofereceu a Maria o disco? (ii)

Raposo postula ainda, conforme Chomsky (1992), a existência de dois tipos de categorias funcionais. Os núcleos L-relacionados (Agr e T) são categorias que desencadeiam a subida de V e N para a checagem de seus traços. Já os núcleos operadores são os que licenciam um especificador com "carga" operatória, ou seja, são categorias que desencadeiam a subida de elementos muito mais relacionados com o núcleo da categoria do que com a checagem de seus traços (C, F, T, Neg, etc). Os traços de operador de um núcleo X são parametrizáveis em "fortes" - portanto, visíveis em FF - ou "fracos" - invisíveis em FF, ocorrendo só em Forma Lógica (FL).

Nesse momento, ele retoma o critério-WH⁸ proposto por Rizzi (1992) e, generalizando-o, apresenta o critério O (operador) em FL:

i. Cada X (+Oz) está numa relação spec-núcleo com um grupo (+Oz);

ii. Cada grupo (+Oz) está numa relação spec-núcleo com um X (+Oz).

e se pergunta onde são licenciados os traços Wh-.

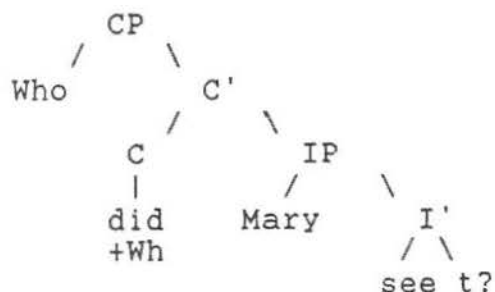
Primeiramente, eles podem ser licenciados num C selecionado por um verbo interrogativo - conforme Chomsky (1992) WH- em C é universalmente forte.

49. (14) a. Perguntaram [com quem C+WH [a Maria tinha conversado]].

b. *Perguntaram [C+WH [a Maria tinha falado com quem]].

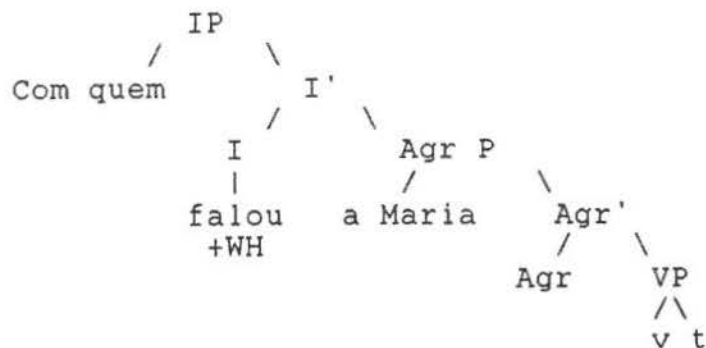
Com relação às orações principais, existem duas possibilidades. Ou eles são licenciados em C (Chomsky, 92) ou em I (Rizzi, 92). Aqui se faz importante a distinção existente entre as línguas no que se refere a atribuição do caso nominativo, pois, se por um lado algumas línguas admitem que o nominativo pode ser atribuído em Spec de IP(TP) ou de AgrP, outras só aceitam que este seja atribuído em Spec de IP. Como consequência disso, nas línguas pertencentes a esse segundo grupo, será necessário projetar-se uma posição adicional para satisfazer-se o critério WH-. É o caso do inglês, por exemplo.

50.



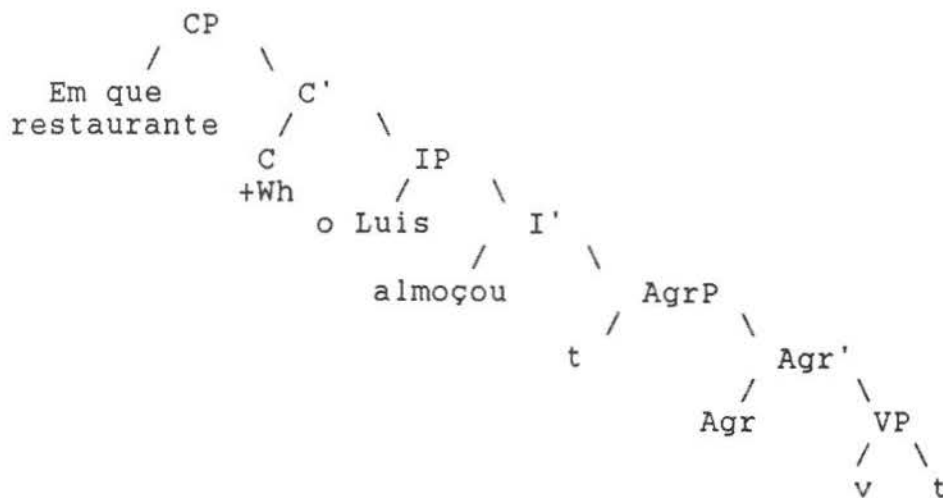
Já em línguas como o PE, o caso pode ser atribuído em Spec de AgrP, liberando o Spec de IP para a satisfação do critério Wh-.

51. (25)

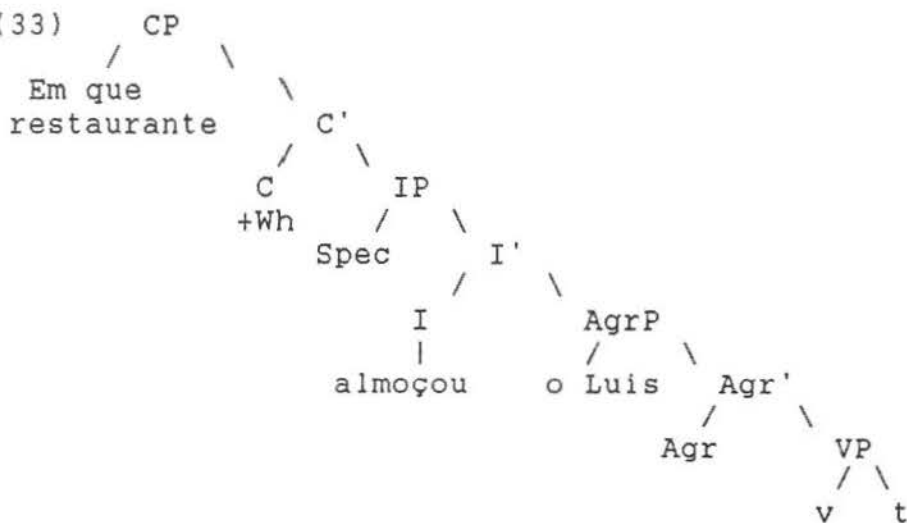


Com relação aos grupos Wh- (que correspondem à representação Q- + N de Âmbar e que não exigem a inversão V-S), Raposo retoma a definição de Pesetsky (1987) que considera-os "fortes" morfologicamente por serem ligados ao discurso (D-linked) em oposição às palavras Wh- que são "fracas" morfologicamente. Por isso, os grupos Wh- têm a propriedade da concordância dinâmica⁹ e, nesse caso, muitas vezes ocorre a projeção de CP. Ambas as sentenças abaixo são possíveis.

52.



53. (33)



Como consequência necessária de (33), tem-se que os traços N-relacionados de T podem ser fracos. Isso ocorre quando T licencia casualmente Agr, e o seu traço N-relacionado forte, está em Agr.

Concluindo, Raposo propõe que se projete CP apenas quando estritamente necessário e isso ocorreria basicamente com os grupos Wh- onde o verbo se encontra posposto ao sujeito.

Parece-me que a proposta de Raposo consegue dar conta, relativamente bem, dos fenômenos relacionados à inversão V-S que ocorre nas interrogativas com Q- no PE. No entanto, creio que sua explicação não pode ser adotada para o PB, visto ser muito comum sentenças interrogativas com palavra Q- onde a inversão V-S não se dá. Veja, por exemplo, a sentença abaixo que é agramatical no PE e perfeitamente boa no PB.

54. (77) Onde o Luis almoçou? (*PE) (ok PB)

NOTAS :

01 Os exemplos utilizados nessa seção foram retirados dos textos aqui resumidos. A numeração que constava em cada um deles é a que aparece entre parênteses no texto. Mantive, também, a terminologia usada por cada autor em seus respectivos textos.

02 Como já foi colocado anteriormente, essa sentença é considerada gramatical desde que sobre o sujeito ou sobre o elemento Q- incida um acento de intensidade.

03 Âmbar adota a Teoria da Regência e Ligação (Government and Binding Theory) proposta por Chomsky (1981) que segue, no essencial a proposta de Chomsky & Lasnik (1977). Com relação ao Princípio das Categorias Vazias, a versão por ela usada é a de Chomsky (1986), que diz o seguinte:

"B é propriamente regido por A se é regido por A e se uma certa ligação existe em A e B".

04 Segundo Âmbar, há realmente em português alguma motivação empírica para essa hipótese. Com efeito, sempre que em Português se quer dar uma interpretação de Foco ao sujeito, sem recorrer à estratégia prosódica de acento de intensidade, o verbo deve reger canonicamente o sujeito. Em frases encaixadas, tal requisito exige que o verbo se desloque para Comp. Daí frases como abaixo:

55. (46)(?) A Ana diz que compra a Maria os bilhetes.

56. (45)(?) (O) que é que ofereceu o Pedro à Joana?

Repare-se que (46) é uma frase com o mesmo estatuto de gramaticalidade que (45) acima, o que constitui um argumento a favor da hipótese de que em (45), tal como em (46), o verbo subiu para um Comp já preenchido pelo complementador que. As frases resultantes são possíveis, mas ligeiramente menos aceitáveis que aquelas em que Comp é preenchido apenas por um dos elementos (complementador ou V-flex).

05 Lobato utiliza, como Âmbar, a versão da Teoria da Regência e Ligação de Chomsky (1986).

06 O conceito ortodoxo de regra estilística, nesse texto, é definido como uma regra que não tem praticamente nenhum contexto categórico para sua aplicação.

07 Raposo apresentou sua proposta de análise das interrogativas Q- durante um curso que ministrou, no primeiro semestre de 1993, na Universidade de Campinas - UNICAMP. Os exemplos, aqui mencionados, foram obtidos nos handouts distribuídos no citado curso.

08 Rizzi (1992) propõe o seguinte Critério-WH:

a) um operador WH deve estar em uma configuração de especificar-núcleo com um X (+WH);

b) um X (+WH) deve estar em uma configuração de especificador-núcleo com um operador WH.

09 Segundo Rizzi (op.cit.), em certas línguas, uma expressão Wh-pode induzir os traços Wh- em C através de concordância spec-núcleo. É o que ele denomina Concordância Dinâmica.

IV. A DEFINIÇÃO DOS FATORES

Os textos apresentados anteriormente suscitam uma série de questões que procurarei responder com esse trabalho e que nortearam a escolha dos fatores a serem considerados na análise. As perguntas são as seguintes:

- 1) Qual é a incidência de ordem VS?
- 2) Se existe um condicionamento estilístico, que tipo de corpus apresenta uma maior produtividade de ordem VS?
- 3) O tipo de elemento Q- é relevante na ocorrência de VS?
- 4) O tipo de verbo e a hipótese da monoargumentalidade se confirmam nos dados?
- 5) A presença ou ausência do expletivo é que está afetando a ordem nas sentenças interrogativas?

A metodologia adotada nessa pesquisa prevê que seja feita uma análise estatística do corpus coletado e, para isso, foi selecionado o pacote VARBRUL. Esse pacote contém basicamente três programas que executam as funções indispensáveis para se fazer a análise e o cruzamento de variáveis. Para que o estudo estatístico seja feito por esse pacote, é necessário que as sentenças sejam codificadas segundo os fatores linguísticos e extra-linguísticos que se quer verificar. O conjunto total de dados do projeto foi dividido em três grupos: (1) dados do Projeto NURC/SP, (2) dados de língua escrita e (3) dados de língua falada. Os fatores utilizados para a codificação dos três corpora são basicamente os mesmos, só se diferenciando no grupo que especifica o tipo de corpus. A seguir, apresentarei os grupos de fatores que foram selecionados nessa pesquisa, procurando justificar a escolha de cada um deles.

1. VARIÁVEL DEPENDENTE

A primeira variável selecionada diz respeito à ordem relativa dos constituintes sujeito e verbo nas sentenças. Dessa forma, esse grupo de fatores ficou assim definido:

- 0: Sujeito Verbo
- 1: Verbo Sujeito

O objetivo principal desse trabalho é verificar a influência da ordem relativa entre o sujeito e o verbo nas construções interrogativas com elemento Q-, portanto as sentenças que possuíam ou sujeito ou verbo nulo foram separadas em dois arquivos diferentes: um para as orações sem sujeito e outro para as sem verbo. Esse procedimento foi necessário para que se pudesse avaliar com mais precisão o papel que esses constituintes desempenhavam nas sentenças em questão sem que fosse preciso fazer qualquer tipo de estipulação com relação ao caráter do elemento ausente. Ou seja, estando o sujeito ou o verbo nulo, ao fazer a codificação das sentenças, eu teria de estipular - logicamente que seguindo alguns critérios linguísticos - qual a natureza desse elemento. Portanto, decidi por separar em arquivos diferentes as construções cujo verbo ou sujeito fosse nulo, para poder obter um estudo estatístico mais expressivo sobre a influência deles na ordem das orações interrogativas em questão. Essas sentenças que apresentam ou o sujeito ou o verbo nulo não serão analisadas no presente trabalho, pois o objetivo central dessa dissertação é descrever os fatores que influenciam ou não a ordem entre sujeito e verbo e, estando um desses elementos ausentes, essa questão não se coloca.

Vejamos agora alguns exemplos da codificação feita:

57. "Que profissões elas exercem?" (NURC/SP, DID 251, 123') = 0

58. "Como é que surgiu a ideia de brincar com logotipos? (JO SOARES 11:30, 06/12/90, SBT) = 1

2. TIPO DE SUJEITO

Esse subgrupo de fatores tem por objetivo verificar a influência dos diversos tipos de sujeito na estruturação das sentenças com elemento Q-. Vários autores (Âmbar, 1988 e Duarte, 1992) apontaram para o fato de que, dependendo da natureza do sujeito (ex. pronome, expressão referencial ou NP complexo), ele tenderia a aparecer antes ou depois do verbo. Dessa forma, os seguintes tipos de sujeito foram considerados na análise:

- A: pronome pessoal ou demonstrativo
- B: pronome de tratamento
- C: expressão referencial simples*
- D: NP complexo

* Esse fator codifica os NP simples conforme exemplo 61 abaixo, sendo que, para simplificar a notação, doravante usarei, ao invés de expressão-referencial simples, apenas expressão-referencial ou, simplesmente, expressão-r.

Exemplos:

59. "Que ela poderia fazer?" (NURC/SP, DID 251, 161') = A
60. "Como a senhora vê o crescimento desse movimento chamado MUNICIPALISTA?" (RODA VIVA, 17/12/90, TV Cultura) = B
61. "O que as mães fazem, hem?" (FALA ESPONTÂNEA, 24/10/90) = C
62. "o senhor sabe como é denominada a pessoa que empresta dinheiro?" (NURC/SP, DID 250, 254') = D

É importante colocar que as sentenças onde o elemento Q- substituiu o sujeito foram separadas num outro arquivo. Pois, nesse contexto, a única ordem possível era Sujeito-Verbo, e esse resultado poderia interferir na análise estatística do corpus e acabar por levar a conclusões não totalmente corretas com relação à quantidade de interrogativas onde não ocorria a inversão. Isto é, como, nas construções em que é o sujeito o elemento interrogado, não há a alternância entre a ordem S-V e a V-S, decidi por separar essas sentenças para que não houvesse qualquer tipo de interferência desses resultados na porcentagem geral de ocorrência de estruturas Sujeito-Verbo. Dessa forma, acredito que poderei obter resultados mais precisos com relação ao tipo de ordem que predomina nesse tipo particular de construção e, portanto, verificar realmente quais os aspectos lingüísticos e extra-lingüísticos que influenciam as interrogativas com elemento Q-.

