

Projeto de Pesquisa:

SOBRE A PERSPECTIVA FUNCIONAL DA FRASE EM PORTUGUÊS:
REVISITANDO O PAPEL DAS “ESTRUTURAS SINTÁTICAS MARCADAS”¹

1. Essência e natureza do problema

É lugar-comum, em trabalhos sobre a sintaxe das línguas naturais que adotam uma perspectiva funcional, a afirmação de que os padrões de forma de uma língua não existem por acidente, mas são antes motivados pelas funções comunicativas a que servem (ver Halliday 1967, 1985, Dik 1978, Givón 1979, 1984, 1990; em português, Pontes 1986, 1987, Naro & Votre 1989, Ilari 1992, entre outros). A idéia básica por trás desta afirmação é que a organização gramatical das línguas reflete distinções de “significado” em sentido amplo, incluindo funções relativas à “estrutura informacional” da frase – isto é, ao modo como a mensagem altera o estado informacional do interlocutor em uma certa situação comunicativa ou “contexto”. Incluem-se entre estas “funções discursivas” – especialmente em se tratando de “estruturas marcadas” como a ordem Verbo-Sujeito e o Deslocamento à Esquerda –, aquelas relativas às distinções entre tema e rema, dado e novo, tópico e foco, tópico e comentário.

Também é lugar-comum em trabalhos funcionalistas observar que há muita divergência não apenas na terminologia adotada para a identificação destas “funções discursivas”, mas principalmente na definição e caracterização de seu papel ou modo de operação no discurso (ver discussão em Ilari 1992, Vallduví 1992, Lambrecht 1995, Vallduví & Engdahl 1996). De fato, uma das dificuldades mais notáveis que tenho enfrentado no desenvolvimento de um outro estudo ainda em andamento (Menuzzi 2001, a sair) é a de aplicar noções como as de “tópico”, “foco”, etc., a ocorrências reais de enunciados no discurso. Considere-se, por exemplo, o trecho abaixo, no qual as “estruturas marcadas” aparecem sublinhadas (editorial do jornal Zero Hora, em Elichirigoity 1997):

(1) [a] No Brasil é o torto princípio da instabilidade que costuma presidir as relações entre os cidadãos e o Estado. Somos um país mutante porque [b] volúvel e insegura parece ser a tecnocracia enquistada nos mais altos escalões decisórios da República. [c] Praticamente a cada ano, trocam-se as normas para o financiamento das safras, no geral em desfavor de quem quer semear e colher riquezas para este país. [d] Modificam-se

¹ O presente projeto se encontra em andamento e tem duração prevista até 2007, sendo financiado por bolsa de produtividade em pesquisa do CNPq (Processo 306743/2004-3).

também as regras de financiamento do sistema habitacional e [e] torna-se a casa própria um sonho distante. [f] Altera-se a declaração do imposto de renda e [g] reduzem-se as possibilidades de abatimentos de despesas. [h] É isso que acaba de ocorrer pela enésima vez.

Qual seria o tópico (Givón 1992, 1993), ou o tema (Halliday 1985), das orações iniciais de (1)? A julgar pelos critérios operacionais sugeridos, o tópico, ou tema, seria o referente incluído no locativo *no Brasil*: é o “referente do discurso” que “ocupa a posição inicial” da frase e, poder-se-ia argumentar, é também aquele que tem alguma “continuidade referencial” (*o Brasil* pode ser a referência do sujeito implícito de (1b)). No entanto, a intuição que se tem é que *no Brasil* oferece apenas o “cenário principal” em que se deve localizar o “assunto geral” do parágrafo.

O melhor candidato a “tópico do discurso” de (1) é também um *referente* da frase inicial, aquele expresso pelo SN *o torto princípio da instabilidade*: afinal, o resto do parágrafo parece ser dedicado à enumeração de fatos que atestam a validade do dito princípio no Brasil e, portanto, servem para “explicar” o que o autor entende pelo tal princípio. Note-se que este “referente tópico” é introduzido por uma estrutura de “foco” – a estrutura clivada em (1a) –, mas não é retomado referencialmente. O efeito de “enumeração” é antes obtido por meio de uma sucessão de estruturas típicas de “foco” (Givón 1979, 1990) ou “téticas” (Lambrecht 1995): “inversões”, ou estruturas de ordem Verbo-Sujeito.

É difícil explicitar como noções como as de “tópico” ou “tema” se relacionam com a noção de “tópico do discurso” num texto como (1) (para uma discussão inicial, ver van Dijk 1981 e Brown & Yule 1983); e também é difícil entender como as diversas inversões encontradas ao longo do texto contribuem para sua organização temática específica; ainda assim, parece claro que as funções expressas por estas e outras estruturas marcadas em (1) são fundamentais para a identificação da organização temática do texto. Este fato pode ser vislumbrado a partir de uma adaptação de (1) em que todas as estruturas marcadas são “regularizadas” para estruturas caracterizadas pela ordem direta, como em (2) (também extraído de Elichirigoity 1997):

(2) [a] O torto princípio da instabilidade costuma presidir as relações entre os cidadãos e o Estado no Brasil. Somos um país mutante porque [b] a tecnocracia enquistada nos mais altos escalões decisórios da República parece ser volúvel e insegura. [c] As normas para o financiamento das safras são trocadas, no geral em desfavor de quem quer semear e colher riquezas para este país, praticamente a cada ano. [d] As regras de financiamento do sistema habitacional também são modificadas, e [e] a casa própria se torna um sonho

distante. [f] A declaração do imposto de renda é alterada, e [g] as possibilidades de abatimentos de despesas são reduzidas. [h] Isso acaba de ocorrer pela enésima vez.

