**Mestrado em Linguística**

# O TÓPICO NO DISCURSO ORAL

**Anotação de diferentes tipos de tópico num corpus de discurso oral**

Diana Margarida Santos Machado

Orientadora: Professora Doutora Fátima Silva

Versão final

Dissertação submetida à faculdade de Letras da Universidade do Porto para obtenção do grau de Mestre em Linguística, realizada sob orientação da professora Doutora Fátima Silva

Esta tese foi realizada com o apoio de Fundos Nacionais através da FCT – Fundação para a Ciência e a Tecnologia no âmbito do projeto COPAS – Contraste e Paralelismo na fala (PTDC/CLE-LIN 120017/2010)

## AGRADECIMENTOS

Para a realização deste contribuíram várias pessoas a diferentes níveis.

Começo por manifestar a minha gratidão aos meus pais, que sempre me apoiaram e me continuam a acompanhar incondicionalmente na realização de todos os meus projetos pessoais e académicos.

Agradeço também à Professora Fátima Silva, por me ter orientado com disponibilidade e entusiasmo neste percurso.

Manifesto igualmente o meu agradecimento aos colegas do projeto COPAS, em especial às Professoras Fátima Oliveira e Ana Isabel Mata e à Helena, pelo incentivo e ajuda que sempre me deram ao longo deste ano.

## RESUMO

Nesta dissertação pretende-se observar os diferentes papéis que o tópico pode assumir no discurso oral espontâneo.

Os dados utilizados na realização deste estudo pertencem ao projeto COPAS, que tem como objetivo determinar o modo como os diferentes tipos de informação são veiculados pelos falantes no discurso oral. Este projeto conta com a participação de investigadores das áreas linguísticas da semântica, sintaxe e prosódia, no sentido de estabelecer interface entre estes domínios para o estudo das questões centrais que ocupam o projeto, nomeadamente a determinação da forma como a entoação e os recursos sintáticos são usados pelos falantes, como se relaciona a estrutura da informação com esses aspetos, e a arquitetura da gramática.

Este estudo encontra-se dividido em dois capítulos principais. O primeiro centra-se na delimitação do enquadramento teórico para a proposta de anotação do *corpus* em análise, incidindo essencialmente sobre questões que se relacionam com o discurso, a estrutura informacional, os conceitos de tópico e foco, e de dado e novo, além de proceder ao levantamento de uma taxonomia de tópicos discursivos.

O segundo capítulo apresenta os procedimentos que levaram à definição de uma terminologia de análise dos diferentes tipos de tópico, a análise do corpus, os resultados e o levantamento de hipóteses que possam justificar os resultados obtidos.

## ABSTRACT

This dissertation aims to observe the different roles that the topic can take on spontaneous oral discourse, specifically CPE FACES.

This project includes the participation of researchers from the areas of language semantics, syntax and prosody, to establish an interface between these areas for the study of the central questions that concern the project, including the determination of how the intonation and syntactic features are used by speakers, how information structure relates with those aspects, and the architecture of grammar.

This study is divided into two main chapters. The first focuses on the delineation of the theoretical framework for the proposed annotation corpus analysis, focusing mainly on issues that relate to the speech, the informational structure, the concepts of topic and focus, and given and new, as well as to survey a taxonomy of discursive topics.

O segundo capítulo apresenta os procedimentos que levaram à definição de uma terminologia de análise dos diferentes tipos de tópico, a análise do corpus, os resultados e o levantamento de hipóteses que possam justificar os resultados obtidos.

# Índice de esquemas

**Capítulo 1 – Tópico, foco e estrutura informacional no discurso: alguns fundamentos teóricos**

Esquema 1.....17
Esquema 2.....25

**Capítulo 2 – Proposta de anotação semântica de um discurso oral**

Esquema 1.....47

# Índice de tabelas

Tabela 1.....36
Tabela 2.....36

# Índice de figuras

Figura 1.....44
Figura 2.....44
Figura 3.....45
Figura 4.....46

Figura 5.....47
Figura 6.....48
Figura 7.....48
Figura 8.....49
Figura 9.....50
Figura 10.....50
Figura 11.....51
Figura 12.....51
Figura 13.....52
Figura 14.....53
Figura 15.....53
Figura 16.....54
Figura 17.....55
Figura 18.....55
Figura 19.....56
Figura 20.....56
Figura 21.....57
Figura 22.....58
Figura 23.....58
Figura 24.....59
Figura 25.....60
Figura 26.....60
Figura 27.....61