3. TIPO DE VERBO

Inúmeros foram os autores (Pontes, 1982; Kato, 1987; Duarte, 1992; entre outros) que colocaram que o tipo de verbo selecionado influenciava diretamente na estruturação das interrogativas com Q-. Por exemplo, muitos estudiosos apontaram para o fato de que os verbos transitivos eram menos favorecedores da inversão V-S, enquanto que os ergativos, os intransitivos e a cópula desencadeavam facilmente a ordem V-S. Dessa forma, as seguintes estruturas verbais foram consideradas:

- 1: cópula
- 2: ergativos
- 3: reflexivo inerente
- 4: transitivo referencial
- 5: transitivo - referencial
- 6: intransitivo

Exemplos:

63. "por que isso é tão surpreendente?" (O QUE SERÁ?, pag. 35,8)=
1
64. "Ora bolas, se até a história já morreu, o que é que sobrou,
então?" (UM DIA VAMOS RIR DISSO TUDO, pag. 112,10) = 2
65. "como é que o público se manifesta (...) depois da peça?"
(NURC/SP, DID 234, 108') = 3
66. "como você vê essa pessoa nessa profissão?" (NURC/SP, D2 62,
1272') = 4
67. "que tipo de problema demora oito anos na justiça?" (CANAL
LIVRE, 14/12/90, TV Bandeirantes) = 5
68. "a sua mãe dormia aonde?" (CANAL LIVRE, 07/12/90, TV
Bandeirantes) = 6

Como cópula foram consideradas as estruturas que apresentavam um verbo copular seguido de um adjetivo, um sintagma nominal, um sintagma preposicional, um advérbio ou um particípio. É importante ressaltar que, no corpus analisado, praticamente só ocorreram estruturas copulares com o verbo ser.

As sentenças que continham um verbo ergativo foram classificadas no segundo subfator. Não distingi entre os diversos tipos de ergativos - apresentativo, incoativo ou causativo - pois isso não me pareceu relevante para o fenômeno em questão.

Um fator codificava as sentenças que continham um verbo reflexivo inerente. É interessante comentar que, em alguns casos, o pronome reflexivo não aparecia lexicalmente expresso, o que dava margem para uma interpretação ambígua. Porém recorri ao contexto em que a sentença aparecia para poder decidir se o verbo em questão era ou não um reflexivo, como no exemplo abaixo:

69. "como é que chama a cria do cavalo?" (NURC/SP, DID 18, 685')
= 3

O quarto fator - verbos transitivos referenciais - foi o mais frequente no corpus e nele foram agrupados tanto os verbos de dois argumentos (transitivos diretos e transitivos indiretos) quanto os de três argumentos (transitivos direto e indireto). Isso porque esses últimos foram extremamente escassos no corpus, além do que, normalmente, apareciam com um dos argumentos elíptico ou em forma de um clítico.

Como transitivos [- referenciais] foram considerados os verbos que pediam um argumento locativo ou temporal. A abertura desse subfator tem por objetivo verificar a afirmação de Kato & Tarallo (1993), onde os autores propõem que a restrição de monoargumentalidade não se aplica quando um dos argumentos selecionados pelo verbo, seja ele nominal ou adverbial, aparece em primeira posição na sentença. Ou seja, com esse subfator, pretendo verificar se a inversão verbo-sujeito é favorecida quando o argumento locativo ou temporal do verbo é interrogado, sendo, portanto, colocado em primeira posição na sentença.

70. "quanto é que custa esse gravador?" (NURC/SP, DID 250, 370')
= 5

O último fator classificava as sentenças que possuíam um verbo intransitivo, pois, como já foi mencionado anteriormente (Kato & Tarallo, 1993; entre outros autores), esse tipo de verbo normalmente favorece a ordem VS.

71. "...como é que eu vou viver melhor daqui uns anos se eu não tenho mais nada hoje?" (NURC/SP, D2 62, 701') = 6

4. TIPO DE VOZ

O objetivo desse grupo de fatores é verificar se o tipo de voz verbal - passiva ou ativa - favorece ou não a inversão verbo-sujeito nas interrogativas com Q-. Dessa forma, esse grupo teve os seguintes fatores:

A: ativa
P: passiva

72. "como que a gente pode fazer alguma coisa?" (SALA DE AULA/PRIMÁRIO) = A

73. "você acham que o jornal é feito por quem?" (SALA DE AULA/PRIMÁRIO) = B

5. TIPO DE CONSTRUÇÃO VERBAL

Esse grupo de fatores procurou examinar se o tipo de construção verbal poderia ou não influenciar a ordem entre o sujeito e o verbo nas interrogativas em questão. Sendo assim, esse grupo possui dois fatores:

S: simples
C: composto

O primeiro fator agrupa todas as construções onde o verbo empregado aparece na sua forma simples, como no exemplo abaixo:

74. "que tipo de programa a senhora gosta?" (NURC/SP, DID 234, 613') = S

Por outro lado, as sentenças onde o verbo aparece numa forma composta receberam o código P. Esse fator considerou as locuções verbais formadas por um verbo auxiliar e um principal.

75. "Encontrei-me com ele na sala de visitas e perguntei-lhe por que ele havia feito aquilo." (REDAÇÃO DO VESTIBULAR/NARRATIVA) = P

6. EXPLETIVO É QUE

Esse grupo de fatores foi criado para analisar o papel do elemento é que como possível desencadeador da ordem SV nas sentenças com elemento Q-. Sendo assim, ele foi subdividido em três variantes possíveis: a primeira constituía-se de construções com o expletivo é que; a segunda possibilidade é quando, na sentença, ocorria somente a partícula que - considerando-se que nesses casos se trata do expletivo é que, porém com o verbo ser elíptico -; e, finalmente, foram agrupadas as sentenças onde o expletivo é que não aparecia.

- 1: presença de "é que"
2: presença de "que"
3: ausência de "é que"

76. "quanto é que o senhor me dá por um milhão a vista?" (NURC/SP, DID 250, 371') = 1

77. "e o que que tinha de mais?" (DE REPENTE, AS TRÊS DA TARDE, pag. 95,2) = 2

78. "quantas vezes o homem foi à lua?" (ESCOLINHA DO PROFESSOR RAIMUNDO, 08/01/91, TV Globo) = 3

7. TIPO DE ELEMENTO Q-

- A: quem
- B: que
- C: que + NP
- D: que tipo de NP
- E: o que
- F: qual
- G: como
- H: por que
- I: quanto
- J: quanto + NP
- K: onde
- L: quando
- M: qual + NP

O objetivo desse grupo de fatores é verificar se alguns tipos de elemento Q- favorecem mais a ordem VS do que outros. Conforme os fatores listados acima indicam, as sentenças foram classificadas de acordo com o tipo de elemento Q- que nelas ocorria, sendo que foram consideradas em separado todas as ocorrências de um elemento Q- acompanhado de um NP, pois, como apontado por vários autores (Âmbar, 1988 e Duarte, 1992, entre outros), nesses contextos a inversão V-S não precisa acontecer.

- 79. "quem cê vai votar?" (JÔ SOARES 11:30, 11/12/90, SBT) = A
- 80. "que que você escreveu?" (SALA DE AULA/PRIMÁRIO) = B
- 81. "que prejuízos essa guerra poderia trazer ao meio ambiente?" (RODA VIVA, 21/01/91, TV Cultura) = C
- 82. "que tipo de doce a senhora gosta?" (NURC/SP, DID 235, 112') = D
- 83. "o que significa essa chave?" (REDAÇÃO DO VESTIBULAR/NARRATIVA) = E
- 84. "qual é a significação que tá em jogo aqui?" (SALA DE AULA, 08/11/90) = F
- 85. "como eu poderia desistir?" (O HOMEM CHEIO DE SONHOS, pag. 44,7) = G
- 86. "Sabe por que é que eu trouxe a televisão? (UM DIA VAMOS RIR DISSO TUDO, pag.47,4) = H
- 87. "quanto é que você tá devendo?" (O HOMEM CHEIO DE SONHOS, pag. 9,10) = I
- 88. "e quanto peso você consegue carregar nas costas?" (JÔ SOARES 11:30, 04/12/90, SBT) = J

89. "esse leite é posto onde?" (NURC/SP, DID 18, 573') = K
90. "quando é que vai ser realizada a primeira, a segunda e a terceira viagem a Marte?" (ESCOLINHA DO PROF. RAIMUNDO, 08/01/91, TV Globo) = L
91. "qual desses pratos o senhor recomendaria?" (RODA VIVA, 10/12/90, TV Cultura) = M

8. ESTATUTO DA ORAÇÃO

Esse grupo codificou as sentenças conforme o elemento Q- foi gerado numa oração independente ou numa subordinada.

- 1: Independente
2: Subordinada

92. "o que é que o senhor vigário vai fazer todas as tardes na casa da D. Mercedes?" (DE REPENTE, AS TRÊS DA TARDE, pag. 142,9) = 1
93. "que tipo de filme a senhora acha que deveria ser passado?" (NURC/SP, DID 234, 437') = 2

9. TIPO DE CORPUS

Como já foi mencionado anteriormente, os dados foram separados em três arquivos. O primeiro deles continha somente sentenças retiradas do corpus do Projeto NURC/SP. O outro arquivo reunia as construções obtidas em textos escritos, tais como, peças de teatro, romances e redações do vestibular. O último arquivo englobava as ocorrências retiradas de situações de fala, por exemplo: entrevistas na televisão, sala de aula de adultos e crianças e, também, fala espontânea. Dessa forma, esse grupo de fatores teve três codificações, conforme seria usado em cada um dos arquivos.

9.1- Dados do Projeto NURC/SP

- A: EF
B: D2
C: DID

94. "...em média quanto ele vai ter no bolso?" (NURC/SP, EF 388, 103') = A
95. "o que é que você faz?" (NURC/SP, D2 62, 99) = B

96. "e qual é a diferença básica que haveria entre esses três museus?" (NURC/SP, DID 137, 142') = C

É importante esclarecer que tipo de contextos de fala o Projeto NURC/SP procura abranger. Primeiramente, temos as Elocuções Formais - EF - que são gravações feitas durante uma aula ou uma conferência. Em seguida, encontram-se os Diálogos entre dois informantes - D2 - que, como o próprio nome diz, se trata da gravação de uma conversa entre dois informantes. Sendo que o grau de formalidade desses diálogos varia conforme o grau de conhecimento entre os informantes e também conforme o sexo e a idade deles. Ou seja, quando os informantes não se conheciam previamente e eram do sexo masculino a conversa tendia a ser mais polida, portanto mais formal. Por outro lado, quando eram duas mulheres, mais ou menos da mesma idade, conversando, mesmo que não se conhecessem antes, o diálogo tendia a ser tornar mais rapidamente informal. Por último, temos o Diálogo entre informante e documentador - DID - que gravou as interações entre um informante e o seu entrevistador. Sendo que, ao contrário do que normalmente é feito nas análises do Projeto NURC/SP, considere também a fala do documentador, pois esse era o contexto mais freqüente de uso de interrogativas com elemento Q-. No entanto, acho importante ressaltar que, normalmente, o documentador prepara uma lista de perguntas que poderiam ser usadas durante a entrevista e isto pode influenciar o tipo de construção interrogativa.

9.2- DADOS DE ESCRITA

B: peça de teatro
C: romance
F: narração
G: dissertação
H: carta

Os fatores acima codificam as sentenças conforme o tipo de contexto do qual elas foram retiradas. Os textos considerados foram duas peças de teatro (O QUE SERÁ? e O HOMEM CHEIO DE SONHOS), dois romances (UM DIA VAMOS RIR DISSO TUDO e DE REPENTE, AS TRÊS DA TARDE) e redações do vestibular da Unicamp, que podiam ser de três formas: dissertação, narração e carta. Todos os textos mencionados anteriormente foram escritos depois de 1980 e o objetivo desse grupo de fatores é verificar se o tipo de texto influencia no uso de uma determinada construção interrogativa. Isto é, ao confrontar textos profissionais - peças teatrais e romances - com textos, vamos dizer, não-técnicos - redações do vestibular -, quero ver se existem diferenças significativas quanto ao uso de VS ou SV em cada um deles. A hipótese é que, nas redações de vestibular, tende-se a usar estruturas mais arcaicas, pois, nesse contexto, existe uma forte pressão pela norma. Já os escritores de

textos mais profissionais possuem maior liberdade de uso da linguagem, muitas vezes optando por utilizar um tipo de escrita bem próximo da linguagem coloquial. Sendo assim, a escolha desses diversos tipos de textos escritos tem por objetivo principal verificar se o contexto de escrita do mesmo influencia no tipo de estruturas nele empregadas.

97. "Mas por que ela deixa?" (O QUE SERÁ?, pag. 25,14) = B
98. "por quanto você vai vender?" (O HOMEM CHEIO DE SONHOS, pag. 21,5) = B
99. "o que foi que ela disse, heim?" (UM DIA VAMOS RIR DISSO TUDO, pag. 30, 8) = C
100. "o que será que a Julieta estava fazendo aquela hora?" (DE REPENTE, ÀS TRÊS DA TARDE, pag. 97,4) = C
101. "como alguém centraria o tiro tão perfeitamente?" (REDAÇÃO DO VESTIBULAR/NARRAÇÃO) = F
102. "onde estaria o super-man para fazer o tempo retroceder?" (REDAÇÃO DO VESTIBULAR/DISSERTAÇÃO) = G
103. "por que alguém escrever-me-ia?" (REDAÇÃO DO VESTIBULAR/CARTA) = H

9.3- DADOS DE FALA

Esse arquivo engloba as ocorrências de interrogativa com elemento Q- obtidas nos seguintes contextos de fala:

A: Canal Livre
J: Jô Soares 11:30
K: Cara a Cara
L: Escolinha do Prof. Raimundo
M: Metrôpolis
N: Roda Viva
O: 40 anos de Televisão
D: sala de aula de criança
E: sala de aula de adulto
I: fala espontânea

Em princípio todos os programas de entrevista na televisão estavam reunidos num único código - A -, porém mais tarde achei interessante separá-los para verificar se o estilo de cada programa estava favorecendo o aparecimento de um determinado tipo de construção. Sendo assim, esse grupo de fatores acabou por

considerar os contextos de fala conforme apresentado acima. Além das entrevistas na televisão, coletei dados em sala de aula de adultos - são sentenças coletadas durante as aulas da graduação (curso que fazia durante o período de coleta desses dados) - e também da fala espontânea. Os dados de sala de aula de criança foram obtidos a partir das anotações e gravações feitas por uma colega durante o seu projeto de Iniciação Científica. Gostaria de expressar aqui o meu agradecimento à Cláudia Regina Castellano Pfeiffer por me ceder seu material para coletar dados.

104. "quantos juízes nós precisamos para cada caso?" (CANAL LIVRE, 14/12/90, TV Bandeirantes) = A
105. "como é que cê faz o seu programa?" (JÔ SOARES 11:30, 10/12/90, SBT) = J
106. "como é que cê se meteu nessa?" (CARA A CARA, 06/01/91, TV Bandeirantes) = K
107. "quando é que vai ser realizada a primeira, a segunda e a terceira viagem à Marte?" (ESCOLINHA DO PROFESSOR RAIMUNDO, 08/01/91, TV Globo) = L
108. "o que que você espera da escalada do rock?" (METRÓPOLIS, 04/12/90, TV Cultura) = M
109. "por que não existe uma maior identificação?" (RODA VIVA, 17/12/90, TV Cultura) = N
110. "o que quer dizer isso?" (40 ANOS DE TV, 04/12/90, TV Cultura) = O
111. "quem é que permite que as notícias sejam escritas no jornal?" (SALA DE AULA/PRIMÁRIO) = D
112. "qual é a especificidade desse campo?" (SALA DE AULA, 09/11/90) = E
113. "que cê vai fazer com essas duas chupeta?" (FALA ESPONTÂNEA, 09/01/90) = I

10. DIMENSÃO DO SUJEITO

O último grupo de fatores a ser verificado é o que considera a dimensão do sujeito. Esse grupo tem por objetivo verificar se o tamanho do sujeito influencia no seu posicionamento antes ou depois do verbo nas sentenças com pronome interrogativo. A dimensão do sujeito foi verificada segundo o número de sílabas que o compunham, sendo que a contagem sílabica levou em conta as

elisões que normalmente ocorrem na fala. Dessa forma, foram os seguintes os fatores selecionados:

- 1: uma sílaba
- 2: duas sílabas
- 3: três sílabas
- 4: quatro sílabas
- 5: cinco sílabas
- 6: seis sílabas
- 7: sete sílabas
- 8: oito sílabas
- 9: nove sílabas
- 0: dez ou mais sílabas

Exemplos:

- 114. "e fora de São Paulo eles representavam essa pecas onde?" (NURC/SP, DID 161, 451') = 1
- 115. "por que alguém escrever-me-ia?" (REDAÇÃO DO VESTIBULAR/CARTA) = 2
- 116. "o que é que é editora?" (SALA DE AULA/PRIMÁRIO) = 3
- 117. "como fica sua carreira?" (O QUE SERÁ?, pag. 4,10) = 4
- 118. "como é que as autoridades prepararam a cidade para o verão?" (CANAL LIVRE, 04/01/90, TV Bandeirantes) = 5
- 119. "como é que vai ser o desempenho dele?" (NURC/SP, EF 377, 348') = 6
- 120. "qual o teu próximo projeto?" (JÔ SOARES 11:30, 14/12/90, SBT) = 7
- 121. "por que novela de televisão faz sucesso?" (NURC/SP, DID 161, 377') = 8
- 122. "na sua opinião a infidelidade conjugal é mais marcada em que camada social?" (CANAL LIVRE, 06/12/90, TV Bandeirantes) = 9
- 123. "o que deve a sociedade tão privilegiada fazer atualmente quanto a isso?" (REDAÇÃO DO VESTIBULAR/DISSERTAÇÃO) = 0

As 708 sentenças que compõem o conjunto de dados desse projeto foram codificadas segundo os fatores apontados acima. Naturalmente, nas etapas posteriores do programa VARBRUL, foi necessário que se fizesse algumas alterações nos grupos, como, por exemplo, a amalgamação de fatores. Entretanto, deixarei para apresentar e justificar as modificações que foram feitas no capítulo seguinte, onde, também, procurarei descrever os resultados obtidos na análise quantitativa dos dados.