O efeito da “regularização” do texto (1) parece evidente: o texto em (2) não possui a “fluidez” e a organização original de (1); parece não mais expressar de modo claro o desenrolar de seu “assunto”. Em outras palavras: não é possível, ou ao menos fica difícil, identificar o “tópico do discurso” em (2) e, especialmente, o modo como é desenvolvido. Isto indica que as “estruturas marcadas” usadas em (1) têm um papel instrumental, ativo, no processamento da informação veiculada pelo discurso e, em particular, no que diz respeito à identificação de sua organização temática. O fato de que é difícil precisar que papel é este revela que é preciso, antes de mais nada, compreendê-lo melhor.

Para isso, seria preciso especificar *explicitamente*: (a) o que são as chamadas “funções discursivas”, como a de “tópico” ou “foco” oracional, por exemplo; (b) que estruturas ou padrões oracionais as expressam; e (c) como aquelas funções contribuem, *especificamente*, para a constituição e o desenvolvimento do “tópico do discurso”, isto é, para a identificação de sua organização temática. As teorias disponíveis da estrutura informacional da frase, entretanto, têm se limitado, em larga medida, a responder às perguntas do tipo (a) e (b), deixando de lado a tarefa de investigar o papel das funções discursivas no processamento do discurso, especialmente no que diz respeito à constituição de uma representação de sua organização temática. Não há, portanto, respostas precisas a questões como (c).

Por isso, a análise que acima esbocei do discurso em (1) permanece, na maioria das teorias disponíveis, no nível da intuição: não há recursos teóricos que permitam “construir” uma representação de (1) que explicitamente formalmente o papel das estruturas de tópico, foco, etc., na constituição de sua organização temática. Identificar este papel, especialmente no que diz respeito às funções discursivas associadas às “estruturas marcadas” do português, e contribuir para uma teoria da representação do discurso que explicitamente este papel: eis a tarefa que o presente projeto pretende investigar.

2. Objetivo principal

Assim, o objetivo principal do projeto é contribuir para o estudo da perspectiva funcional da frase portuguesa, agora com vistas à elaboração de uma teoria explícita da estrutura do discurso. Esta teoria deve apresentar recursos que não se limitem a identificar as principais articulações da estrutura informacional da frase e a caracterizar funções discursivas como as de “tópico”, “foco”, “tema”, “rema”, etc., e sua relação com estruturas sintáticas específicas (como em Vallduví 1992, Lambrecht 1995, ou Büring 2003b; para o português, Ilari 1992). A teoria deve também fornecer meios para construir uma “representação básica da organização temática” do texto/discurso, representação na qual se

possa explicitar a contribuição específica das funções de “tópico”, “foco”, etc., para a identificação da organização temática do discurso.

3. Estado atual do conhecimento sobre o problema

3.1 Estruturas marcadas e sua “função discursiva”

Estruturas “marcadas” são usadas para sinalizar que uma frase é “não-típica”, isto é, que o conjunto de relações entre funções sintáticas, semânticas e discursivas que expressa não é a “norma”, para usar o termo de Givón (1979; ver também Givón 1984, 1990; em português, Ilari 1992, entre outros). Por exemplo, a primeira frase do parágrafo em (1) é caracterizada: (a) pelo deslocamento do locativo *no Brasil* para a posição inicial, talvez para sinalizar seu papel de “tópico”, ou de “elemento tematizado”; e (b) pela clivagem do SN *o torto princípio da instabilidade* – sujeito da oração clivada, mas predicado do cópula *ser* –, possivelmente para sinalizar seu papel de “foco”. A situação típica, ou a “norma”, segundo Givón, seria o sujeito ser o tópico ou o elemento tematizado, e o locativo, parte do “comentário” (que, na primeira frase de (1), é expresso pela oração clivada: ver discussão sobre “estruturas apresentacionais bi-oracionais em Lambrecht 1994).

Como exemplo mencionado mostra, as estruturas marcadas sinalizam alterações nas correspondências funcionais típicas por meio de “marcas lingüísticas”, isto é, por elementos do código lingüístico não presentes em estruturas não-marcadas: mudanças na entoação e na ordem típica das palavras, com o auxílio eventual de elementos gramaticais como pronomes, verbos “funcionais” como o cópula, etc. As principais “estruturas marcadas” do português – ou, ao menos, aquelas que receberam alguma atenção na literatura lingüística e que serão ponto de partida deste projeto – são todas sinalizadas com tais marcas: a ordem inversa verbo-sujeito, as estruturas de “deslocamento-à-esquerda” e de “topicalização contrastiva”, e as orações “clivadas” ou “cindidas”, para usar o termo de Ilari (1992). Revisitemos algumas das propriedades destas construções a fim de ilustrar o modo como têm sido caracterizadas.