## Índice de quadros

Quadro 1.....64

## Índice de gráficos

Gráfico 1.....61
Gráfico 2.....62
Gráfico 3.....63

# Índice

Introdução.....1

**Capítulo 1 – Tópico, foco e estrutura informacional no discurso: alguns fundamentos teóricos**.....7

1. Discurso e universo discursivo.....8
2. O conceito de estrutura informacional: da frase ao discurso.....9
3. Os conceitos de tópico e foco.....13
   3.1.O conceito de tópico.....14
   3.2.O conceito de foco.....19
4. Os conceitos de dado e novo.....21
5. Tipos de tópico no discurso.....27
6. O estatuto informacional dos referentes no discurso.....31

**Capítulo 2 – Proposta de anotação semântica de um discurso oral**.....39

1. Apresentação do *corpus*.....42
2. Metodologia.....42
   2.1.Metodologia geral.....42
   2.2.Método de análise dos exemplos do corpus.....43
3. Taxonomia de tópicos para a anotação discursiva.....46
   3.1.Continuing topic ⇨ CONT.....47
   3.2.Familiar topic ⇨ F.....49
   3.3.Shifting topic ⇨ SHIFT.....52
      3.3.1. Rough-shifting topic ⇨ SHIFT (RSHIFT).....52
      3.3.2. Smooth-shifting topic ⇨ SHIFT (SSHIFT).....54
   3.4.Contrastive topic ⇨ C.....57
      3.4.1.Contrastive 1 ⇨ C (C1).....57

3.4.2. Contrastive 2 ⇨ C (C2).....59
4. Apresentação de discussão dos resultados.....60
4.1.Apresentação dos resultados.....60
4.2.Breve discussão dos resultados.....64

Considerações finais.....67
Referências bibliográficas.....71

# INTRODUÇÃO

No decorrer do seu quotidiano, os falantes utilizam as trocas verbais como um meio de transmissão de informação e de partilha de conhecimentos sobre o mundo que os rodeia. Deste modo, o discurso deve ser organizado para que os interlocutores compreendam os conteúdos bem como as mudanças das características das diversas entidades e eventos referidos. Neste sentido, a estrutura informacional – isto é, a forma como o discurso se encontra organizado – representa as relações entre as entidades do mundo exterior e as relações entre os elementos linguísticos que permitem a compreensão do enunciado (cf. Myackykov, Garrod & Schepeers 2009:2).

Grosz & Sidner (1986:177) consideram que o discurso é o resultado da interacção de três componentes: a estrutura linguística (os elementos utilizados pelo locutor para veicular a informação e as relações que se estabelecem entre eles), a estrutura intencional (a intenção/objetivo que o locutor quer atingir ao produzir a mensagem) e o estado de atenção (capacidade que cada falante possui de reconhecer e organizar a informação tendo em conta a sua importância no discurso). Segundo Grosz & Sidner (1986:177) são estas três componentes que permitem ao interlocutor determinar o significado do que foi dito e o seu contributo para o discurso.

Durante a produção do discurso, o falante utiliza a sua capacidade de atenção e de reconhecimento para organizar a informação consoante o seu grau de relevância (Myackykov, Garrod & Schepeers 2009:2), ou seja, os elementos são apresentados de acordo com o seu contributo e com a sua importância para o discurso.

As várias línguas do mundo apresentam formas distintas de marcação da informação relevante, tal como Myackykov, Garrod & Scheepers (2009:2) afirmam: "[...] the languages develop and provide devices that are specifically responsible for this representation". Assim, no caso da Língua Portuguesa, um dos meios utilizados é a distinção de um tópico e de um foco dentro da frase. Esta divisão permite, portanto, distinguir a informação que o locutor pretende destacar da restante informação. Contudo, para que os falantes consigam identificar o tópico e o foco dentro de um enunciado é necessário recorrer ao contexto em que a frase se encontra inserida (contexto de enunciação), de modo a que seja possível relacionar os tipos de informação (Barbosa 2005:341).

O contexto de enunciação é um fator de grande importância na análise semântica, dado que a produção discursiva encontra-se sempre relacionada com as coordenadas enunciativas *quem*, *o quê*, *onde* e *porquê* (Degand & Simon 2009:5). Van Dijk (1997:196-197, referido por Degand & Simon 2009:5) considera, por isso, que as representações discursivas são gradualmente construídas ao longo do discurso pelos falantes e que formam parte inerente do contexto à medida que este se vai desenvolvendo.

Assim, o tema desta dissertação relaciona-se com os diferentes comportamentos que o tópico pode assumir no contexto de discurso oral espontâneo, tendo como objetivos centrais i) a delimitação de uma taxonomia de tópicos; ii) a aplicação dessa taxonomia a uma amostra do corpus CPE FACES através de um processo de anotação; iii) a aferição do grau de eficácia da etiquetagem para a caracterização do corpus selecionado; iv) a validação dessa etiquetagem para os corpora inscritos no projeto.