V. ANÁLISE DOS FATORES

Como já foi dito anteriormente, o conjunto de dados dessa pesquisa se compõe de 708 sentenças. Esse corpus foi dividido em três subconjuntos conforme o contexto no qual as sentenças foram obtidas - Projeto NURC/SP, Dados de Língua Falada e Dados de Língua Escrita. Essa separação objetivou ter-se uma visão mais clara da estruturação desse tipo particular de interrogativa segundo os diversos contextos de uso. As sentenças que não possuíam um sujeito lexicalmente expresso foram colocadas num arquivo em separado. O mesmo procedimento foi tomado com as construções onde o verbo estava elíptico. Assim como as sentenças em que o elemento Q-interrogativo substituiu o sujeito foram, também, agrupadas num outro arquivo.

Dessa forma, os dados, cuja análise apresentarei a seguir, são compostos por sentenças onde o sujeito e o verbo aparecem lexicalmente preenchidos e o elemento Q- não substitui o sujeito da mesma. Para facilitar a apresentação dos resultados, mantereí a separação do corpus feita anteriormente, sendo que primeiro mostrarei análise dos dados do Projeto NURC/SP, depois serão os dados de Língua Escrita e, por último, os dados de Língua Falada. No próximo capítulo, procurarei mostrar os resultados convergentes e divergentes entre os diversos tipos de corpus.

A análise estatística tem como função calcular as probabilidades associadas aos fatores de cada variável estudada, com base nas frequências observadas, além de apresentar uma seleção estatística das diversas variáveis analisadas. Um outro aspecto desse pacote é que ele trabalha com níveis diversos de análise, efetuando-se comparações entre os valores probabilísticos atribuídos aos fatores das variáveis. Esse procedimento é bastante interessante, pois ele nos permite verificar precisamente a interferência entre as variáveis, ocasionada por codificação superposta. Quando não há sobreposições ou enviesamentos entre variáveis, não se constata interferências e as probabilidades do primeiro nível permanecem as mesmas, ou bastante semelhantes, até o último nível de análise. Esta seria a situação lingüística e matematicamente ideal, mas não constitui a realidade dos fatos. Se lingüisticamente os fatos sempre ocorressem dessa forma, não haveria necessidade de sofisticação matemática na análise dos dados: o cálculo de percentagens seria suficiente.

No caso de haver sobreposições entre variáveis, o programa atribui pesos probabilísticos relativos em função da importância estatística de cada uma delas, com base, por exemplo, numa distribuição equilibrada dos dados entre os diversos fatores de uma variável.

Essa versão do programa, além de apresentar um instrumental estatístico mais elaborado, permite-nos assim ver com clareza toda a questão de codificações superpostas, conduzindo-nos a um entendimento mais apurado da distribuição dos dados lingüísticos e permitindo-nos análises lingüísticas cada vez melhores.

Antes de começar a apresentar os resultados obtidos com cada um dos conjuntos de dados apontados acima, gostaria de explicar alguns procedimentos relativos ao pacote VARBRUL. Esse programa não procede com a análise estatística se duas situações ocorrerem: KNOCKOUT e SINGLETON. O primeiro caso se dá quando, com um determinado fator, apenas uma ordem se apresenta. Por exemplo, se em todas as sentenças com verbo intransitivo só ocorre a ordem Sujeito Verbo, nesse caso houve KNOCKOUT, pois com relação a um certo fator apenas uma das possibilidades da variável dependente ocorre. Quando essa situação se apresenta, têm-se três maneiras de resolvê-la: ou desconsideramos essas sentenças ou fazemos sua amalgamação com outro fator semelhante ou, ainda, inventamos uma sentença com esse mesmo tipo de verbo, mas cuja ordem seja Verbo Sujeito. Na análise do meu corpora, eu encontrei vários knockouts e em cada situação uma maneira de solucionar o problema foi a adotada, no entanto, deixarei para justificar cada uma das decisões tomadas quando estiver apresentando os resultados de cada grupo de fatores, pois, acredito que assim fica mais fácil de entender as razões do que foi feito. Quanto o caso de SINGLETON, ele se dá quando, num determinado grupo de fatores, apenas uma variável ocorre. Isto é, se, por exemplo, no grupo que considera o tipo de voz do verbo só tivéssemos sentenças com o verbo na ativa teríamos o SINGLETON. Na minha análise, essa situação só ocorreu em um dos corpora.

Vou agora apresentar os resultados obtidos com os dados do PROJETO NURC/SP.

1. DADOS DO PROJETO NURC/SP

O corpus do Projeto NURC/SP que aqui será analisado contém 187 sentenças e os resultados gerais da quantificação feita mostram que a ordem Verbo-Sujeito ocorre em 37% dos dados, sendo que a probabilidade dessa ordem acontecer quando todos os fatores relevantes são considerados é de .27. É necessário agora caracterizar melhor essa ocorrência e verificar como ela se dá. Meu objetivo aqui é, justamente, descrever os resultados obtidos no tratamento quantitativo dos dados e tentar compor o quadro dos fatores que estão associados à produção da ordem V-S nas estruturas com elemento Q-.

A análise probabilística seleciona as variáveis mais relevantes para a compreensão do fenômeno pela pesagem de fator a

fator. A listagem hierarquizada por grau de relevância apontou os seguintes grupos de fatores: tipo de sujeito, tipo de verbo, tipo de corpus e tipo de elemento Q-. Os demais fatores foram considerados irrelevantes para a ordem encontrada nas sentenças com elemento Q-, no entanto, apresentarei, também, os resultados obtidos para essas variáveis e farei uma breve discussão sobre eles. Primeiramente, porém, mostrarei a estatística para as variáveis relevantes.

1.1- Tipo de sujeito

As variáveis analisadas nesse grupo codificavam o sujeito conforme ele era um pronome pessoal ou demonstrativo, um pronome de tratamento, uma expressão referencial ou um NP complexo e os primeiros resultados mostram o seguinte:

TIPO DE SUJEITO	FREQUÊNCIA V-S		FREQUÊNCIA S-V	
NP complexo	100%	19/19	0%	0/19
Expressão-R	78%	47/60	22%	13/60
Pronome pessoal	5%	4/84	95%	80/84
Pronome tratamento	0%	0/24	100%	24/24

Tabela 01: Frequência de ordem S-V e V-S em relação ao tipo de sujeito.

A tabela acima mostra a frequência tanto da ordem S-V quanto V-S com relação ao tipo de sujeito da sentença e as seguintes conclusões podem ser tiradas:

1) Pode-se notar uma distribuição complementar entre os tipos de sujeito. Se, por um lado, a expressão referencial e o NP complexo favorecem a ordem Verbo-Sujeito, por outro lado, os pronomes pessoais e os de tratamento preferem a ordem Sujeito-Verbo.

2) Em todas as sentenças em que o sujeito era um pronome de tratamento a ordem que ocorria era Sujeito-Verbo. Assim como em todas as construções cujo sujeito era um NP complexo o verbo aparecia anteposto ao mesmo.

A situação apontada na observação nº 2 acima não pode persistir, pois mantendo-se assim a análise probabilística não poderá ser feita, uma vez que, nesse grupo, dois fatores apresentam um índice de 100% com relação a ordem relativa entre sujeito e verbo - NP complexo = 100% de VS e pronome de tratamento = 100% de SV - o que caracteriza uma situação de KNOCKOUT. Para solucionar esse problema, optei por fazer amalgamações. Ou seja, como os pronomes pessoais também favoreciam fortemente a ordem SV, coloquei os pronomes de tratamento junto com estes. Procedimento semelhante

fiz com os NP complexos, pois amalgamei estes com as expressões referenciais que apresentavam um mesmo comportamento em relação à ordem Sujeito Verbo. A tabela abaixo, portanto, mostra a frequência e o peso relativo de V-S segundo o tipo de sujeito depois de feitas todas as amalgamações descritas acima.

TIPO DE SUJEITO	FREQÜÊNCIA	PESO RELATIVO
Expressão-R	84% 66/ 79	.95
Pronomes	4% 4/108	.10

Tabela 02: Frequência e Peso relativo de V-S segundo o tipo de sujeito.

A tabela acima mostra um comportamento complementar entre os dois tipos de sujeito considerados. Enquanto as expressões-referenciais nitidamente favorecem a ordem Verbo-Sujeito, apresentando, no corpus estudado, uma frequência de 84% dessa ordenação. Por outro lado, os pronomes claramente inibem esse tipo de construção. Além disso, podemos, inclusive, dizer que uma interrogativa com elemento Q- cujo sujeito é uma expressão-referencial tem uma probabilidade de .95 de ocorrer com a ordem Verbo-Sujeito. Voltarei novamente a esse grupo de fatores depois que tiver apresentado a influência dos tipos de verbo nessas construções quando mostrarei o cruzamento desses dois grupos.

1.2- Tipo de verbo

A primeira análise do corpus do NURC/SP obteve os seguintes resultados:

TIPO DE VERBO	FREQÜÊNCIA DE V-S	FREQÜÊNCIA DE S-V
Ergativos	100% 1/ 1	0% 0/ 1
Cópula	94% 51/ 54	6% 3/ 54
Reflexivos	58% 7/ 12	42% 5/ 12
Trans. -ref.	13% 1/ 8	87% 7/ 8
Trans. +ref.	9% 10/109	91% 99/109
Intransitivos	0% 0/ 3	100% 3/ 3

Tabela 03: Frequência de V-S e S-V conforme o tipo de verbo.

Tendo em vista a tabela 03, podemos verificar que os verbos ergativos e a cópula favorecem fortemente a anteposição do verbo em relação ao sujeito. Por outro lado, os verbos transitivos [- e +] referenciais e os intransitivos normalmente aparecem depois do sujeito. Além disso, os reflexivos apresentam um comportamento basicamente neutro com relação ao sujeito, pois a frequência de ocorrência antes do sujeito é pouco menor do que o seu aparecimento

depois deste.

Assim como aconteceu com os fatores que consideravam o tipo de sujeito, nesse grupo também ocorreram dois knockouts (verbos ergativos = 100% de V-S e verbos intransitivos = 100% de S-V). Porém, como as variáveis acima representavam um número muito pequeno de ocorrências (1 ergativo e 3 intransitivos), optei por eliminar localmente essas sentenças. Ou seja, nos cálculos posteriores desse programa, as construções com esses dois tipos de verbos não serão consideradas quando estiver sendo feita a análise desse grupo de fatores. Porém essa eliminação só se dará nesse contexto e o programa considerará essas sentenças quando estiver analisando os outros tipos de fatores.

Além dessa modificação, decidi amalgamar os dois tipos de transitivos considerados na análise (transitivos - referenciais e transitivos + referenciais), pois ambos os tipos de verbos apresentavam um comportamento bastante semelhante com relação ao seu aparecimento antes ou depois do sujeito. Esse resultado é muito interessante, pois corrobora a tese da monoargumentalidade de Kato & Tarallo. Somando-se a isso, tem-se que o número de ocorrências com transitivos [- referenciais] é bastante restrito - apenas 8 sentenças - e, persistindo essa situação, essas construções deveriam ser eliminadas do corpus por não poderem ser consideradas representativas. Portanto, os dois tipos de transitivos foram unificados e considerados como uma única variante.

TIPO DE VERBO	FREQÜÊNCIA V-S		PESO RELATIVO
Cópula	94%	51/ 54	.97
Reflexivos	58%	7/ 12	.69
Transitivos	9%	9/117	.15

Tabela 04: Frequência e Peso Relativo de V-S segundo o tipo de verbo.

Os resultados acima mostram que a cópula desencadeia a ordem V-S em 94% dos casos, sendo que a probabilidade de ocorrência dessa ordem chega a .97 com esse verbo. Os reflexivos mantêm o seu comportamento quase neutro, aparecendo antes do sujeito em 58% dos casos. O peso relativo desse tipo de verbo, no entanto, aponta para uma tendência dele ocorrer antes do sujeito (.69 de probabilidade). Comportamento completamente diferente é o verificado com os verbos transitivos, pois estes inibem claramente o aparecimento do sujeito em posição pós-verbal, uma vez que essa situação só ocorreu em 9% dos dados. O peso relativo desse verbo mostra que a ordem Verbo-Sujeito possui uma chance de .15 de acontecer. Abaixo mostrarei o cruzamento entre esse grupo de fatores e o que considera o tipo de sujeito, com o objetivo de

verificar o comportamento dessas duas variáveis mais de perto.

TIPO DE VERBO	T I P O D E		S U J E I T O		TOTAL
	Expressão-R (84%)		Pronomes (4%)		
Cópula (94%)*	98%	49/50	50%	2/ 4	94% 51/54
Reflexivos (58%)	71%	5/ 7	40%	2/ 5	58% 7/12
Transitivos (9%)	55%	11/20	0%	0/97	9% 11/117
TOTAL	84%	65/77	4%	4/106	38% 69/183

Tabela 05: Frequência de V-S a partir do cruzamento de tipo de verbo com tipo de sujeito.

* os números entre parênteses representam os percentuais originais de frequência.

O cruzamento entre grupos de fatores é importante para se verificar precisamente a força explanatória dessas variáveis, pois se os percentuais gerados no cruzamento forem bem próximos dos obtidos anteriormente isso significa que os fatores apontados são muito relevantes para a análise. E o que se observa na tabela 05 acima é que, sendo o sujeito uma expressão-referencial, ocorre uma queda nos percentuais de ordem V-S a medida que passamos de um verbo copular (98%), para um reflexivo (58%) e, finalmente, para um transitivo (55%). Este último, inclusive, não apresenta nenhuma sentença com o verbo anteposto quando o sujeito é um pronome. Ou seja, anteriormente tínhamos visto que a cópula é o tipo de verbo que mais favorece a ordem V-S e esse resultado se confirma plenamente quando o sujeito é uma expressão-referencial. Ocorre uma queda quando este é um pronome (50%), entretanto o número de orações onde esses dois fatores aparecem simultaneamente é muito pequeno (apenas 4). Por outro lado, os transitivos tendiam a apresentar a ordem S-V, e a tabela acima mostra que, em todos os contextos onde o verbo se encontrava anteposto ao sujeito, este era uma expressão-referencial. Isto quer dizer que, quando duas variáveis tidas como favorecedoras da ordem S-V acontecem numa mesma oração, obtive 100% de aplicação da não inversão Verbo-Sujeito - caso em que o verbo é um transitivo e o sujeito um pronome. Os reflexivos inerentes apresentavam um comportamento mais neutro em relação à ordem e o que se verifica nessa tabela é que o tipo de sujeito influencia decisivamente na posição verbal, pois, sendo o sujeito uma expressão-r, este tende a aparecer depois do verbo.

1.3- Tipo de corpus

O cálculo de peso relativo das variáveis selecionou o tipo de corpus como um fator relevante para a ordem do sujeito e do verbo nas sentenças com elemento Q-. Vejamos, portanto, os números para esse grupo de fatores.

TIPO DE CORPUS	FREQUÊNCIA V-S		PESO RELATIVO
Elocução Formal	43%	9/ 21	.36
Diálogo Informante/ documentador	43%	52/122	.70
Diálogo 2 Informan- tes	21%	9/ 43	.11

Tabela 06: Frequência e peso relativo de V-S segundo o tipo de corpus.