Sobre a ordem inversa verbo-sujeito, ilustrada em (1) acima por (1a) a (1g), encontram-se vários estudos na literatura lingüística brasileira; sobre sua motivação funcional, as principais hipóteses encontradas são as de que esta estrutura é usada quando o sujeito da oração é o “foco” da oração (cf. Pontes 1986, 1987, baseada em Givón 1979), ou quando o sujeito é não-tópico e oração pertence ao *background* discursivo – hipótese que incluiria a situação em que o sujeito é foco, mas não se limitaria a ela (cf. Naro & Votre 1999; também Berlinck 1997). Tanto quanto eu saiba, não há estudo empírico da distribuição preferencial desta construção no discurso em português; a literatura teórica costuma apontar que, em outras línguas, muito freqüentemente é usada para a apresentação

de novos referentes do discurso, tendendo por isso a preferir a posição inicial de parágrafo (ver Givón 1990, 1993; Lambrecht 1994).

As estruturas de “Deslocamento-à-Esquerda” freqüentemente envolvem, como a inversão verbo-sujeito, uma alteração na ordem típica dos constituintes: há a “adjunção” à margem esquerda da frase de um constituinte que pode ter, mas não precisa ter, uma função sintática interna à frase. Caso ele possua uma tal função, fala-se de um constituinte “deslocado-à-esquerda”; caso não a possua, fala-se de um “tópico pendente” (inglês *hanging topic*). Em ambos os casos, o constituinte tende a ter uma curva entoacional própria, não sendo integrado ao contorno do resto da frase (cf. Givón 1990, 1993 e Lambrecht 1994), o que é freqüentemente – mas nem sempre – sinalizado graficamente pelo uso de vírgulas. Além disso, no caso de um “constituinte deslocado-à-esquerda”, há ainda a possibilidade de o SN nele contido ser retomado anaforicamente por pronomes ou mesmo descrições definidas.

Ilustra o deslocamento-à-esquerda uma frase como (1c) acima, mas também (1a) – embora, neste último exemplo, o “deslocamento” não seja sinalizado por vírgula. Do mesmo modo, no discurso oral abaixo transcrito, as frases (3a) e (3b) exemplificam casos de deslocamento-à-esquerda sinalizados pela pontuação, enquanto que as frases (3f) e (3g) exemplificam casos não sinalizados:

(3) J.B.: Quanto se gasta efetivamente com um preso?

A.T. Jr.: [a] Esses 685 reais da Detenção, tem muito preso falando assim: “Por que não dão isso aí pra mim por mês, que eu não vou mais fazer nada. Não vou mais assaltar.” [b] Pelo nosso cálculo, papel higiênico a família tem de trazer, [c] sabonete também (tem de trazer), [d] roupa ele tem de conseguir, [e] nem a calça é mais fornecida, e tem de ser no padrão, tem de ser bege. Então, vamos desconsiderar funcionário, aparato, burocracia e tal, o preso sai por 10 reais por dia. Vai que ele precise de uma aspirina (...) vai a 13 reais, que ele saia a 20 reais, certo? [f] Na Penitenciária é a mesma coisa, e [g] nos presídios do interior não difere muito, ele sai por volta de 20 reais. Mesmo que fossem 20 reais por dia, ainda sobram 85 reais por mês (...)

Quanto à motivação funcional das estruturas de deslocamento-à-esquerda, a literatura costuma apontar para o fato de que o constituinte deslocado é “temático” e “tópico” – tende a ser informação dada ou acessível e a apontar o referente mais proeminente do fluxo introduzido pela construção –; pode, ainda, envolver algum tipo de “contraste” com elementos tópicos precedentes (Pontes 1987, Ilari 1992, Pezzati 1998 e outros). Givón (1990, 1993) vê nesta última propriedade o elemento fundamental do deslocamento-à-esquerda; por isso, seria usado tipicamente em pontos de “ruptura temática”, isto é, que sinalizam uma nova unidade temática do discurso: um novo tópico discursivo, um novo

episódio em uma narrativa, um novo turno na conversação, etc. Isto explicaria sua distribuição no texto: o fato de o constituinte deslocado estar em regra distante da última menção do seu referente no texto, e da construção ocupar preferencialmente as posições inicial e medial de parágrafo.

A topicalização contrastiva é muito similar, ao menos superficialmente, ao deslocamento-à-esquerda: também envolve elemento “deslocado” para a margem esquerda da frase. A diferença, do ponto de vista da forma, é que nesses casos o constituinte não pode ser retomado anaforicamente por pronome ou descrição definida, e é o elemento “focal” da frase, recebendo seu acento mais proeminente e estando integrado ao contorno entoacional dela (diferença que, curiosamente, não é admitida e, portanto, explorada por Pontes 1987). Do ponto de vista funcional, também há diferenças. Embora o constituinte “topicalizado”, como o “deslocado-à-esquerda”, também seja em regra “temático” e “tópico”, não precisa ser – pode, inclusive, ser um elemento não referencial (cf. Pontes 1987, Pezzati 1998). Além disto, é necessariamente acompanhado de contraste e sua distribuição textual é diferente da dos constituintes deslocados-à-esquerda: ocupa normalmente a posição medial de parágrafo e não pode ser usado em posição inicial; por envolver contraste, tende a estar a curta distância do referente ao qual está ancorado (cf. Givón 1990, 1993). De acordo com estas características, poderíamos dizer que as frases (4b) a (4e) acima ilustram uma cadeia de “topicalizações contrastivas”.