Os dados utilizados na realização deste estudo pertencem ao projeto *COPAS – Contraste e Paralelismo na Fala*, que tem como objetivo determinar o modo como os diferentes tipos de informação são veiculados pelos falantes no discurso oral. Este projeto conta com a participação de investigadores das áreas linguísticas da semântica, sintaxe e prosódia, no sentido de estabelecer interface entre estes domínios para o estudo das questões centrais que ocupam o projeto, nomeadamente a determinação da forma como a entoação e os recursos sintáticos são usados pelos falantes, como se relaciona a estrutura da informação com esses aspetos, e a arquitetura da gramática.

O trabalho desenvolvido quer no âmbito do projeto COPAS, quer nesta dissertação, pretende ainda procurar e determinar semelhanças – e também diferenças – entre os resultados obtidos através da análise do *corpus*, bem como determinar de que forma interagem as diferentes áreas da linguística.

A dissertação apresentada encontra-se organizada em dois capítulos. O primeiro centra-se na delimitação do enquadramento teórico para a proposta de anotação do *corpus* em análise, incidindo essencialmente sobre questões que se relacionam com o discurso, a estrutura informacional, os conceitos de tópico e foco, e de dado e novo, além de proceder ao levantamento de uma taxonomia de tópicos discursivos. O segundo capítulo inclui os procedimentos de determinação das categorias semânticas para a anotação do corpus, parâmetros de anotação do corpus, o método que permitiu a análise dos exemplos, os resultados e uma breve discussão dos mesmos, na qual se levantam várias hipóteses que possam explicar os resultados obtidos e o comportamento dos diferentes tipos de tópico anotados. Por fim, apresentam-se as considerações finais e as referências bibliográficas consultadas durante o desenvolvimento do trabalho.

# CAPITULO 1

## Tópico, foco e estrutura informacional no discurso: alguns fundamentos teóricos

> Any model of discourse is in need of identifying its underlying units, the so-called 'building blocks' of the discourse structure (Degand & Simon 2009).

Partindo deste princípio, definimos como objeto central deste capítulo o enquadramento teórico que subjaz à proposta de etiquetagem para anotação semântica da amostra do *corpus* oral CPE-Faces, que será apresentada de forma circunstanciada na segunda parte deste trabalho. Por conseguinte, apresenta como objetivos centrais: i) a delimitação e explicitação dos conceitos relevantes no quadro da análise da organização informacional do discurso oral em análise; ii) a apresentação de diferentes propostas para a delimitação do tópico discursivo na amostra do *corpus* considerada; iii) a relação entre a descrição linguística do fenómeno em estudo e a sua operacionalização ao nível do trabalho de anotação.

Daqui decorrem duas observações preliminares de grande importância. Por um lado, tendo em conta que a anotação discursiva se realiza em *interface* com a anotação sintática e prosódica do mesmo *corpus*, através da utilização de um conjunto de etiquetas que procuram descrever o comportamento de segmentos discursivos em cada uma das áreas de análise consideradas, a descrição realizada neste capítulo circunscreve-se exclusivamente ao domínio da consideração da estrutura informacional, do tópico e do foco numa perspetiva semântico-discursiva, o que significa assumir, desde o início, a não sobreposição necessária entre as diferentes áreas no que se refere ao significado dos termos utilizados, bem como ao escopo discursivo ao qual se aplicam. Por outro, a reflexão apresentada não tem pretensão de ser exaustiva, o que seria impossível, dadas as limitações deste trabalho e, sobretudo, a complexidade do tema e a extensão da literatura

a ele já consagrada, mas pretende apenas sustentar e fundamentar as opções teóricas e metodológicas tomadas.

## 1. Discurso e universo discursivo

O conceito de discurso é definido de forma distinta em diversas propostas teóricas em função dos princípios e objeto ao qual se se refere.

No contexto deste estudo, considera-se o discurso como uma prática verbal que assume sempre um valor interindividual, consistindo, ao nível da receção e da produção, um processo cognitivo complexo que envolve diferentes tipos e níveis de operações e conhecimentos.

Como afirma Charaudeau (1997 :6),

> Un texte est le résultat d'un acte de langage produit par un sujet dans une situation d'échange sociale contractuelle. Du fait qu'il est un acte de langage, il se caractérise par les propriétés générales de tout fait langagier à savoir sa matérialité signifiante (...) organisée en systèmes, ses règles de formation et construction linguistique (...), ses procédés d'organisation discursive. Du fait qu'il est produit dans une situation contractuelle, il dépend pour sa signification de ce qui caractérise une situation à savoir : une finalité-visée énonciative, une identité des partenaires de l'échange, un propos comme contenu thématique de l'échange, un dispositif particulier comme circonstances matérielles de l'échange of a discourse provides indispensable cues to its structure.