Na Elocução Formal(EF) e no Diálogo entre Informante e Documentador (DID), 43% das ocorrências apresentavam a ordem Verbo-Sujeito. Provavelmente isso se dê porque, nesse último tipo de corpus (DID) foram consideradas também a fala do documentador, uma vez que a maioria das perguntas partia dele, e estas, muitas vezes, eram preparadas com antecedência - nesses casos, a tendência é que se preste maior atenção a estrutura da sentença, aproximando-a mais da fala padrão cuidada. Portanto, temos que tanto EF quanto DID apresentaram um comportamento mais cuidado no que diz respeito à ordem dos constituintes sujeito e verbo, sendo que o Diálogo entre 2 Informantes (D2) mostrou-se mais coloquial, preferindo marcadamente a ordem S-V (79% dos casos).

Analisando-se agora as probabilidades, percebe-se um aumento significativo no peso relativo no que se refere ao DID. Note-se que, enquanto a ordem V-S ocorre em 43% dos casos, quando considera-se o peso relativo dessa variável em relação a todas as outras, observa-se que a probabilidade dessa ordem ocorrer em sentenças do DID é de .70. Com as EF, obtem-se uma ligeira queda na probabilidade - índice de .36 -, pois a percentagem de ordem V-S era de 43% . Já, com as estruturas do D2, o peso relativo indica que a chance do verbo aparecer antes do sujeito é de .11.

Apresentarei, a seguir, duas tabelas de cruzamento de fatores. Uma cruza o tipo de corpus com o tipo de sujeito e, a outra, esse primeiro fator com o tipo de verbo.

TIPO DE CORPUS	T I P O D E S U J E I T O					Total
	Expressão-R (84%)		Pronomes(4%)			
EF (43%)*	80%	8/ 10	9%	1/ 11	43% 9/21	
DID (43%)	88%	50/ 57	3%	2/ 65	43% 52/122	
D2 (21%)	67%	8/ 12	3%	1/ 31	21% 9/43	
TOTAL	84%	66/ 79	4%	4/107	38% 70/186	

Tabela 07: Frequência de V-S a partir do cruzamento de tipo de sujeito com tipo de corpus.

TIPO DE CORPUS	T I P O D E V E R B O					Total
	Cópula (94%)		Transitivos (9%)		Reflexivos (58%)	
EF (43%)*	100%	9/ 9	0%	0/10	0% 0/ 1	45% 9/20
DID (43%)	100%	34/34	14%	11/77	60% 6/10	42% 51/121
D2 (21%)	73%	8/11	0%	0/29	100% 1/ 1	22% 9/41
TOTAL	94%	51/54	9%	11/116	58% 7/12	38% 69/182

Tabela 08: Frequência de V-S a partir do cruzamento de tipo de verbo com tipo de corpus.

* os números entre parênteses representam os percentuais originais de frequência.

As tabelas acima apontam na direção de que tanto o tipo de sujeito quanto o tipo de verbo influenciam mais a ordem nas interrogativas do que o tipo de corpus propriamente dito. Veja que, em ambos os cruzamentos, os resultados se aproximam mais dos obtidos anteriormente para essas duas variáveis. Por exemplo, vimos que, nos dados do D2, a ordem V-S ocorria em 21% dos casos, sendo que a probabilidade dela ocorrer chegava a atingir o índice de .11, porém, ao fazer o cruzamento desse fator com o que codificava o tipo de sujeito, percebemos que o verbo aparece anteposto em 67% das sentenças cujo sujeito é uma expressão-referencial. Da mesma forma que, observando a tabela 08, notamos que, no corpus do DID, todas as sentenças com verbo transitivo apresentam o verbo posposto ao sujeito, apesar do estudo comparativo entre as variáveis atribuir peso .70 à ordem V-S, nesse contexto de fala.

O que podemos concluir até aqui é que, mesmo no contexto de fala - único fator extra-linguístico sendo considerado -, a tipologia do sujeito e do verbo é que têm influenciado mais diretamente a estruturação das interrogativas com elemento Q-.

1.4- Tipo de elemento Q-

A primeira tabela mostra a análise preliminar feita considerando-se todos os tipos de elemento Q- que apareceram no corpus do PROJETO NURC/SP.

TIPO DE ELEMENTO Q-	FREQUÊNCIA V-S		FREQUÊNCIA S-V	
Qual	96%	26/27	4%	1/27
Quem	50%	3/ 6	50%	3/ 3
Como	44%	23/52	56%	29/52
O que	27%	12/45	73%	33/45
Quanto	25%	1/ 4	75%	3/ 4
Que	17%	2/12	83%	10/12
Que + NP	14%	1/ 7	86%	6/ 7
Por que	11%	1/ 9	89%	8/ 9
Onde	10%	1/10	90%	9/10
Que tipo de NP	0%	0/13	100%	13/13
Quando	0%	0/ 2	100%	2/ 2

Tabela 09: Frequência de V-S e S-V conforme o tipo de elemento Q-.

Primeiramente, gostaria de lembrar que se encontram representadas na tabela acima apenas as sentenças onde o elemento Q- não substitua o sujeito. Pois, nas construções em que o pronome interroga sobre este, a ordem encontrada é sempre Sujeito-Verbo e essa regularidade poderia enviesar os resultados estatísticos, atribuindo à ordem S-V um peso maior do que ele poderia ter realmente.

Sendo assim, temos que o pronome qual fortemente favorece a anteposição do verbo, uma vez que em 96% dos casos é essa a ordem obtida. Já os elementos quem e como apresentam um comportamento basicamente neutro, pois as percentagens de ordem S-V e V-S são bastante próximas. No entanto, é preciso ressaltar que o número reduzido de ocorrências com quem - somente três sentenças de cada - pode não refletir o que realmente acontece com as construções com esse tipo de pronome. Deixarei, porém, para verificar mais de perto o comportamento desse elemento Q- quando estiver analisando o cruzamento desse grupo de fatores com os que consideram o tipo de sujeito e o tipo de verbo.

Os demais tipos de elemento Q- ocorrem, na grande maioria das sentenças, com a ordem Sujeito-Verbo (o que = 27%, quanto = 25%, que = 17%, que + NP = 14%, por que = 11%, onde = 10%, quanto e que tipo de NP = 0%). Nesse grupo, aconteceram 2 knockouts - com quanto e que tipo de NP. Nos dois casos, optei por fazer amalgamações: anexe o que tipo de NP ao que + NP porque ambos os

pronomes apresentavam um mesmo comportamento com relação à ordem relativa entre sujeito e verbo; e, como, com vários fatores, as células apresentavam um número bem reduzido de ocorrências, acabei decidindo unir as variáveis por que, quanto, quando, e onde num único fator, pois assim eu aumentava o tamanho da célula, eliminava um knockout e isso, de forma alguma, afeta estatisticamente essa variável, uma vez que os resultados acima apontam para um comportamento bem semelhante entre esses tipos de elemento Q-. Vejamos, agora, como ficou a frequência e o peso relativo de V-S quando se considera os diversos tipos de elemento Q-, após as amalgamações feitas.

TIPO DE ELEMENTO Q-	FREQUÊNCIA V-S		PESO RELATIVO
Qual	96%	26/27	.89
Quem	50%	3/ 6	.73
Como	44%	23/52	.46
O que	27%	12/45	.53
Que	17%	2/12	.86
Porque/quanto/onde/ quando	12%	3/25	.05
Que + NP	5%	1/20	.45

Tabela 10: Frequência e Peso Relativo de V-S segundo o tipo de elemento Q-.

Os resultados acima apontam para o fato de que, probabilisticamente falando, os pronomes qual, que e quem favorecem à ordem Verbo-Sujeito, enquanto que o como, o o que e que + NP apresentam um comportamento quase neutro em relação à posição do sujeito e do verbo na sentença e, somente, os elementos porque/quanto/onde/quando inibem a anteposição o verbo. No entanto, é importante ressaltar que o número de ocorrências de V-S com que e quem é bastante reduzido, sendo necessário comparar esses resultados com os dos demais corpora. Inclusive um dado que chama bastante a atenção na tabela 10 é que a percentagem de anteposição do verbo quando usamos que é de apenas 17%, porém a probabilidade dessa ordem ocorrer quando comparamos todos os fatores chega a .86. Dessa forma, acho interessante cruzar esse grupo de fatores com os resultados que consideram o tipo de sujeito e de verbo.

TIPO DE ELEMENTO Q-	T I P O D E		S U J E I T O		Total	
	Expressão-R (84%)		Pronomes (4%)			
Qual (96%)*	100%	25/25	50%	1/ 2	96%	26/27
Quem (50%)	100%	3/ 3	0%	0/ 3	50%	3/ 6
Como (44%)	88%	22/25	4%	1/27	44%	23/52
O que (27%)	85%	11/13	3%	1/32	27%	12/45
Que (17%)	100%	1/ 1	9%	1/11	17%	2/12
Porque...(12%)	30%	3/ 7	0%	0/15	12%	3/25
Que + NP (5%)	50%	1/ 2	0%	0/18	5%	1/20
TOTAL	84%	66/79	4%	4/108	37%	70/187

Tabela 11: Frequência de V-S a partir do cruzamento de tipo de elemento Q- com tipo de sujeito.

* os números entre parênteses representam os percentuais originais de frequência.

A tabela acima é bastante interessante, pois mostra altos índices de V-S quando o sujeito é uma expressão-referencial, e, sendo o sujeito um pronome, a percentagem de anteposição do verbo cai significativamente - exceção do caso de qual que é igual a 50%, porém a célula é bastante pequena para ser considerada representativa. Têm-se, com quem, uma distribuição complementar, pois os três casos de V-S acontecem quando o sujeito é uma expressão-referencial e, por outro lado, sendo o sujeito um pronome, este sempre aparece antes do verbo. Esses números atribuem ao sujeito um peso maior do que ao tipo de elemento Q- na definição da ordem nas estruturas interrogativas.

TIPO DE ELEMENTO Q-	T I P O D E		V E R B O		Reflexivos		Total
	Cópula (94%)		Transitivos (9%)		(58%)		
Qual(96%)*	100%	26/26	0%	0/ 1	0%	0/ 0	96%
Quem (50%)	100%	2/ 2	25%	1/ 4	0%	0/ 0	50%
Como (44%)	94%	15/16	19%	5/27	43%	3/ 7	46%
O que(27%)	100%	8/ 8	8%	3/36	100%	1/ 1	27%
Que (17%)	0%	0/ 0	0%	0/10	100%	2/ 2	17%
Porque(12%)	0%	0/ 2	5%	1/19	50%	1/ 1	9%
Que+NP(5%)	0%	0/ 0	5%	1/20	0%	0/ 0	5%
TOTAL	94%	51/54	9%	11/117	58%	7/12	38%

Tabela 12: Frequência de V-S a partir do cruzamento de tipo de verbo com tipo de elemento Q-.

* os números entre parênteses equivalem aos percentuais originais de frequência.

É interessante notar que, nas sentenças onde a cópula acontece, é quase que categórica a ocorrência de ordem Verbo-Sujeito. Veja, principalmente, o caso do qual, pois há apenas uma construção onde o verbo não é anteposto ao sujeito e é precisamente a que ocorre com um verbo transitivo - todas as demais apresentam ordem V-S e o verbo é uma cópula. Todos os demais casos indicam claramente que, mais uma vez, o tipo de verbo afeta diretamente a ordenação dos constituintes nas interrogativas. Observe também o caso do elemento que, que muito nos intrigou anteriormente, a tabela acima mostra que, em todas as sentenças cujo verbo é transitivo, só encontramos a ordem S-V e, as ocorrências de V-S se limitam aos casos de verbos reflexivos.

1.5- Os demais grupos de fatores

Agora discutirei brevemente os grupos de fatores que foram considerados explicativamente fracos pelo cálculo probabilístico, sendo, portanto, excluídos do conjunto de variáveis importantes para a compreensão desse fenômeno.

Os dois primeiros grupos de fatores - tipo de voz verbal e tipo de construção verbal - procuraram verificar se características internas do verbo poderiam beneficiar uma determinada estruturação frasal. Os resultados são os seguintes:

TIPO DE VOZ	FREQUÊNCIA V-S		FREQUÊNCIA S-V	
ativa	36%	65/180	64%	115/180
passiva	71%	5/ 7	29%	2/ 7

Tabela 13: Frequência de V-S e S-V segundo o tipo de voz verbal.

Os números acima mostram que as construções passivas apresentam uma tendência a ocorrerem com a ordem Verbo-Sujeito - 71% dos dados -, porém, é necessário ressaltar que essa célula é bem pequena para ser considerada representativa. Vejamos se com os demais corpora esse mesmo comportamento das passivas se verifica para, só então, podermos levantar algumas hipóteses.

TIPO DE CONSTRUÇÃO	FREQUÊNCIA V-S		FREQUÊNCIA S-V	
simples	41%	62/153	59%	91/153
composto	24%	8/ 34	76%	26/ 34

Tabela 14: Frequência de V-S e S-V segundo o tipo de construção verbal.

O construção verbal simples apresenta um comportamento quase neutro em relação à ordem sujeito-verbo nas sentenças, apontando para uma ligeira preferência pela ordem S-V (59%) em detrimento da construção V-S (41%). Já, as orações com tempo verbal composto preferencialmente ocorrem com o sujeito anteposto ao verbo.

As duas tabelas acima mostram que características internas ao verbo parecem não afetar, diretamente, o posicionamento deste na sentença. Inclusive, quando fazemos o cruzamento desses dois grupos de fatores com a variável tipo de verbo, percebemos que o que realmente está afetando a ocorrência da ordem S-V ou V-S é exatamente o fato do verbo ser uma cópula, um transitivo ou um reflexivo inerente.

Com relação à próxima variável - expletivo "é que" - Duarte (1992) observa que o aparecimento da ordem Sujeito-Verbo coincide com o surgimento desse expletivo em 1882. Vejamos como, atualmente, se comporta essa variável.

EXPLETIVO "É QUE"	FREQUÊNCIA V-S		FREQUÊNCIA S-V	
ausência de é que	46%	51/112	54%	61/112
presença de é que	33%	14/ 42	67%	28/ 42
presença de que	15%	5/ 33	85%	28/ 33

Tabela 15: Frequência de V-S e S-V conforme a presença ou ausência do expletivo "é que".

O que a tabela 15 mostra é que, no PB atual, o uso ou não da partícula expletiva "é que" não tem afetado a estruturação das sentenças interrogativas. Esse resultado confirma a afirmação feita por Duarte no texto acima, pois a autora realmente relaciona o surgimento da ordem S-V como o uso desse expletivo, porém isso se deu em 1882. Os dados dela também mostram que, no corpus sincrônico, tal condicionamento não se verifica.

A hipótese que se procurou testar com a consideração da variável que analisava o tipo de oração de onde o elemento Q- se deslocou é que, nas sentenças subordinadas, a ordem S-V deveria ser mais facilmente encontrada. Lembre-se que, como já foi dito anteriormente, Âmbar(1988) coloca que, para o PE, nesse tipo de construção subordinada não haveria nenhum tipo de restrição para a ocorrência da ordem S-V - exceção para o caso de que. Os números, no entanto, mostram um comportamento diferente do esperado no PB.

TIPO DE ORAÇÃO	FREQUÊNCIA V-S	FREQUÊNCIA S-V
subordinada	63% 12/ 19	37% 7/ 19
independente	35% 58/168	65% 110/168

Tabela 16: Frequência de V-S e S-V segundo o estatuto da oração.

Observe que os resultados acima mostram que, nas orações subordinadas, a ordem preferencial é Verbo-Sujeito com uma percentagem de 63% e o inverso acontece com as orações independentes, pois a frequência maior equivale precisamente à ordem S-V (65%). No entanto, o cálculo do peso relativo desse fator em comparação com os demais concluiu por dizer que essa é uma variável explicativamente fraca, sendo que, inclusive, o cruzamento com os demais fatores demonstrou que esta variável muitas vezes se submetia à força de seus co-ocorrentes a ponto de ter, em alguns casos, sua hierarquia interna modificada.

Resta um último fator a ser considerado e é o que analisa se a dimensão do sujeito pode afetar a estruturação das sentenças interrogativas. O grupo foi reorganizado através de uma série de amalgamações. Ou seja, um fator agrupava os sujeitos que apresentavam uma ou duas sílabas, um outro reunia os que tinha de três a seis sílabas e o último os sujeitos de sete ou mais.

DIMENSÃO DO SUJEITO	FREQUÊNCIA V-S	FREQUÊNCIA S-V
1/2 sílabas	9% 9/ 97	91% 88/97
3/4/5/6 sílabas	54% 29/ 54	46% 25/54
7/8/9/+ sílabas	89% 32/ 36	11% 4/36

Tabela 17: Frequência de V-S e S-V segundo a dimensão do sujeito.