3.2 Uma teoria da estrutura informacional

Como já se disse antes, a visão geral das abordagens funcionalistas é a de que as estruturas marcadas sinalizam que as relações entre funções sintáticas, semânticas e discursivas na frase não são as “típicas”. Em se tratando das “funções discursivas”, presume-se que o que acontece é que há diferentes maneiras de “indicar como a informação veiculada pela linguagem entra no estado informacional do ouvinte no momento de produção do enunciado. Quando comunicam uma proposição ϕ , os falantes podem realizá-la por meio de diferentes estruturas sentenciais, de acordo com suas crenças com respeito ao conhecimento e ao estado de atenção do ouvinte com relação a[os elementos de] ϕ ” (Engdahl & Vallduví 1996; ver também Vallduví 1992, 1994; Givón 1990, 1992, 1993; Lambrecht 1994; Ilari 1992, entre outros.)

Por exemplo, as frases assinaladas em (4B), (5B) e (6B) abaixo expressam uma mesma “proposição”, isto é, possuem as mesmas condições de verdade: todas elas descrevem o fato de que Paulo odeia chocolate. Ainda assim, as frases diferem em forma, e esta forma é própria para o contexto porque veicula uma “partição informacional” exigida por ele. As

diferentes “partições informacionais” são “marcadas” seja por meio da entoação, como em (4B) e (5B), seja pela entoação e a ordem de palavras, como em (6B):²

- (4) A: O que o Paulo odeia?
B: [Ele odeia CHOCOLATE].
- (5) A: (É verdade que) o Paulo adora chocolate?
B: Não, [ele ODEIA chocolate].
- (6) A: E de doce, o Paulo gosta?
B: [*Chocolate* ele ODEIA], mas *o resto* ele COME.

Note-se o uso da forma apropriada à “partição informacional” exigida pelo contexto é fundamental para a aceitabilidade do discurso: não se poderia utilizar (5B) como uma resposta para (4A), ou (4B) como resposta para (6A), por exemplo. É neste sentido que se diz que a sintaxe “reflete” exigências informacionais do discurso. Indicar, portanto, “como a informação veiculada pela linguagem entra no estado informacional do ouvinte no momento de produção do enunciado” exige, portanto, uma teoria que: (a) explicita as “unidades informacionais” e suas funções no processamento do discurso; e (b) identifique as estruturas sentenciais da língua que codificam estas unidades.

Recentemente apareceram trabalhos que procuram desenvolver teorias explícitas, formalizadas, tanto dos elementos que constituem o chamado “componente informacional” da linguagem (por exemplo, Vallduví 1992, 1994; Vallduví & Engdahl 1996, Lambrecht 1995, Erteschik-Shir 1998) quanto da própria organização do discurso (por exemplo, Kamp & Reyle 1992, Roberts 1996, Buring 2003a, 2003b). Um dos objetivos deste projeto é, justamente, comparar tais teorias com relação a sua capacidade de caracterizar as funções das estruturas marcadas, especialmente com respeito a seu papel na constituição da estrutura temática do discurso. No que segue, apresentarei uma destas teorias – a teoria de Vallduví (1992, 1994) – a fim de ilustrar o tipo de problema que colocam.

Para Vallduví, a frase é dividida em duas unidades informacionais básicas, *foco* e *ground*, sendo que o *ground* pode se subdividir em duas subunidades, *link* e *tail*.³ A formalização destas unidades é baseada num modelo para diálogo. Segundo este modelo, a interação comunicativa verbal é concebida como um sistema de atualização de informação

² Letras maiúsculas sinalizam o “acento-A”, ou “acento nuclear”, da frase; itálicas sinalizam o “acento-B”, isto é, a curva entoacional típica de “tópicos”, cf. Jackendoff (1972); ver Ilari (1992) para uma descrição destes traços entoacionais em português.

³ Todas estas noções possuem correspondentes próximos nas distinções mais tradicionais entre rema e tema, novo e dado, foco e tópico, etc.; ver Vallduví & Engdahl 1996 para discussão.

(ing. *information update*), isto é, como um sistema baseado em representações dos “estados informacionais” dos interlocutores e em operações de mudança (“atualização”) destes estados.⁴ Estes estados são representados por “arquivos” (*files*), que são coleções de “cartões” de arquivo (*file cards*). Cada cartão corresponde a um “referente do discurso” – isto é, uma das entidades às quais o discurso se refere. Os cartões são compostos por um conjunto de “registros” (*records*) que representam, por meio de funções proposicionais, relações e atributos do referente. As diferentes unidades básicas da estrutura informacional de uma frase atuam, então, como *instruções* para a atualização dos registros nos cartões.