A tabela acima mostra um comportamento bem regular dessa variável em relação à ordem Sujeito-Verbo, pois a medida que o número de sílabas do sujeito aumenta cresce igualmente a frequência com que o mesmo aparece em posição pós-verbal. Sendo assim, parece-me importante questionar por que esse fator não foi considerado relevante para o fenômeno que está sendo estudado. Vejamos o que nos diz o cruzamento dessa variável com o grupo de fatores que analisa o tipo de sujeito.

DIMENSÃO DO SUJEITO	T I P O D E		S U J E I T O			
	Expressão-R (84%)		Pronomes (4%)		Total	
1/2 sílabas(9%)*	83%	5/ 6	4%	4/91	9%	9/97
3/4/5/6 síl.(54%)	73%	29/40	0%	0/14	54%	29/54
7/8/9/+ síl.(89%)	97%	32/33	0%	0/ 3	89%	32/36
TOTAL	84%	66/79	4%	4/108	37%	70/187

Tabela 18: Frequência de V-S a partir do cruzamento de tipo do sujeito e dimensão do sujeito.

* os números entre parênteses representam os percentuais originais de frequência.

A tabela 18 mostra que o que está favorecendo a ordem V-S é exatamente o tipo do sujeito independentemente de sua dimensão, pois a expressão-referencial apresenta um alto índice de pós-posição em relação ao verbo qualquer que seja o seu tamanho. Diametralmente oposta é a situação dos sujeitos pronominais, uma vez que, nesses casos, é praticamente categórica a ocorrência da ordem S-V. Portanto, esse cruzamento de fatores mostra que é o tipo de sujeito que realmente está influenciando a estruturação nas construções interrogativas.

2. DADOS DE LÍNGUA ESCRITA

Esse corpus contém 214 sentenças e foi obtido a partir de diversos textos escritos, tais como, peças teatrais, romances e redações de vestibular da UNICAMP. Todos esses textos foram escritos depois de 1980.

Nesse corpus, a percentagem de ordem Verbo-Sujeito foi ainda menor do que nos dados do Projeto NURC/SP, ocorrendo em apenas 21% das orações. Porém, ao observarmos o valor do peso relativo dessa ordem quando as variáveis relevantes são consideradas obtemos o índice de .03. Parece estranho que justamente os dados de língua escrita apresentem um índice tão baixo de inversão V-S, verifiquemos, portanto, o que os resultados estatísticos nos dizem.

As sentenças desse corpus receberam o mesmo tratamento estatístico que foi apresentado anteriormente, salvo uma ou outra diferença que será mostrada oportunamente, os grupos de fatores que foram considerados acima - inclusive com, praticamente as mesmas amalgamações - serão utilizados para o tratamento quantitativo que ora descreverei.

O índice de inversão verbo-sujeito nas sentenças de língua escrita foi ainda menor do que no corpus do PROJETO NURC/SP

e isto pode ser um indício que as gravações apresentam uma fala muito mais monitorada - portanto, formal - do que o próprio texto escrito. Porém, é necessário verificar-se precisamente em que contextos a inversão está acontecendo na escrita para poder, então, traçar um paralelo entre essas duas situações de uso da língua. Sendo assim, gostaria de dizer que o cálculo probabilístico selecionou as seguintes variáveis como relevantes no estabelecimento da ordem relativa entre sujeito e verbo: tipo de verbo, tipo de sujeito, tipo de elemento Q- e estatuto da oração, respectivamente. Vamos agora aos números:

2.1- Tipo de verbo

Esse grupo de fatores teve que sofrer algumas alterações em relação ao que foi considerado no corpus anterior. A seguir apresentarei, na tabela 19, os resultados de frequência de V-S e S-V antes de fazer qualquer modificação no arranjo interno do grupo.

TIPO DE VERBO	FREQUÊNCIA V-S		FREQUÊNCIA S-V	
cópula	78%	35/ 45	22%	10/ 45
transitivo + referencial	6%	8/143	94%	135/143
transitivo - referencial	0%	0/ 15	100%	15/ 15
reflexivo	0%	0/ 5	100%	5/ 5
intransitivo	25%	1/ 4	75%	3/ 4
ergativos	0%	0/ 2	100%	2/ 2
TOTAL	21%	44/214	79%	170/214

Tabela 19: frequência de V-S e S-V segundo o tipo de verbo.

Dentre as variáveis analisadas acima, somente a cópula se mostrou favorecedora da ordem Verbo-Sujeito com uma percentagem de 78% de ocorrência. Todos os demais fatores se apresentaram inibidores dessa ordem, inclusive, em alguns casos, tendo a aplicação categórica da anteposição do sujeito em relação ao verbo - transitivo [- referencial], reflexivo inerente e ergativo.

Os três knockouts que aconteram acima foram resolvidos da seguinte forma: (1) amalgamei as sentenças com transitivo [- referencial] com as de transitivo [+ referencial], adotando, assim, o mesmo procedimento usado na análise anterior; (2) desconsidere localmente as sentenças cujo verbo ou era um reflexivo inerente, ou um intransitivo ou um ergativo. As orações

com verbo intransitivo, apesar de não apresentarei knockout, foram eliminadas porque a célula era muito pequena para ser considerada significativa. Portanto, após essas modificações, a frequência e o peso relativo de V-S ficou da seguinte forma:

TIPO DE VERBO	FREQÜÊNCIA V-S	PESO RELATIVO
cópula	78% 35/ 45	.99
transitivos	5% 8/158	.22
Total	21% 43/203	79% 160/203

Tabela 20: Frequência e peso relativo de V-S segundo o tipo de verbo.

Os números acima indicam que a cópula é um fator altamente favorecedor da ordem Verbo-Sujeito, sendo que esta ocorreu em 78% dos dados. Além disso, quando é feita a análise probabilística que considera todas as variáveis, o peso relativo desse fator aumenta mais ainda, chegando a atingir o índice de .99. Com os verbos transitivos, o inverso se dá, pois a ordem preferencial é Sujeito-Verbo (95% dos casos). No entanto, probabilisticamente falando, a chance dessa ordem ocorrer com os verbos transitivos é de .22.

2.2- Tipo de sujeito

O segundo fator considerado explicativamente forte é o tipo de sujeito. A seguir, apresento os valores obtidos para a frequência e o peso relativo de V-S conforme os diversos sub-fatores verificados nesse grupo (aqui foram feitas as mesmas alterações internas que aconteceram na análise do corpus do NURC/SP).

TIPO DE SUJEITO	FREQÜÊNCIA V-S	PESO RELATIVO
Expressão-r	66% 31/ 47	.96
Pronomes	8% 13/167	.29

Tabela 21: Frequência e peso relativo de V-S segundo o tipo de sujeito.

Os resultados acima mostram que, quando o sujeito é uma expressão-referencial, este aparece em posição pós-verbal em 66% dos casos e a probabilidade geral dele ocorrer chega a .96. Por outro lado, o sujeito pronominal preferencialmente acontece antes

do verbo (92% dos dados, com peso relativo de .29). Vejamos, agora, como fica o quadro quando consideramos simultaneamente a influência dos fatores tipo de sujeito e tipo de verbo.

TIPO DE VERBO	T I P O D E		S U J E I T O		Total	
	Pronomes (8%)		Expressão-r (66%)			
Cópula(78%)*	52%	11/ 21	100%	24/24	78%	35/ 45
Transitivos (5%)	1%	2/137	29%	6/21	5%	8/158
TOTAL	8%	13/158	67%	30/45	21%	43/203

Tabela 22: Frequência de V-S a partir do cruzamento de tipo de sujeito com tipo de verbo.

* os números entre parênteses equivalem aos percentuais originais de frequência.

A tabela acima indica que, quando o sujeito é pronominal e o verbo uma cópula, em 52% dos casos a ordem obtida era V-S. O pronome é um fator inibidor dessa ordem, porém a cópula é altamente favorecedora e, quando ambos atuam simultaneamente, o verbo parece estar influenciando mais, pois a percentagem obtida com o cruzamento está mais próxima da conseguida anteriormente pela cópula, quando esses fatores eram considerados em separado. Observe, também, o comportamento das sentenças onde o sujeito é uma expressão-referencial (fator favorecedor da ordem V-S) e o verbo é transitivo (fator inibidor da ordem V-S), nesses casos o tipo do verbo, também, parece estar atuando mais fortemente, uma vez que 29% (índice obtido com o cruzamento dos dois fatores) está mais próximo de 5% (percentagem anterior conseguida pelos verbos transitivos) do que de 67% (percentagem de expressões-r antepostas ao verbo). Em resumo, ambos os fatores influenciam fortemente a estruturação das sentenças interrogativas com elemento Q-, no entanto, quando forças opostas se confrontam (por exemplo, o verbo favorecendo a ordem V-S e o sujeito não), a tabela 21 acima indica que o tipo de verbo escolhido pesará mais na ordenação desses constituintes do que o tipo de sujeito, pelo menos no que se refere a este corpus.

2.3- Tipo de elemento Q-

Esse grupo de fatores recebeu a mesma reestruturação interna feita anteriormente, portanto, apresento, a seguir, as percentagens e os pesos relativos de V-S para essa variável.

TIPO DE ELEMENTO Q-	FREQUÊNCIA V-S			PESO RELATIVO
Qual	88%	7/	8	.96
Quem	57%	4/	7	.85
Como	41%	14/	34	.89
Que + NP	38%	5/	13	.56
Que	25%	1/	4	.86
O que	13%	10/	78	.69
Porque/quanto/quando/ onde	4%	3/	70	.07

Tabela 23: Frequência e peso relativo de V-S conforme o tipo de elemento Q-.

A tabela acima mostra que o qual é um grande favorecedor da ordem V-S, apresentando uma percentagem de 88% dos dados e o peso relativo desse sub-fator é de .96. Os elementos quem e qual parecem ter um comportamento quase neutro em relação a esse fenômeno, obtendo, respectivamente os seguintes índices: 57% e 41% de frequência. Porém, quando observamos a resultado probabilístico, verificamos que este coloca esses sub-fatores como favorecedores da anteposição do verbo (.85 e .89, respectivamente). Os demais fatores são tidos como inibidores da ordem V-S, segundo as percentagens obtidas por cada um deles. No entanto, probabilisticamente, o único sub-fator que realmente favorece a ordem S-V é o que agrupa os seguintes elementos Q-: porque, quanto, quando e onde. Vejamos agora o cruzamento desse grupo com as variáveis que consideram o tipo de verbo e o tipo de sujeito.

TIPO DE ELEMENTO Q-	T I P O D E		V E R B O			Total	
	Cópula (78%)		Transitivo (5%)				
Qual (88%)*	100%	7/ 7	0%	0/	1	88%	7/ 8
Quem (57%)	80%	4/ 5	0%	0/	1	57%	4/ 6
Como (46%)	90%	9/10	22%	4/	18	46%	13/ 28
Que + NP (38%)	63%	5/ 8	0%	0/	5	38%	5/ 13
Que (25%)	100%	1/ 1	0%	0/	2	33%	1/ 3
O que (13%)	86%	6/ 7	6%	4/	71	13%	10/ 78
Porque... (4%)	43%	3/ 7	0%	0/	60	4%	3/ 67
Total	78%	35/45	5%	8/158		21%	43/203

Tabela 24: Frequência de V-S a partir do cruzamento de tipo de verbo com tipo de elemento Q-.

* os números entre parênteses representam os percentuais originais de frequência.

Os resultados acima mostram que as construções interrogativas são fortemente influenciadas pelo tipo de verbo, mais até do que pelo tipo de elemento Q-. Veja que, quando o verbo

era uma cópula, obteve-se altos índices de ordem V-S, exceção feita ao casos dos elementos Q- porque, quanto, quando e onde, mas, mesmo nesse caso, teve-se a anteposição do verbo em 43% dos dados. Inclusive, quando ambos os fatores eram favorecedores da ordem V-S (caso em que o verbo era uma cópula e o elemento Q- era qual), a percentagem chegou a 100% das ocorrências. Note, também, que nos casos que a tabela 23 apontava para um comportamento basicamente neutro em relação à ordem quando se considerava os valores percentuais e, por outro lado, o estudo probabilístico dizia que eram fatores favorecedores da ordem V-S (casos de quem, como e que, principalmente), esse resultado, aparentemente, estranho parece ser explicado na tabela 24. Pois, tomando o exemplo do como, temos que 90% das sentenças cujo verbo era uma cópula obtemos a ordem V-S, em contra partida, quando o verbo é transitivo, essa ordem é conseguida em apenas 22% dos casos. Ou seja, na verdade é o tipo de verbo que mais fortemente está influenciando a estruturação das interrogativas com elemento Q- e, como o número de sentenças com verbos transitivos é quase quatro vezes maior do que o com a cópula, a probabilidade do verbo aparecer anteposto ao sujeito nas orações com como se torna bem mais significativa do que a percentagem simples indicava. Vamos agora verificar a influência do tipo de sujeito.

TIPO DE ELEMENTO Q-	T I P O D E			S U J E I T O					
	Pronomes (8%)			Expressão-r (66%)			Total		
Qual (88%)*	0%	0/	1	100%	7/	7	88%	7/	8
Quem (57%)	25%	1/	4	100%	3/	3	57%	4/	7
Como (41%)	14%	3/	22	92%	11/	12	41%	14/	34
Que + NP(38%)	38%	5/	13	0%	0/	0	38%	5/	13
Que (25%)	25%	1/	4	0%	0/	0	25%	1/	4
O que (13%)	5%	3/	62	44%	7/	16	13%	10/	78
Porque..(4%)	0%	0/	61	33%	3/	9	4%	3/	70
Total	8%	13/	167	66%	31/	47	21%	44/	214

Tabela 25: Frequência de V-S a partir do cruzamento de tipo de sujeito com tipo de elemento Q-.

* os números entre parênteses representam os percentuais originais de frequência.

Observe que a tabela acima mostra uma oposição bem clara: os valores mais altos de ordem V-S são encontrados quando o sujeito é uma expressão-referencial, inclusive mostrando dois casos (qual e quem) de aplicação categórica da regra de inversão. Por outro lado, o sujeito pronominal apresenta percentagens baixas de anteposição do verbo. Além disso, os resultados também mostram que essa variável - tipo de elemento Q- - está subordinada à variável tipo de sujeito, pois um sub-fator - porque, quanto, quando e onde

- que é altamente favorecedor da ordem S-V apresenta um índice relativamente alto (33%) de anteposição do verbo quando o sujeito é uma expressão-referencial que, por sua vez, é um fator altamente favorecedor da ordem V-S.

2.4- Estatuto da oração

O quarto grupo de fatores considerado relevante para o fenômeno em questão é o que verifica o tipo de oração - independente ou subordinada - de onde o elemento Q- foi extraído. A hipótese é de, nas orações subordinadas, existem menos restrições de ocorrência da ordem S-V (conforme Âmbar 1988). Vamos aos primeiros resultados:

ESTATUTO DA ORAÇÃO	FREQÜÊNCIA V-S	PESO RELATIVO
Subordinada	11% 1/ 9	.01
Independente	21% 43/205	.55

Tabela 26: Frequência e peso relativo de V-S segundo o estatuto da oração.

Em princípio, parece-me surpreendente que esse fator tenha sido considerado explicativamente forte em relação à ordem Verbo-Sujeito nas sentenças interrogativas com elemento Q-, pois creio que é apenas o comportamento desses dois constituintes nas orações independentes que afetou o resultado geral. Veja que a célula que reúne as orações subordinadas contém somente 9 ocorrências. Apesar disso, podemos ver que, nesse corpus, os números confirmam a hipótese inicial, pois a ordem V-S ocorre em uma só sentença subordinada e a probabilidade geral dela ocorrer nessa situação é de .01. Por outro lado, nas orações independentes, essa probabilidade sobe para .55, sendo que a percentagem de anteposição do verbo nesse corpus foi de 21% . Vamos agora verificar o comportamento dessa variável, cruzando-a com os outros fatores explicativamente fortes.

ESTATUTO DA ORAÇÃO	T I P O D E		V E R B O		Total
	Cópula (78%)		Transitivo (5%)		
Subordinada (11%)	33%	1/ 3	0%	0/ 6	11% 1/ 9
Independente (21%)	81%	34/42	5%	8/152	22% 42/194
Total	78%	35/45	5%	8/158	21% 43/203

Tabela 27: Frequência de V-S a partir do cruzamento de tipo de verbo com estatuto da oração.

* os números entre parênteses representam os percentuais originais de frequência.