Foco é definido como a “informação nova” de uma sentença S, ou seu “potencial de atualização”, isto é, a única contribuição que S faz ao estado informacional do ouvinte no momento de realização do enunciado. Esta atualização pode se dar por adição de informação nova, ou por substituição de informação presente; o foco pode, portanto, executar sinalizar ao interlocutor dois tipos de “instrução”: UPDATE-ADD(*informação*) ou UPDATE-REPLACE(*informação*).

Ground é o elemento que o falante presume já estar presente neste estado informacional e serve como “guia” para o foco, garantindo a “ancoragem” e a “inserção” apropriadas da informação no estado informacional do ouvinte. Toda a sentença possui necessariamente algum foco, mas o ground só está presente se o contexto exige.

Como se disse antes, o ground pode possuir duas subunidades, cada qual com uma função específica. O *link* serve para estabelecer (abrir ou ativar) um cartão de arquivo particular, isto é, indica em que cartão o foco de S deve ser registrado, correspondendo à instrução GOTO(*cartão*); se o local de registro é o mesmo da sentença anterior, a sentença não precisa de uma expressão que marque o *link* – a instrução *default* é, portanto, a manutenção do cartão atualmente aberto.⁵ Um *tail*, quando presente na sentença, aponta que registro particular no cartão deve ser atualizado; neste caso, a informação veiculada pelo *foco* deve modificar o registro designado pelo *tail* – isto é, a instrução veiculada ao interlocutor é UPDATE-REPLACE (*informação, registro*). Este é o modo não-*default* de atualização; o modo *default* se dá quando a informação veiculada pelo *foco* deve ser simplesmente adicionada como um registro, em cujo não há um *tail* presente em S, e a instrução ao interlocutor é, simplesmente, UPDATE-ADD(*informação*).

⁴ A teoria de Vallduví não é, portanto, primariamente concebida para dar conta do fluxo de informação de outros tipos de discurso mais elaborados, como o texto narrativo, etc. Presume-se, entretanto, que os elementos fundamentais do processamento informacional do discurso sejam os mesmos, ainda que textos mais “organizados” possam envolver níveis de estruturação não imediatamente necessários à análise de trocas simples entre perguntas e respostas.

⁵ Isto codifica no sistema a idéia de que o “tópico contínuo” é o tópico “não-marcado”, isto é, expresso por “formas fracas” – elementos não-proeminentes entoacionalmente, como pronomes não-acentuados e elementos nulos. Ver Givón (1990, 1992) e Lambrecht (1994), entre outros, para discussão.

Em resumo, o modelo permite quatro tipos de organização informacional da sentença, correspondendo a quatro tipos de atualização do estado informacional:

(7) Instruções do componente informacional:

- a. *link – foco* : [GOTO(*cartão*), UPDATE-ADD(*informação*)]
- b. *link – foco – tail* : [GOTO(*cartão*), UPDATE-REPLACE(*informação*)]
- c. *só foco* : [UPDATE-ADD(*informação*)]
- d. *foco – tail* : [UPDATE-REPLACE(*informação, registro*)]

O modelo de Vallduví permite dar conta de várias das possibilidades e combinações de interações dialogais e, especialmente, das escolhas feitas entre várias estruturas disponíveis em diversas línguas – em particular, o catalão. Aplicando o modelo ao caso dos diálogos do português antes ilustrados, teríamos análises como as seguintes:

(8) Frase em (4B) acima:

- a. Não há *link* (herdado da frase anterior: $x = Paulo$);
- b. *tail*: $\lambda x \lambda y$ [x odeia y];
- c. *foco*: $y = chocolate$
- d. instrução: [UPDATE-ADD($y = chocolate, \lambda x \lambda y$ [x odeia y])]

(9) Frase em (6B) acima:

- a. *link*: $y = chocolate$
- b. *tail*: $\lambda y \lambda R$ [Paulo R y]
- c. *foco*: $R = odeia$
- d. instrução: [GOTO(*chocolate*), UPDATE-REPLACE($R = odeia, \lambda y \lambda R$ [Paulo R y])]

Esta é uma análise *explícita* de noções como as de “tópico”, “foco”, etc., e permite, portanto, não apenas ter uma idéia de sua “função” no fluxo de informação; permite também formular de maneira precisa relações entre estas funções e as estruturas sentenciais que as expressam. Para dar um exemplo: de acordo com esta teoria, as construções de deslocamento-à-esquerda expressam estruturas informacionais em que há um link e um foco e podem ou não ter um tail. Oferecendo uma caracterização precisa das funções codificadas nas sentenças, o modelo permite também levantar questões relativas à “naturalidade” da relação entre construção e funções: por exemplo, *por que* o deslocamento-à-esquerda seria apropriado para expressar a função de link? etc. E, finalmente, por ser uma teoria explícita, permite formular de modo preciso os problemas que enfrenta e, portanto, as questões que podem ser investigadas objetivamente.

3.3 Questões e problemas

Como indicarei brevemente, uma teoria que simplesmente identifica as funções discursivas expressas na frase, mas não o modo como estas funções contribuem para a identificação da organização temática do discurso, fornece um quadro incompleto do papel delas e, portanto, do papel das estruturas sintáticas marcadas.

Na análise intuitiva que fiz do exemplo em (1) acima, sugeri que uma cadeia de estruturas téticas com sujeitos incluídos no foco (as estruturas de ordem VS) permite, de algum modo, contribuir para a “progressão temática” do texto. No caso de (1), a cadeia contribui para a progressão do tema relacionado a um “tópico referencial”, *o torto princípio da instabilidade*, de modo que pode ser vista como parte de uma “unidade tópica” do discurso. É interessante que a transposição das estruturas de ordem VS para a ordem SV – e a conseqüente promoção dos sujeitos do foco a tópicos – “quebra” o fluxo da progressão temática, cf. (2). Este caráter “ruptor” das “estruturas téticas” na progressão temática pode ser visto também no uso das estruturas de deslocamento à esquerda no texto (3), em que (3a), (3f) e (3g) sinalizam subtemas do tema principal “o custo do preso”. Note-se que, em (3), seria o *não* deslocamento do elemento tético que “desorganizaria” o fluxo informacional, tornando a identificação da estrutura temática extremamente difícil.

Mas as observações acima não podem ser interpretadas com indicando que “estruturas téticas” nunca possam convergir em uma mesma direção temática, nem que estruturas de foco nunca possam ser utilizadas para uma “ruptura” temática. Considere, por exemplo, o trecho abaixo (adaptado de Lambrecht 1994, p. 160):

(10) Eu acabei meu curso secundário sendo considerado um aluno médio. Nunca me esforcei para conseguir mais do que isso. [a] *História*(,) eu achava uma matéria muito chata. [b] *Em matemática*(,) eu nunca fui bem. [c] Eu gostava de CIÊNCIAS e [d] FUTEBOL era meu “prato predileto”, é claro.

Como se vê, as “estruturas téticas” de deslocamento à esquerda em (10a) e (10b) contribuem para a progressão de uma mesma unidade temática (em que se caracteriza o falante como um “aluno médio”). Por outro lado, a ruptura desta unidade ocorre justamente em (10c), em que *ciências* não aparece deslocado à esquerda – embora pertencendo às coisas que o falante fazia na escola e, portanto, parte do conjunto de referentes a que *história* e *matemática* pertencem –, mas como parte do foco da frase. Note-se que, embora *ciências* em (10) apareça em circunstâncias frequentemente associadas ao do deslocamento à esquerda (ponto de ruptura, de um lado, e parte de uma lista de itens em contraste, de

outro), o uso desta estrutura é absolutamente impossível em (10) (substituição de (10c) por “*De ciências(,) eu gostava*” – entoação análoga a de (10a,b) – torna o discurso esquisito).⁶

O que vemos nos exemplos que acabo de mencionar é que estruturas de tópico e de foco contribuem de modo decisivo para o fluxo do texto e, em particular, para a identificação de sua organização temática: afinal, como indiquei, a alteração destas estruturas, nos contextos apresentados, leva a problemas no processamento e, em alguns casos, à inaceitabilidade do discurso. Mas é claro que uma teoria que se limite a explicitar o procedimento pelo qual uma estrutura de foco “adiciona informação em um arquivo”, por exemplo, pouco esclarece sobre os fenômenos que discuti aqui: por que a “adição de informação a um arquivo” em alguns casos contribui para a progressão temática, e em outros, para a ruptura desta progressão? Evidentemente, o mesmo problema surge para as estruturas de tópico.

Tanto quanto sei, estas questões não foram endereçadas pelas teorias disponíveis da estrutura informacional da frase e de sua relação com a organização do discurso; mas, claramente, elas são fundamentais se o que se quer é, realmente, “entender” o uso das articulações informacionais – e, portanto, o papel das estruturas marcadas – no processamento do discurso. Basta mencionar que, se as observações acima estão corretas, a compreensão usual de várias outras observações – mencionadas na seção 3.1 acima – não pode ser mantida como tal. Isto vale particularmente para a observação de que as estruturas marcadas tendem a preferir certas posições no discurso: estruturas de foco ocorreriam em abertura de parágrafos; o deslocamento-à-esquerda, tanto em abertura de parágrafos quanto em posição medial; a topicalização contrastiva, basicamente no interior de parágrafos. Como se viu com os exemplos acima, estas podem ser tendências estatísticas que sugerem algo da natureza destas estruturas – e das funções que expressam –; mas o uso delas vai além e precisa de uma explicação que não se limite a estas “normas”.⁷

⁶ Para confirmar que o que se vê em (10) não é um uso “não-típico” de estruturas de foco e tópico, considere-se o seguinte trecho de Rubens Fonseca (*Agosto*, p. ...):

“[a] *Lacerda* se despediu do major *Vaz* e caminhou com o filho para a porta da garagem do edifício. *Vaz* foi em direção ao carro. *Alcino* atravessou a rua e atirou em *Lacerda*, que correu para o interior da garagem. [b] O estrondo do revólver ao disparar surpreendeu *ALCINO*, que por instantes ficou sem saber o que fazer. Notou então que o major se aproximara e agarrava sua arma. Novamente *Alcino* acionou o gatilho. O major continuou agarrando o cano do revólver até que *Alcino*, num repelão, soltou a arma dos dedos que a prendiam, caindo com o esforço que fizera.”