A tabela 27 mostra que o único caso de V-S em oração subordinada ocorre quando o verbo é uma cópula, no entanto, mesmo sendo este um fator altamente favorecedor dessa ordem, essa só aconteceu em uma das três sentenças nessa mesma situação. Esse resultado parece indicar que, nesse caso, o estatuto da oração teve mais influência na estruturação da sentença - veja que o índice obtido no cruzamento está mais próximo da frequência original da oração subordinada do que o valor conseguido pela cópula anteriormente. Já as orações independentes se mostram mais susceptíveis à influência dos outros fatores, tais como o tipo de verbo.

ESTATUTO DA ORAÇÃO	T I P O D E		S U J E I T O		Total
	Pronomes (8%)		Expressão-r (66%)		
Subordinada (11%)*	0%	0/ 8	100%	1/ 1	11% 1/ 9
Independente(21%)	8%	13/159	65%	30/46	21% 43/205
Total	8%	13/167	66%	31/47	21% 44/214

Tabela 28: Frequência de V-S a partir do cruzamento de tipo de sujeito com estatuto da oração.

* os números entre parênteses representam os percentuais originais de frequência.

Os resultados acima indicam que a variável tipo de sujeito não é afetada diretamente pelo estatuto da oração, pois os valores obtidos com o cruzamento desses dois fatores se aproximam muito mais dos percentuais originais conseguidos por essa variável.

TIPO DE ELEMENTO Q-	E S T A T U T O D A O R A Ç Ã O					
	Subordinada (11%)		Independente (21%)		Total	
Qual (88%)*	100%	1/ 1	86%	6/ 7	88%	7/ 8
Quem (57%)	0%	0/ 1	67%	4/ 6	57%	4/ 7
Como (41%)	0%	0/ 0	41%	14/34	41%	14/34
Que+NP(38%)	0%	0/ 0	38%	5/13	38%	5/13
Que (25%)	0%	0/ 0	25%	1/ 4	25%	1/ 4
O que (13%)	0%	0/ 4	14%	10/74	13%	10/78
Porque(4%)	0%	0/ 3	4%	3/67	4%	3/70
Total	11%	1/ 9	21%	43/205	21%	44/214

Tabela 29: Frequência de V-S a partir do cruzamento de estatuto da oração com tipo de elemento Q-.

* os números entre parênteses representam os percentuais originais de frequência.

Os valores da tabela 28 indicam que o estatuto da oração não influencia tanto a estruturação das sentenças interrogativas quanto o faz o tipo de elemento Q-. Isso pode ser observado comparando-se os resultados obtidos no cruzamento que, em quase todos os casos, correspondem aos percentuais originais de frequência conseguidos pela variável tipo de elemento Q-.

2.5- Os demais grupos de fatores

Resta-nos agora abordar o comportamento das demais variáveis em relação a ordem Sujeito-Verbo no corpus de língua escrita. São elas: tipo de voz, tipo de construção verbal, expletivo "é que", tipo de corpus e dimensão do sujeito. Com exceção do grupo de fatores que considera o tipo de corpus, todas as demais variáveis também foram excluídas, no corpus do Projeto NURC/SP, do conjunto de fatores importantes para a explicação do fenômeno em questão. Portanto, primeiramente apresentarei os resultados obtidos com o grupo de fatores que codificava o tipo de corpus para verificar o porquê de não ter sido ele considerado relevante para a explicação da ordem relativa entre sujeito e verbo. Em seguida, abordarei, rapidamente, as demais variáveis.

TIPO DE CORPUS	FREQUÊNCIA V-S			FREQUÊNCIA S-V		
Dissertação	67%	6/	9	33%	3/	9
Narração	60%	3/	5	40%	2/	5
Romance	37%	15/	41	63%	26/	41
Peça teatral	13%	20/	157	87%	137/	157

Tabela 30: Frequência de V-S e S-V segundo o tipo de corpus.

Apesar desse grupo não ter sido considerado relevante para a explicação do fenômeno da ordenação entre sujeito e verbo nas sentenças interrogativas, ele mostra uma distribuição interessante entre as diversas variantes estudadas acima. Primeiramente gostaria de explicar que, na modalidade de redação de vestibular carta, só foi encontrada uma ocorrência de interrogativa com elemento Q-, portanto essa sentença foi eliminada localmente dessa análise.

Podemos observar, na tabela 30, que os textos de redação de vestibular apresentam valores bem próximos, sendo que, em ambas as modalidades, os escritores deram preferência para a ordem V-S. Já as outras duas variantes apontam para a outra direção, inclusive, vemos que as peças teatrais mostram um alto índice de ordem S-V. Esse resultado é bastante interessante, pois mostra que, numa situação de pressão pela norma (prova de vestibular), aumenta significativamente a ocorrência da inversão verbo-sujeito. Apesar de pequenas, as células correspondem a uma expectativa minha, pois, mesmo antes de coletar dados nesse contexto de uso da língua escrita, eu esperava encontrar um maior uso de ordem V-S nessa situação. Essa hipótese parte do fato de que, num contexto de uma prova escrita de vestibular, os usuários de uma língua tendem a prestar mais atenção à estrutura geral dessa e, mesmo inconscientemente, procuram respeitar as regras aprendidas durante todos os anos na escola. Inclusive, é muito comum ocorrerem erros ocasionados por hipercorreção.

Já os romances, ao apresentarem um índice de 37% de inversão Verbo-Sujeito, mostram-se mais próximos da tendência geral que é estruturar as interrogativas, preferencialmente, com o sujeito anteposto ao verbo. Mais ainda do que estes, as peças teatrais dão preferência para essa ordenação (apenas 13% dos dados apresentam a ordem V-S). Vários autores já apontaram para o fato de que as peças teatrais tendiam a usar uma linguagem mais próxima da falada por aquela comunidade e os valores obtidos na tabela 30 confirmam essa hipótese ao mostrar altas percentagens de S-V nesse tipo de texto escrito. Sendo assim, apesar desse grupo de fatores não ter sido selecionado como relevante para explicar o fenômeno da ordem nas sentenças interrogativas com elemento Q-, creio que ele mostra que o contexto de escrita de um texto influencia, mesmo que inconscientemente, o usuário da língua no momento de sua

construção.

Com relação às demais variáveis consideradas nessa análise, é interessante comentar que os grupo que verificava o tipo de voz verbal não pôde ser incluído dentre os fatores estudados, pois todas as sentenças possuíam verbo na voz ativa. Quando apenas um dos fatores ocorre num grupo caracteriza-se o SINGLETON e, a menos que se elimine esse problema, por exemplo, inventando um dado, essa variável não poderá ser usada no cálculo probabilístico. Não achei que seria necessário criar um dado para eliminar esse SINGLETON, dessa forma optei por não incluir esse grupo de fatores no estudo probabilístico. No entanto, não deixa de ser interessante notar que, num total de 214 ocorrências, não apareceu nenhuma oração onde o verbo estivesse na voz passiva.

Agora apresentarei os valores obtidos com o fator expletivo "é que". Esse grupo não foi selecionado como explicativamente relevante para o fenômeno, porém mostrou uma distribuição interna diferente da conseguida com os dados do NURC/SP.

EXPLETIVO "É QUE"	FREQÜÊNCIA V-S	FREQÜÊNCIA S-V
presença de que	50% 1/ 2	50% 1/ 2
ausência de é que	25% 40/162	75% 122/162
presença de é que	6% 3/ 50	94% 47/ 50

Tabela 31: Freqüência de V-S e S-V segundo o uso do expletivo "é que".

Tirando os resultados obtidos com o sub-fator presença de que, que possui uma célula muito pequena, temos que, na grande maioria das sentenças onde o expletivo "é que" aparece, a ordem resultante é Sujeito-Verbo. Inclusive, mesmo quando esse expletivo está ausente, a percentagem de freqüência cresce, mas não o suficiente para tornar a ordem V-S preferencial nesse contexto. Ou seja, a presença do expletivo pode até favorecer fortemente a não anteposição do verbo em relação ao sujeito, porém, quando este se encontra ausente, a freqüência de V-S aumenta mas não a ponto de se mostrar significativa.

Os valores conseguidos com a variável tipo de construção verbal não trouxeram nada de novo para a análise, apenas mostram índices ainda menores de anteposição do verbo (7% nas estruturas com verbo composto e 27% nas com verbo simples). Mas essa diferença pode ser explicada pelo fato da ordem V-S ocorrer com menor freqüência no corpus de língua escrita (21%) do que no do Projeto NURC/SP (37%).

Com relação ao último grupo de fatores - dimensão do sujeito - uma observação deve ser feita antes de se verificar os resultados obtidos. Foi necessário inverter uma sentença, pois, quando o sujeito possuía 7 ou mais sílabas, foi categórico o uso da ordem V-S e, não querendo eliminar esse fator da análise, optei por essa solução. A sentença foi a seguinte: "O que aquele menino de azul marinho comprou?". Agora, vejamos os resultados.

DIMENSÃO DO SUJEITO	FREQUÊNCIA V-S	FREQUÊNCIA S-V
1/2 sílabas	10% 18/176	90% 158/176
3/4/5/6 sílabas	59% 16/ 27	41% 11/ 27
7/8/9/+ sílabas	91% 10/ 11	9% 1/ 11

Tabela 32: Frequência de V-S e S-V segundo a dimensão do sujeito.

A tabela 32 acima mostra resultados bem próximos dos obtidos no mesmo grupo de fatores para os dados do Projeto NURC/SP. Ou seja, apesar de apresentar uma distribuição muito equilibrada entre os diversos sub-fatores, essa variável não foi selecionada como explicativamente forte para o fenômeno, pois o que realmente afeta a ordem relativa entre sujeito e verbo é exatamente o tipo de sujeito e não a sua dimensão. Além disso, quando o sujeito é pronominal, ele normalmente é bem curto, de uma ou duas sílabas. Sendo assim, essa variável só deve atuar internamente dentro do sub-fator expressão-referencial.

3. DADOS DE LÍNGUA FALADA

O corpus de língua falada se compõe de 307 sentenças que foram obtidas em programas de entrevista na televisão, sala de aula de adultos e de crianças e, também, fala espontânea. Nesse contexto de uso da língua, a ordem V-S ocorre em 43% dos dados e a probabilidade dessa ordem acontecer é de .43 quando todos as variáveis relevantes para esse fenômeno estão atuando. Os fatores selecionados como explicativamente fortes foram, respectivamente, tipo de verbo, dimensão do sujeito, estatuto da oração e tipo de sujeito. Apresentarei, a seguir, os resultados obtidos em cada um desses grupos de fatores.

3.1- Tipo de verbo

A análise dessa variável mostrou os seguintes valores para cada tipo de verbo considerado.

TIPO DE VERBO	FREQUÊNCIA V-S	FREQUÊNCIA S-V
Cópula	89% 110/123	11% 13/123
Ergativos	73% 8/ 11	27% 3/ 11
Intransitivos	33% 1/ 3	67% 2/ 3
Reflexivos	30% 3/ 10	70% 7/ 10
Transitivos	5% 8/160	95% 152/160

Tabela 33: Frequência de V-S e S-V conforme o tipo de verbo.

A tabela acima mostra que, nessa análise, optei por fazer menos amalgamações, mantendo o grupo de fatores praticamente igual ao original. A única alteração que fiz foi unir os dois tipos de verbos transitivos - [+ referenciais] e [- referenciais] - como eu já havia feito anteriormente com o intuito de eliminar knockout. Agora, é interessante observar que os verbos ergativos e intransitivos, que muitos autores consideravam como sendo bem semelhantes, nesse corpus, apresentam um comportamento oposto - os ergativos favorecendo a ordem V-S (73%) e os intransitivos a ordem S-V (67%). Na verdade, as sentenças com verbos intransitivos deveriam ser amalgamadas com as de verbos transitivos (semelhanças numéricas e lingüísticas), porém acabei decidindo mantê-lo separado para observar como ele se comporta em relação ao demais fatores, mesmo sendo essa uma célula muito pequena. Com relação aos demais verbos, podemos observar que a cópula continua favorecendo fortemente a anteposição do verbo, os transitivos, por outro lado, inibem-na e os reflexivos, nesse corpus, têm preferido as estruturas onde o sujeito aparece antes do verbo.

Vejamos, agora, o que o cálculo probabilístico pode nos dizer:

TIPO DE VERBO	FREQUÊNCIA V-S	PESO RELATIVO
Cópula	89% 110/123	.93
Ergativos	73% 8/ 11	.66
Intransitivos	33% 1/ 3	.73
Reflexivos	30% 3/ 10	.49
Transitivos	5% 8/160	.12

Tabela 34: Frequência e peso relativo de V-S segundo o tipo de verbo.

O resultado mais interessante da tabela acima é que os intransitivos, probabilisticamente falando, são mais favorecedores da ordem V-S do que os próprios ergativos, sendo que os valores percentuais indicavam na direção contrária. Dessa forma,

temos que o único tipo de verbo que realmente exige o sujeito antes dele numa interrogativa com elemento Q- é o transitivo. Além desses verbos, só os reflexivos inerentes mostram, na tabela acima, uma ligeira preferência por essa ordem (.49). Sendo assim, percebemos que, de um lado estão os verbos copulares, os ergativos e os intransitivos que apresentam uma alta probabilidade de ocorrerem antes do sujeito nesse contexto de interrogativa (respectivamente, .93, .66 e .73) e, do outro lado estão os verbos transitivos que tendem a acontecerem depois do sujeito (.12) e, quase no meio exato desses dois grupos, temos os reflexivos inerentes que se mostram sutilmente a favor da ordem S-V.

3.2- Dimensão do sujeito

É a primeira vez que essa variável é selecionada como explicativamente relevante para o fenômeno aqui analisado. Nos casos anteriores, o grupo que considerava o tipo de sujeito sempre foi escolhido como um dos dois fatores mais importantes para a questão da ordem. Nesse corpus, ele foi o quarto grupo a ser selecionado e, em seu lugar, apareceu a variável dimensão do sujeito. Vejamos os valores obtidos:

DIMENSÃO DO SUJEITO	FREQÜÊNCIA V-S	PESO RELATIVO
1/2 sílabas	12% 20/171	.27
3/4/5/6 sílabas	68% 52/ 76	.62
7/8/9/+ sílabas	95% 58/ 61	.89

Tabela 35: Frequência e peso relativo de V-S conforme a dimensão do sujeito.

A tabela 35 mostra que apenas os sujeitos de uma ou duas sílabas favorecem a não anteposição do verbo (ordem S-V em 88% dos casos). O sujeito tendo três ou mais sílabas já tende a acontecer depois do verbo e, se ele tiver mais de sete, a frequência dessa ordem cresce mais ainda. O estudo probabilístico confirma os valores percentuais, pois, segundo esse cálculo, quanto maior for o sujeito maior será a probabilidade de acontecer depois do verbo, pelos menos nos dados de língua falada.

TIPO DE VERBO	D I M E N S Ã O D O S U J E I T O					
	1/2 sílabas (12%)		3/4/5/6 síl. (68%)		7/8/9/+ síl. (95%)	
Cópula (89%)*	60%	15/ 25	96%	45/47	98%	50/51
Ergativos (73%)	50%	1/ 2	67%	2/ 3	83%	5/ 6
Intransitivos(33%)	33%	1/ 3	-		-	
Reflexivos (30%)	17%	1/ 6	33%	1/ 3	100%	1/ 1
Transitivos (5%)	1%	2/134	17%	4/23	67%	2/ 3
Total	12%	20/171	68%	52/76	95%	58/61

Tabela 36: Frequência de V-S a partir do cruzamento de dimensão do sujeito com tipo de verbo.

* os números entre parênteses correspondem aos percentuais originais de frequência.

Os valores acima indicam que o tipo de verbo influencia mais a sua posição em relação ao sujeito do que a dimensão deste último. No entanto, note que, quando o sujeito possui mais de sete sílabas, os percentuais de ordem V-S, com todos os tipos de verbo, são muito altos - 67% é o menor índice obtido e ocorre justamente com os verbos transitivos, que são os maiores inibidores dessa ordem. Observando-se o comportamento dos sujeitos de três a seis sílabas é que se percebe a atuação da variável tipo de verbo, pois quando este é favorecedor da ordem V-S, os sujeitos tendem a ocorrer depois dos verbos (veja os índices para os verbos copulares e os ergativos). Por outro lado, quando a situação inversa se dá, os sujeitos normalmente aparecem antepostos ao verbo (verifique os percentuais dos reflexivos e, principalmente, dos transitivos).