Toda a seqüência em [a] é caracterizada por uma sucessão de frases de estrutura tópico-comentário em que os tópicos são personagens contribuindo para o progresso da ação; a cena é vista de uma perspectiva ampla, sem predomínio do ponto de vista de qualquer dos personagens. No ponto [b], há uma ruptura temática: a perspectiva da cena muda, concentrando-se na luta entre *Alcino* e *Vaz*, percebida a partir do ponto de vista de *Alcino*: pois neste ponto, *Alcino* é reintroduzido por uma estrutura de foco, e não como tópico.

⁷ Um outro problema para a teoria de Vallduví reside no fato de que algumas vezes é difícil justificar o uso de certas estruturas marcadas com base nas distinções funcionais que o modelo fornece. Isto acontece, por exemplo, com o caso de estruturas de deslocamento-à-esquerda e/ou topicalização contrastiva em que o

4. Síntese da metodologia

As tarefas básicas que guiarão o trabalho proposto neste projeto são as seguintes: (a) levantamento bibliográfico e estudo de teorias explícitas, formalizadas, da organização do discurso e das funções discursivas envolvidas no uso das estruturas sintáticas marcadas; (b) levantamento de dados e constituição de um corpus com um número satisfatório de ocorrências de quatro das estruturas marcadas do português – a ordem inversa verbo-sujeito, o deslocamento-à-esquerda, a topicalização contrastiva e as orações clivadas; (c) descrição e análise dos dados sob a ótica de diferentes modelos e tentativa de reformulação/elaboração destes modelos a fim de acomodar os problemas encontrados (como, por exemplo, a organização hierárquica em níveis de tópicos); (d) produção de relatórios/artigos sistematizando e apresentando os resultados do trabalho.

No que diz respeito aos métodos de investigação a serem empregados, o projeto seguirá procedimentos científicos usuais (formulação de hipóteses explícitas a partir de um conjunto restrito de dados, verificação da hipótese em um conjunto ampliado de dados, correção da teoria ou abandono e substituição se necessário, nova ampliação do conjunto de dados, etc.). É importante esclarecer, entretanto, que um dos recursos heurísticos fundamentais a ser utilizado é o estudo quantitativo de ocorrências naturais das diferentes estruturas marcadas que são objeto deste projeto. O objetivo destes estudos será estabelecer a distribuição preferencial daquelas estruturas em português, a fim de verificar se as observações ocasionais encontradas na literatura se confirmam e se são, realmente, parte do problema que a teoria da organização informacional do discurso deve resolver.

Ainda em relação aos métodos de investigação, é preciso justificar brevemente o fato de o projeto ser orientado ao estudo de um corpus de português *escrito*. Há duas razões para esta decisão. A primeira é que o texto escrito tende a ser mais elaborado que o discurso oral, e isto pode revelar aspectos da organização discursiva – e do uso das estruturas marcadas – que podem não estar presentes ou não ser tão proeminentes no discurso menos planejado. A segunda razão é que o projeto pretende fornecer subsídios para o processamento automático de textos digitalizados, em cujo caso de novo o que é relevante

elemento deslocado não parece nem ser o foco, nem ser o *link* da frase (como no caso dos elementos sublinhados em (1i) e (12) acima.

Com relação a outros modelos da estrutura de discurso, é interessante mencionar a teoria desenvolvida por Roberts (1996) e Büring (2003a, 2003b) em função do fato de ela preencher uma das lacunas mais notáveis na teoria de Vallduví – ela fornece uma formalização apropriada para a hierarquia de tópicos. Mas tem, por outro lado, um problema sério: parte do pressuposto de que todo o tipo de texto que se organiza em termos de uma hierarquia tópica envolve uma estrutura subjacente de diálogo. Embora a idéia possa fazer sentido para certos tipos de textos – em particular, textos argumentativos – torna-se extremamente artificial na análise de outros, como os textos narrativos.

são os textos escritos, especialmente os encontrados em formato digital na internet, como será o caso dos textos que comporão o corpus do presente projeto.

5. Resultados esperados

Entre os dois principais resultados que espero obter com o presente projeto estão: (a) o desenvolvimento de uma teoria realista da representação do discurso e de seu processamento, especialmente por meio da tentativa de explicitação/formalização de “funções discursivas” envolvidas no uso de “estruturas sintáticas marcadas”, e (b) a identificação e a análise apropriada do papel de diferentes estruturas sentenciais marcadas do português no processamento do discurso. Uma das conseqüências dos resultados (a) e (b) deve ser, por exemplo, uma explicação adequada das tendências distribucionais no uso das estruturas marcadas (por exemplo, suas posições preferenciais no parágrafo, etc.).