3.3- Estatuto da oração

Esse grupo de fatores também foi apontado como relevante para o fenômeno da inversão V-S nas interrogativas com elemento Q- quando da análise do corpus anterior - Dados de Língua Escrita.

ESTATUTO DA ORAÇÃO	FREQUÊNCIA V-S		PESO RELATIVO
Independente	43%	126/295	.52
Subordinada	33%	4/ 12	.13

Tabela 37: Frequência e peso relativo de V-S segundo o estatuto da oração.

Outra vez as orações subordinadas confirmam a hipótese inicial de que, nesse contexto, a ordem S-V é mais freqüente, pois os resultados acima mostram que a anteposição do verbo só ocorreu em 33% dos casos e a probabilidade disso acontecer é de .13. As orações independentes mantêm um comportamento basicamente neutro, sendo que em 43% dos dados a ordem utilizada foi V-S e, por sua vez, a probabilidade dela ocorrer é de .52.

TIPO DE VERBO	E S T A T U T O D A O R A Ç Ã O		
	Independente (43%)	Subordinada (33%)	Total
Cópula (89%)*	92% 108/117	33% 2/ 6	89% 110/123
Ergativos(73%)	70% 7/ 10	100% 1/ 1	73% 8/ 11
Intransitivos(33%)	33% 1/ 3	-	33% 1/ 3
Reflexivos (30%)	22% 2/ 9	100% 1/ 1	30% 3/ 10
Transitivos (5%)	5% 8/156	0% 0/ 4	5% 8/160
Total	43% 126/295	33% 4/12	42% 130/307

Tabela 38: Frequência de V-S a partir do cruzamento de estatuto da oração com tipo de verbo.

* os números entre parênteses representam os percentuais originais de freqüência.

A cópula é o único tipo de verbo que parece ser afetado pelo estatuto da oração, uma vez que, sendo ela fortemente influenciadora da ordem V-S, quando ocorre na sentença subordinada a ordem mais freqüente é justamente a oposta, ou seja, S-V. Esse resultado mostra a força que o estatuto da oração tem para influenciar a ordenação dos constituintes nas interrogativas, isto é, nesse tipo de construção a ocorrência da anteposição do verbo é fortemente inibida. Com as sentenças independentes, a ordem é mais afetada pelo tipo de verbo. Observe que os resultados obtidos com o cruzamento das variáveis se aproximam mais dos conseguidos com os fatores que codificam o tipo de verbo.

3.4- Tipo de sujeito

Esse grupo de fatores obteve os seguintes valores freqüenciais e probabilísticos:

TIPO DE SUJEITO	FREQÜÊNCIA V-S	PESO RELATIVO
Expressão-r	73% 116/159	.63
Pronomes	9% 14/148	.36

Tabela 39: Freqüência e peso relativo de V-S conforme o tipo de sujeito.

A tabela acima mostra que as expressões-r preferem nitidamente a ordem V-S, pois esta acontece em 73% desses casos. Uma vez que esse tipo de sujeito ocorre em mais da metade do corpus (159 de um total de 307 sentenças), esse resultado faz com que, nos dados de língua falada, encontremos o maior índice de probabilidade da ordem V-S ser a utilizada nas sentenças interrogativas (.43). No entanto, os números acima não nos trazem grandes novidades, pois, nos demais corpora, essa variável também foi selecionada como explicativamente forte para o fenômeno. Porém, parece-me importante entender o porquê desse grupo de fatores estar em quarta posição nesse corpus, sendo que nos outros, ele apareceu na primeira (NURC/SP) e na segunda (Dados de Língua Escrita). Portanto, apresentarei o cruzamento dessa variável com a que codifica a dimensão do sujeito.

DIMENSÃO DO SUJEITO	T I P O D E S U J E I T O		Total
	Pronomes (9%)	Expressão-r (73%)	
1/2 sílabas (12%)*	10% 14/144	22% 6/ 27	12% 20/171
3/4/5/6 síl.(68%)	0% 0/ 4	72% 52/ 72	68% 52/ 76
7/8/9/+ síl.(97%)	-	97% 58/ 60	97% 58/ 60
Total	9% 14/148	73% 116/159	42% 130/307

Tabela 40: Freqüência de V-S a partir do cruzamento de tipo de sujeito com dimensão do sujeito.

* os números entre parênteses representam os percentuais originais de freqüência.

A tabela acima mostra que é a dimensão do sujeito a variável mais forte a influenciar a anteposição do verbo, pois, quando o sujeito é pronominal, o seu número de sílabas é pequeno (normalmente 1 ou 2) e, na maioria dos casos, ele ocorre antes do verbo. Já quando o sujeito é uma expressão-referencial - que é a maioria nesse corpus - a sua dimensão varia bastante e isso influencia a posição deste na sentença. Veja que quando é um sujeito curto, ele, em geral, aparece antes do verbo. Sendo ele de tamanho médio - 3 a 6 sílabas - ele tende a ocorrer com a ordem V-S (72%). Por fim, se ele for longo - 7 sílabas ou mais -, em quase

todos os casos, ele também vai aparecer nessa posição. Portanto, pode-se perceber que a variável dimensão do sujeito foi colocada em posição hierarquicamente superior à variável tipo de sujeito porque a primeira, ao atuar basicamente dentro do fator expressão-referencial, que nesse corpus, ao contrário dos outros, representa a maioria dos casos de sujeito, possibilita uma distinção mais fina da atuação do sujeito no estabelecimento da ordem nas sentenças interrogativas. Ou seja, a variável dimensão do sujeito mostra que existem dois pólos opostos, um que fortemente favorece a ordem S-V (sujeitos de 1 ou 2 sílabas) e outro que nitidamente prefere a ordem V-S (sujeitos de 7 ou mais sílabas). Entre esses pólos estão os casos onde o sujeito possui um tamanho intermediário (3 a 6 sílabas), estes também fazem uso preferencial da anteposição do verbo, porém observe que a probabilidade dessa ordem ocorrer nesse contexto já é bem menor (.62). Sendo assim, temos que o sujeito possui um papel altamente importante na estruturação das interrogativas com elemento Q-.

3.5- Os demais grupos de fatores

Os fatores que não foram apontados como explicativamente relevantes para a ordem nas estruturas com Q- desse corpus são: tipo de voz, tipo de construção verbal, expletivo "é que", tipo de elemento Q- e tipo de corpus. Como nas descrições anteriores, as duas primeiras variáveis não trouxeram praticamente nada de novo para análise. Uma ressalva apenas se faz necessária: mais uma vez houve uma baixíssima ocorrência de voz passiva nos dados coletados. Nesse corpus, de um total de 307 sentenças, somente sete apresentaram o verbo na passiva, sendo que, nesse caso, apenas duas orações estavam na ordem V-S.

O grupo de fatores expletivo "é que" apresentou a seguinte distribuição interna.

EXPLETIVO "É QUE"	FREQUÊNCIA V-S		FREQUÊNCIA S-V	
presença "é que"	49%	49/100	51%	51/100
ausência "é que"	48%	71/147	52%	76/147
presença "que"	17%	10/60	83%	50/ 60

Tabela 41: Frequência de V-S e S-V conforme o uso do expletivo "é que".

A tabela acima mostra que, realmente, a presença ou não do expletivo "é que" não está afetando diretamente a ordem relativa entre sujeito e verbo. Observe que os valores obtidos para V-S desses fatores são bem próximos (49% para a presença e 48% para a ausência). Já a presença de "que" aparece como sendo um fator que fortemente inibe essa ordem, apresentando um índice de apenas 17%

dos dados. Ora, o que se esperaria seria um comportamento similar entre a presença de "é que" e a presença de "que", pois creio que este último é uma versão do primeiro só que com o verbo ser elíptico. No entanto, não é isso que encontramos. Podemos concluir, portanto, que essa variável parece não estar influenciando a estruturação das sentenças interrogativas.

O próximo fator a ser descrito é o que codifica o tipo de elemento Q-. Essa mesma variável foi selecionada como importante para os outros dois corpora. Vejamos agora como esse corpus se comporta em relação a esse grupo de fatores.

TIPO DE ELEMENTO Q-	FREQUÊNCIA V-S	FREQUÊNCIA S-V
Qual	98% 46/47	2% 1/47
Como	57% 54/95	43% 41/95
Quem	38% 5/13	62% 8/13
O que	22% 12/54	78% 42/54
Que	19% 3/16	81% 13/13
Que + NP	14% 4/28	86% 24/28
Porque/quanto/quando/ onde	11% 6/54	89% 48/54

Tabela 42: Frequência de V-S e S-V conforme o tipo de elemento Q-.

Os resultados acima mostram que o elemento Q- qual favorece fortemente a ordem V-S, existindo apenas uma sentença com esse pronome que não utiliza essa ordem. Dentre as demais variantes, temos que o como ligeiramente prefere essa mesma ordem com uma percentagem de 57% dos casos. Os outros pronomes interrogativos aparecem, na grande maioria dos casos, com o verbo posposto ao sujeito. Apesar disso, essa variante não foi considerada explicativamente relevante para o fenômeno em questão.

O último fator que falta analisar é o que codifica o tipo de corpus e foram os seguintes os valores obtidos para essa variável.

TIPO DE CORPUS	FREQUÊNCIA V-S			FREQUÊNCIA S-V		
40 anos de TV	83%	5/	6	17%	1/	6
Cara a Cara	71%	5/	7	29%	2/	5
Escolinha do Prof. Raimundo	67%	2/	3	33%	1/	3
Sala aula/adulto	63%	10/	16	38%	6/	16
Jô Soares 11:30	50%	55/	110	50%	55/	110
Metrópolis	50%	1/	2	50%	1/	2
Roda Viva	39%	7/	18	61%	11/	18
Canal Livre	37%	37/	101	63%	64/	101
Fala espontânea	20%	2/	10	80%	8/	10
Sala aula/criança	18%	6/	33	82%	27/	33

Tabela 43: Frequência de V-S e S-V conforme o tipo de corpus.

A tabela acima mostra que, nos seguintes contextos, a ordem V-S foi a mais usada: 40 anos de TV, Cara a Cara, Escolinha do Prof. Raimundo e sala de aula de adulto. O Jô Soares 11:30 e o Metrôpolis apresentam um comportamento neutro em relação à ordem S-V e os demais tipos de corpus favorecem a não inversão do verbo em relação ao sujeito. No entanto, é necessário ressaltar que, em vários dos contextos acima, as células apresentavam um número muito pequeno de ocorrências.

Algumas observações parecem interessantes, tais como a ocorrência equilibrada de ambas as ordens S-V e V-S no programa do Jô Soares 11:30. Isso talvez se dê porque esse apresentador molda muito facilmente a sua fala ao tipo de entrevistado com quem ele vai conversar. Ou seja, se o entrevistado for uma pessoa aparentemente culta, ele tende a usar uma fala mais cuidada. Por outro lado, se é uma pessoa mais simples e despojada, a sua fala se torna também mais coloquial.

Uma outra situação peculiar é o comportamento oposto verificado nas duas situações de sala da aula. Na sala de aula do primário, a ordem S-V foi largamente usada (82% dos casos) enquanto que na sala de aula dos adultos o inverso se deu, com uma percentagem de 63% de ordem V-S. Talvez isso se explique porque, nesse último contexto, foram coletadas sentença durante uma aula na faculdade, situação essa bem próxima das Elocuções Formais do corpus do NURC/SP que também favoreciam essa ordem.

Ao coletar os dados, eu supunha que alguns tipos de programas, tais como o Roda Viva, utilizariam um número maior de construções V-S pela própria natureza do programa (o entrevistado ficava no centro de uma roda de jornalistas e estes dirigiam-lhe os mais diversos tipos de questões, sendo que o apresentador também colocava perguntas feitas pelo público). E, por outro lado, um programa do tipo do Canal Livre - transmitido num horário mais cedo (17:00) e com a participação ao vivo do público - teria uma

predominância de construções sem a anteposição do verbo. No entanto, o que se observa na tabela anterior é que ambos os programas apresentam índices bem próximos, favorecendo a ordem S-V. Sendo assim, a minha hipótese inicial não foi confirmada.

VI. CONCLUSÃO

Essa seção tem por objetivo reunir os aspectos convergentes e divergentes que a análise estatística de cada corpus levantou e, também, tentar responder às várias perguntas que foram colocadas ao longo de todo o texto. Começarei retomando a questão central levantada no início desse trabalho.

Quais os fatores, externos e internos, que mais influenciam a ordem relativa entre sujeito e verbo nas interrogativas com elemento Q-?

Antes de tentar responder a essa pergunta, gostaria de relembrar também as outras questões que foram colocadas no início do capítulo IV, pois todas elas estão diretamente relacionadas a essa primeira e serão respondidas juntamente.

- 1) Qual a incidência de ordem VS?
- 2) Se existe um condicionamento estilístico, que tipo de corpus apresenta uma maior produtividade de ordem VS?
- 3) O tipo de elemento Q- é relevante na ocorrência de VS?
- 4) O tipo de verbo e a hipótese da monoargumentalidade se confirmam nos dados?
- 5) A presença ou ausência do expletivo é que está afetando a ordem nas sentenças interrogativas?

Sem dúvida alguma as variáveis que mais diretamente afetam a estruturação das sentenças interrogativas são tipo de verbo e o tipo de sujeito. Esses fatores foram apontados em todos os corpora como explicativamente fortes para o fenômeno em questão.

Com relação ao tipos de verbos, temos que a classe dos transitivos [- e + referenciais] são inibidores da ordem V-S, ocorrendo o sujeito, quase sempre, antes. Por outro lado, os verbos copulares, ergativos e intransitivos tendem a aparecer antes do sujeito. Como já foi apontado anteriormente - capítulo V -, esses resultados confirmam a hipótese da monoargumentalidade. E os reflexivos inerentes apresentam um comportamento basicamente neutro, sendo que a sua posição na sentença parece ser mais relacionada ao tipo de sujeito que nela ocorre.

Em se tratando do tipo de sujeito, observamos que os pronominais tendem a aparecer antes do verbo, sendo que os pronomes de tratamento basicamente ocorrem, todos, na ordem S-V. Ao contrário desses, os sujeitos que são uma expressão-referencial

simples ou um NP complexo favorecem fortemente a anteposição do verbo. Esse resultado é confirmado quando analisamos a dimensão do sujeito, pois vemos que, quanto maior o seu tamanho, maior é também a probabilidade dele ocorrer depois do verbo.

Outro fator que foi apontado várias vezes como influenciador da estruturação das interrogativas é o estatuto da oração. Aqui temos um resultado interessante. Enquanto os dados de língua escrita e falada confirmam a hipótese de que, nas sentenças encaixadas, existem menos restrições para ocorrência da ordem S-V, por outro lado, os resultados do Projeto NURC/SP apontam que, nesse contexto, a ordem V-S é a mais frequente (63%). Porém, a tabela abaixo pode explicar esse paradoxo:

TIPO DE SUJEITO	E S T A T U T O		D A O R A Ç Ã O		Total
	Independente (35%)		Subordinada (63%)		
Expressão-r(84%)*	82%	55/ 67	92%	11/12	84% 66/79
Pronomes (4%)	3%	3/101	14%	1/ 7	4% 4/108
Total	35%	58/168	63%	12/19	37% 70/187

Tabela 44: Frequência de V-S a partir do cruzamento de estatuto da oração com tipo de sujeito, no corpus do NURC/SP.

* os números entre parênteses representam os percentuais originais de frequência.

A tabela 44 mostra que o alto índice de ordem V-S obtido nas sentenças subordinadas se deve à natureza do sujeito que ali ocorre, pois em 92% dos casos (11 entre 12 orações) este é uma expressão-referencial que, como vimos acima, influencia fortemente a anteposição do verbo. Sendo assim, o resultado discrepante que encontramos no corpus do NURC/SP se deve mais ao tipo de sujeito que predominantemente aparece nessas sentenças encaixadas.

O tipo de elemento Q- também foi apontado como um fator determinante da posição do sujeito e do verbo. Temos que o qual, em todos os corpora, ocorre quase que categoricamente com a ordem V-S. No entanto, com esse tipo de pronome interrogativo, normalmente usamos uma cópula e é a sua presença que ocasiona tão alto índice de anteposição do verbo nesse contexto. Com relação aos demais elementos Q-, podemos observar que o porque, o quanto, o quando e o onde preferem a não anteposição do verbo. Já o como apresenta um comportamento basicamente neutro, ocorrendo com frequência quase igual em ambas as ordens. Os demais tipo de elemento Q- se mostram inconstantes: ora preferindo a ordem V-S (NURC/SP e dados de língua escrita), ora a ordem S-V (dados de língua falada). Mais uma vez é importante colocar que só foram consideradas as sentenças onde o elemento Q- não substitua o

sujeito. Na verdade, essa variável é muito afetada pelos fatores co-ocorrentes com ela, principalmente o tipo de verbo e o tipo de sujeito.

O último fator apontado como explicativamente forte para o fenômeno foi o tipo de corpus. No entanto, essa variável só foi selecionada na análise dos dados do Projeto NURC/SP. Porém, ela se mostrou relevante também para os diversos contextos de língua escrita, pois, nesse corpus, pudemos perceber uma distribuição complementar entre as situações de escrita. As sentenças coletadas a partir das redações do vestibular mostraram uma predominância da ordem V-S, sendo que, por outro lado, nos romances, a ordem S-V foi a preferida e, mais ainda, nas peças teatrais, a anteposição do verbo quase não existiu. Dessa forma, pode-se perceber que foi confirmada a hipótese inicial segundo a qual, nos textos de redações do vestibular, tende-se a usar uma modalidade mais próxima da norma gramatical. Sendo assim, podemos concluir que o tipo de texto que está sendo escrito influencia o tipo de estruturas nele empregadas.

Com relação aos dados da língua falada, temos que, contrariamente ao que seria de se esperar, esse é o contexto que obteve a maior frequência de anteposição do verbo (42%, contra 37% do Projeto NURC/SP e 21% nos dados de língua escrita). Mas, é também o corpus onde mais ocorreram sujeitos não pronominais - mais da metade das ocorrências possui como sujeito uma expressão-referencial e, nesse contexto (sujeito = expressão-r), em 73% dos casos a ordem adotada foi V-S. Assim sendo, o resultado inesperado pode ser explicado pelo tipo de sujeito mais utilizado nesse corpus.

De uma forma geral, a construção sem inversão sujeito-verbo é a preferida em todos os tipos de corpus. Analisando-se os resultados probabilísticos, temos que, no contexto de língua falada, ocorre a maior incidência de V-S (.42). Em seguida, encontramos os dados do Projeto NURC/SP, que apresenta um índice de .27 de ocorrência do verbo anteposto ao sujeito. Por último, com um resultado bem próximo de zero, .03, temos as sentenças coletadas de textos escritos atuais. A princípio podemos achar esses resultados um tanto quanto inesperados, porém, mais adiante, pretendo discutir quais seriam as possíveis descrições sintáticas para cada uma das ordens encontradas nesse tipo de estrutura interrogativa e, então, acredito poder explicar adequadamente esse quadro obtido.

Resumindo, podemos dizer que os fatores linguísticos que influenciam a estruturação das sentenças interrogativas são: tipo de verbo, tipo de sujeito, estatuto da oração e, também, tipo de elemento Q-. E o fator extra-linguístico considerado nesse trabalho - tipo de corpus - também se mostrou relevante para a questão ordem sujeito-verbo.

Nesse momento, gostaria de retomar o texto de Kato (1993) onde a autora procura explicar todas as possibilidades de ordenação encontradas nas interrogativas com elemento Q- da seguinte forma:

1) Com o expletivo é que é possível tanto SV quanto VS: para explicar essa situação, a autora utiliza a proposta de Lopes Rossi (1993) segundo a qual, nessas construções, a extração se faz da sentença clivada. Dessa forma estaria explicada tanto a ordem VS quanto SV que ocorre com é que, sendo que esta última seria, na verdade, a mais comum, porém a primeira também poderia ocorrer.

2) Sentenças interrogativas com ordem SV: Kato analisa essas estruturas como sendo uma consequência da extração das sentenças clivadas nas quais houve o apagamento tanto da cópula quanto do complementizador que.

3) Inversão VS nas interrogativas Q-: Para a autora, essas construções não são derivadas de clivadas e sua proposta é que, nesse contexto, a extração se faz de estruturas que apresentam um deslocamento à direita com pronome nulo. Ou seja, o que ocorre na verdade é uma falsa inversão.

Isto posto, gostaria de confrontar as explicações de Kato resumidas acima com os resultados empíricos por mim obtidos.

O primeiro contexto abordado pela autora é o que coloca que tanto VS quanto SV podem ocorrer com o expletivo é que. Os meus dados mostram que isto realmente se verifica, sendo que, como apontado anteriormente, a ordem SV é a preferida.

124. "e o que a senhora considera uma boa peça teatral?" (NURC/SP, DID 234, 32').

125. "como está o homem?" (DE REPENTE, AS TRÊS DA TARDE, pag. 191,4).

126. "E o que é que essa Lucinha tá fazendo?" (O HOMEM CHEIO DE SONHOS, pag. 8,7).

127. "como é que surgiu a idéia de brincar com logotipos?" (JÔ SOARES 11:30, 06/12/90, SBT).

128. "que que você escreveu?" (SALA DE AULA/PRIMÁRIO)

129. "do que que é composta a ceita de Natal?" (NURC/SP, DID 235, 287').

Kato postula que as estruturas interrogativas com o expletivo é que são derivadas das sentenças clivadas, conforme

proposta de Lopes Rossi (1993) - exemplos repetidos abaixo para facilitar:

130. Foi você que leu o livro.
Quem foi que leu o livro?

131. É um livro que você leu.
O que é que você leu?

Os pares acima mostram sentenças clivadas com vários tipos de elementos focalizados e sentenças interrogativas com é que incidindo sobre esses elementos focalizados (grifo meu). Parece-me bem plausível a proposta de Lopes Rossi e creio que ela consegue explicar adequadamente a possibilidade de uso de ambas as ordens S-V e V-S nesse tipo de construção.

A segunda possibilidade de ordem apontada por Kato é a colocação do sujeito antes do verbo nas sentenças interrogativas sem o expletivo é que e sua hipótese é de que, nessa situação, o que ocorreu foi o apagamento desse expletivo. Na análise dos três tipos de corpora, várias foram as sentenças que apresentaram essa estrutura.

132. "que profissional você colocaria pra sua proteção?" (NURC/SP, DID 251, 511').

133. "por que tantas crianças perambulam pelas ruas ao invés de estarem estudando nas escolas ou em seus lares?" (REDAÇÃO DO VESTIBULAR/DISSERTAÇÃO).

134. "o que o governo deveria e poderia fazer para reverter essa situação?" (CANAL LIVRE, 07/12/90, TV Bandeirantes).

Quanto à explicação sintática dada por Kato para essa estrutura, os resultados estatísticos pouco podem contribuir. O que temos é que, na ordem S-V, predominam os sujeitos pronominais e os verbos transitivos. Porém, a proposta de Kato pode ser reforçada pelos resultados obtidos por Duarte (1992), pois, nesse texto, a autora mostra, usando dados diacrônicos, que até o aparecimento do expletivo é que em 1882 todas as interrogativas diretas exibiam a ordem V-S. A partir dessa data, percebe-se um aumento sensível no uso da ordem SV, chegando, no dias atuais, a superar a anteposição do verbo. Ou seja, se a ordem SV só começou a ocorrer no PB depois que a partícula é que começou a ser usada nas interrogativas, parece-me clausível atribuir-se ao apagamento desse elemento a não anteposição do verbo em relação ao sujeito nesses estruturas.

O último tipo de estruturação encontrado é a ordem VS nas sentenças em que o expletivo é que não aparece. Kato propõe que, nesse contexto, ocorre o que ela chama de falsa inversão - a extração se faz de estruturas que apresentam um deslocamento à

direita com pronome nulo. As sentenças abaixo - repetidas aqui para facilitar a exposição - seriam a base da extração e o sujeito se encontraria em posição final.

135. (Eles) Estão aqui os meninos.

(Ela) Faturou vinte mil cruzeiros a nossa barraca.

Sendo que, como os sujeitos nulos parecem estar desaparecendo do PB (conforme Duarte, 1993), a tendência é que se revele a falsa inversão que ocorre nessas estruturas.

136. Onde eles estão os meninos?

Quanto ela faturou a sua barraca?

Voltando agora aos dados dessa pesquisa, gostaria de observar que tipos de verbo e sujeito aparecem na ordem V-S nas estruturas sem é que. Os resultados estatísticos mostram que o verbo predominante, nesse contexto, é a cópula. Vejamos alguns exemplos:

137. "como era ela?" (O QUE SERÁ?, pag. 13,9)

138. "como é isso hoje?" (FALA ESPONTÂNEA)

139. "há um tipo aí que é estéril qual é ele?" (NURC/SP, DID 18, 720')

140. "e do que era esse comercial?" (O QUE SERÁ?, pag. 39,3)

141. "então e qual é a dieta do seu regime?" (NURC/SP, DID 235, 76')

142. "como foi o episódio da sua cassação em 1964?" (RODA VIVA, 10/12/90, TV Cultura)

O que está influenciando a ordem nessas sentenças é exatamente o verbo copular utilizado, pois independentemente do tipo de sujeito a ordem V-S é amplamente majoritária. Sendo que a co-ocorrência do verbo copular com um sujeito pronominal se deu em menos de 3% dos dados dentre todos os três corpora (19 ocorrências em 708 sentenças). Isso quer dizer que é bem restrita essa situação e, observando-se mais de perto essas estruturas, vemos que se tratam de expressões quase que cristalizadas na língua.

143. "como estará ele?" (UM DIA VAMOS RIR DISSO TUDO, pag. 39,1).

Vejamos agora a situação das sentenças interrogativas que apresentam a ordem V-S e cujo sujeito é uma expressão-referencial. A análise estatística mostrou-nos que esse

é um fator altamente favorecedor da inversão V-S, portanto, vamos rever alguns exemplos onde ocorre a ordem V-S, mas o verbo não é uma cópula.

144. "desculpe a pergunta foi mal feita que profissões exercem essas pessoas?" (NURC/SP, DID 251, 124').
145. "(...) do que se compõe o cinema?" (NURC/SP, DID 251, 117').
146. "o que não vai falar o povo?" (DE REPENTE, ÀS TRÊS DA TARDE, pag. 23,9).
147. "o que significa essa chave?" (REDAÇÃO DO VESTIBULAR/NARRATIVA).
148. "que consequências pode ter nessa área econômica esse conflito?" (RODA VIVA, 21/01/91, TV Cultura).

Os exemplos acima mostram que as expressões-referenciais que aparecem depois do verbo, nos meus dados, possuem, em praticamente todos os casos, o traço mais definido. Além disso, elas ocorrem em posição final de sentença - observe, principalmente, o ex. 148 onde um sintagma adverbial aparece entre o verbo e o sujeito. Esses dois fatores corroboram a hipótese de Kato de que, nesses contextos, o que ocorre é uma falsa inversão.

Se considerarmos também as sentenças cujo verbo é uma cópula, aumentam ainda mais o número de estruturas V-S onde o sujeito é uma expressão-r definida.

149. "como é o nome dela?" (NURC/SP, DID 234, 82').
150. "ahn:: vai me ahn vai lhe custar alguma coisa e qual vai ser esse custo?" (NURC/SP, EF 388, 403').
151. "a senhora lembra como era o uniforme da senhora?" (NURC/SP, D2 396, 179').
152. "e do que era esse comercial?" (O QUE SERÁ?, pag. 39,3).
153. "qual é a sua graça?" (DE REPENTE, ÀS TRÊS DA TARDE, pag. 70,9).
154. "onde estaria o super-man para fazer o tempo retroceder?" (REDAÇÃO DO VESTIBULAR/DISSERTAÇÃO).
155. "como é a história da exposição que você fez e trocou os trabalhos todos os dias?" (JÔ SOARES 11:30, 06/12/90, SBT).
156. "qual é o objeto para ele?" (SALA DE AULA, 08/11/90).
157. "como seria suprida essa diferença?" (CANAL LIVRE, 03/01/91,

Se considerarmos que, nessas estruturas, não ocorre a inversão verbo-sujeito, mas sim um deslocamento à direita com pronome co-referente nulo, podemos entender também porque a ordem V-S é mais freqüente no corpus de língua falada do que no de língua escrita. Isso porque as estruturas de deslocamento à direita são bem mais comuns na fala do que na escrita, onde o usuário pode voltar no texto e refazer as sentenças se achar que está havendo ambigüidade. Esse recurso não é possível na fala, o que ocasiona um uso maior de sintagmas nominais no final da oração com o objetivo de esclarecer eventuais dúvidas. E o fato desse sintagma ter um caráter definido só reforça a hipótese da falsa inversão. É importante comentar que os dados não foram codificados levando-se em conta o traço mais ou menos definido do sujeito, pois a hipótese de que esse fenômeno poderia estar influenciando diretamente a ordem das interrogativas só foi levantada quando a análise estatística já se encontrava em estágio bem avançado e a recodificação dos dados se apresentava muito problemática. No entanto, usando de recursos do próprio programa VARBRUL, separei as sentenças com ordem V-S, sem o expletivo é que, e procedi com uma análise qualitativa dos mesmos. E é essa análise que me possibilita dizer que pouquíssimas são as ocorrências onde o sintagma nominal que aparece posposto ao verbo não possui o traço mais definido e, quando isso ocorre, o verbo que é usado é uma cópula.

A conclusão final que podemos tirar de tudo o que foi dito anteriormente é que a chamada a inversão Verbo-Sujeito nas interrogativas com elemento Q- é um fenômeno extremamente restrito no PB atual, aplicando-se basicamente quando o verbo é uma cópula. Nos demais casos, o que encontramos é uma falsa inversão onde um sintagma nominal aparece deslocado à direita com um pronome co-referente nulo (pro). Nas estruturas que contém o expletivo é que, por serem derivadas das sentenças clivadas, ambas as ordens V-S e S-V são igualmente possíveis.

REFERÊNCIAS BIBLIOGRÁFICAS:

- ÂMBAR, M.(1988) Para uma Sintaxe da Inversão Sujeito-Verbo em Português (dissertação de doutorado). Universidade de Lisboa.
- BERLINCK, R. de A.(1988) A ordem V SN no Português do Brasil: Sincronia e Diacronia. Dissertação de mestrado. IEL, UNICAMP.
- BRANDÃO, C.(1963) Sintaxe Clássica Portuguesa. Belo Horizonte.
- CASTILHO, A. & PRETTI, D.(1986) A Linguagem Falada Culta na Cidade de São Paulo. Vol. I - Elocuções Formais.
- .(1987) A Linguagem Falada Culta na Cidade de São Paulo. Vol. II - Diálogos entre dois informantes.
- CHOMSKY, N.(1981b) Lectures on Government and Binding. Foris Publications, Dordrecht.
- .(1986b) Barriers. The MIT Press, Cambridge, Mass.
- CHOMSKY, N. & LASNIK(1977) "Filters and Control" in Linguistic Inquiry, 8.
- CUNHA, C.(1976) Gramática do Português Contemporâneo. Editora Bernardo Alvares SA, Belo Horizonte.
- DUARTE, M.E.L.(1992) "A perda da ordem V(erbo) S(ujeito) em interrogativas qu- no português do Brasil" in D.E.L.T.A., vol. 8, número especial.
- KATO, M.A.(1987) "Inversão da ordem SV em interrogativas no português: uma questão sintática ou estilística?" in D.E.L.T.A., vol. 3, número 2.
- .(1992) "Variação Sintática e Estilo" in Cadernos de Estudos Lingüísticos, número 22.
- .(1993) "Word Order Change: the case of Brazilian Portuguese Wh- questions" (mimeo)
- KATO, M.A. & TARALLO, F.(1993). "The loss of VS syntax in Brazilian Portuguese" (mimeo)
- LOBATO, M.L.P.(1988) "Sobre a regra de anteposição do verbo no português do Brasil" in D.E.L.T.A., vol. 4, número 1.

- LOPES ROSSI, M.A.G.(1993) "Estudo Diacrônico sobre as interrogativas do português" in Viagem diacrônica pelas fases do Português Brasileiro, Editora da Unicamp.
- LYONS, J. (1987). Linguagem e Lingüística. Editora Guanabara, Rio de Janeiro, RJ.
- PONTES, E.(1982) "A Ordem VS em Português" in Ensaios de Lingüística, número 7.
- PRETTI, D. & URBANO, H.(1988) A Linguagem Falada Culta na Cidade de São Paulo, vol. III - Entrevistas.
- RIZZI, L.(1992) "Residual Verb Second and the Wh-criterion". Université de Geneve. (mimeo)
- SAID ALI, M.(1964) Gramática Secundária da Língua Portuguesa. Edições Melhoramentos, São Paulo.
- TARALLO, F. & KATO, M.A. (1989). "Harmonia trans-sistêmica: variação intra- e inter-lingüística" in Preedição, número 5.