Como resultados secundários do projeto, espero apresentar subsídios para: (a) teorias do texto e do discurso interessadas em aspectos como a análise da estrutura de tópicos de um texto e a relação entre estrutura de tópicos e tipo de texto; (b) modelos de processamento automático do discurso escrito que busquem maior eficiência na “manipulação” de informação textual (por exemplo, para a busca da hierarquia de tópicos de um texto, o que pode contribuir para a produção automática de resumos; para a resolução de aspectos da anáfora dependentes da estrutura tópica do texto; etc.); (c) sistemas de tradução automatizada, que têm como uma das principais fontes da baixa qualidade de seus resultados a ausência de uma análise comparativa da estrutura informacional da língua-fonte e da língua-alvo.

Referências bibliográficas

- Berlinck, Rosane de Andrade (1997) *Nem Tudo o que é Posposto é Novo: Estatuto Informacional do Sujeito em Português*. Alfa 41, 57-78.
- Büring, Daniel (2003a) *On D-Trees, Beans, And B-Accents*. *Linguistics and Philosophy* 26, 511-545.
- Büring (2003b) *Semantics, Intonation and Information Structure*. Ms., UCLA. To appear in Gillian Ramchand & Charles Reiss, eds., *Interfaces*.
- Dik, Simon (1978) *Functional Grammar*. North Holland, Amsterdã.
- Elichirigoity, Maria Teresinha Py (1997) *Estruturas Sintáticas Marcadas por Reordenação em Dissertações de Vestibulandos*. Dissertação de mestrado, UCPel, Pelotas.
- Engdahl, Elisabet & Vallduví, Enric (1996) *Information Packaging in HPSG*. *Edinburgh Working Papers in Cognitive Science* 12, 1-32.
- Erteschik-Shir, Nomi (1997). *The Dynamics of Focus Structure*. Cambridge University Press, Cambridge (Inglaterra).
- Givón, Talmy (1979) *On Understanding Grammar*. Academic Press, Nova Iorque.

- Givón, Talmy (1984) *Syntax: A Functional-Typological Introduction*, Vol I. John Benjamins, Amsterdã.
- Givón, Talmy (1990) *Syntax: A Functional-Typological Introduction*, Vol II. John Benjamins, Amsterdã.
- Givón, Talmy (1992) The grammar of referential coherence as mental processing instructions. *Linguistics* 30, 5-55.
- Givón, Talmy (1993) *English Grammar: A Function-Based Introduction*, Vols. I e II. John Benjamins, Amsterdã
- Halliday, Michael A. K. (1985) *An Introduction to Functional Grammar*. Arnold, Londres.
- Hopper, Paul & Thompson, Sandra (1980) Transitivity in Grammar and Discourse. *Language* 56, 251-299
- Ilari, Rodolfo (1992) *A Perspectiva Funcional da Frase Portuguesa*. Ed. da UNICAMP, Campinas.
- Jackendoff, Ray (1972) *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge (EUA).
- Kamp, Hans, and Uwe Reyle (1993) *From Discourse to Logic*. Kluwer, Dordrecht.
- Lambrecht, Knud (1994) *Information Structure and Sentence Form: Topic, Focus, and the mental representations of discourse referents*. Cambridge University Press, Cambridge (Inglaterra).
- Menuzzi, Sergio (a sair) *A Ordem Verbo-Sujeito no Português do Brasil: Para uma Comparação das Abordagens Formalistas e Funcionalistas*. A sair na Revista da ANPOLL.
- Menuzzi, Sergio (2001) *A Ordem Verbo-Sujeito no Português do Brasil: Por uma Avaliação Empírica das Diferentes Análises*. Ms., PUCRS, Porto Alegre.
- Naro, Anthony & Votre, Sebastião (1989) Mecanismos Funcionais do Uso da Língua. *Revista DELTA* 5, 169-184.
- Naro, Anthony & Votre, Sebastião (1999) Discourse Motivations for Linguistic Regularities: Verb-Subject Order in Spoken Brazilian Portuguese. *Probus* 11, 75-100.
- Neves, Maria Helena de Moura (1997) *A Gramática Funcional*. Martins Fontes, São Paulo
- Pezatti, Erotilde Goreti (1998) Constituintes Pragmáticos em Posição Inicial: Distinção entre Tema, Tópico e Foco. *Alfa* 42, 133-150.
- Pontes, Eunice (1986) *Sujeito: da Sintaxe ao Discurso*. Ática, São Paulo.
- Pontes, Eunice (1987) *O Tópico no Português do Brasil*. Pontes, Campinas.
- Roberts, Craige (1996) Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. In J. H. Yoon & A. Kathol, eds., *OSU Working Papers in Linguistics* 49: Papers in Semantics, 91-136.
- Vallduví, Enric (1992) *The Informational Component*. New York: Garland.
- Vallduví, Enric (1994) Updates, files, and focus-ground. In Peter Bosch & Rob van der Sandt, eds., *Focus and Natural Language Processing*. IBM Working Papers of the Institute for Logic and Linguistics 8, Vol. 3, 649-658. IBM Deutschland, Heidelberg.
- Vallduví, Enric & Engdahl, Elisabet (1996) The linguistic realization of information packaging. *Linguistics* 34, 459-519.