Por isso, o discurso não pode ser analisado tendo apenas em conta as unidades linguísticas que são produzidas. Na verdade, a conversação está dependente de vários fatores, entre os quais o estatuto do locutor e do interlocutor, o contexto de enunciação, a relação entre os interlocutores, que são determinantes para o sucesso e a coerência dos conteúdos veiculados. Este conjunto de condições, que intervém sempre na produção de qualquer discurso, constitui aquilo a que Lambrecht (1994:36) designa por universo discursivo, um modelo, no qual os fatores acima mencionados (e outros) podem ser agrupados em duas grandes categorias.

A primeira categoria diz respeito aos elementos externos ao texto/discurso, isto é, intervenientes do discurso (locutor e interlocutores) e à situação de enunciação (momento, espaço e circunstâncias em que o discurso é produzido). Os elementos externos ao texto/discurso não se estabelecem apenas a partir de representações discursivas, dado que podem estar presentes na própria situação de comunicação, como acontece com as expressões deíticas. Este tipo de expressões permite que o falante designe diretamente "elements of the text-external world by «pointing» to them [...]"(Lambrecht 1994:38). Neste sentido, é possível encontrar algumas expressões deíticas que designem o locutor e o interlocutor (como por exemplo os pronomes pessoais *eu*, *tu*), o momento da produção do discurso (como é o caso dos advérbios de localização temporal *hoje*, *agora*, *ontem*, *amanhã*) e o local onde o discurso é produzido bem como outras localizações que forem sendo mencionadas pelos interlocutores (por exemplo as expressões de localização espacial *aqui*, *ali*).

A segunda categoria apresentada por Lambrecht (1994:37) corresponde aos elementos internos do texto/discurso e é constituída pelas expressões linguísticas (sejam elas frases ou palavras) e os significados que lhes estão associados. Esta categoria relaciona-se com o lado abstrato das representações mentais criadas nas mentes dos interlocutores durante o processo de comunicação, sendo através da manipulação destas representações que se veicula informação.

A manifestação destas entidades depende da forma como as suas representações se estabelecem no discurso, o que se relaciona de forma estreita com a sua organização informacional e, por essa via, com a estrutura informacional do discurso.

## 2. O conceito de estrutura informacional: da frase ao discurso

O conceito de estrutura informacional (EI) tem sido entendido de formas variáveis em função do quadro teórico em que é definido e explorado. Na impossibilidade de recuperar neste âmbito todas as propostas formuladas nesse universo teórico, convocamos apenas as representações desse conceito que de forma mais nítida articulam a estrutura da informação com as condições nas quais se verifica o fluxo informacional, não só a nível frásico, mas sobretudo a um nível discursivo, aquele que nos concerne.

O contributo do Círculo Linguístico de Praga para o tratamento da estrutura da informação é extremamente importante, salientando-se os nomes de, entre outros, Mathesius (1929), Firbas (1964) e Daneš (1974). Resumidamente, as propostas dos dois últimos autores citados vêm no seguimento de Mathesius (1929), que, reformulando o contraste gramatical entre sujeito e predicado, introduziu o contraste entre tema e rema, entendidos numa perspetiva de enfoque comunicativo. Firbas deu continuidade a este enfoque, numa articulação entre o conceito de estrutura informacional e o de dinamismo comunicativo. O autor define-o como um processo que determina como a informação é organizada na frase, mostrando como cada unidade de um enunciado contribui dinamicamente para a produção global da frase, gerando um determinado efeito comunicativo. Embora tenha uma relação mais ou menos visível com os níveis sintático e semântico, não implica correspondência absoluta entre os elementos que correspondem a cada um destes níveis. Na sua essência, o dinamismo comunicativo reflete o grau de importância atribuída a cada segmento da frase, que pode ser dividido em duas partes: o tema (que falamos) e rema (o que dizemos). O primeiro deve conter os elementos conhecidos e aceites, enquanto o segundo apresenta novas informações, é geralmente aceite que o sujeito, mesmo que represente um meio muito importante de conectividade, tem um nível comunicativo baixo quando comparado com que o rema, que é responsável pela transmissão de nova informação no texto. Daneš, por sua vez, expande o estudo da dinâmica comunicativa da frase para unidades mais complexas e estabelece uma relação entre tema e rema e informações conhecidas e novas informações, analisando os princípios que sustentam a progressão temática do texto.

Halliday (1967), que usou pela primeira vez o termo, considera que a estrutura informacional de um texto é constituída por uma sequência de unidades informativas que não correspondem necessariamente aos constituintes presentes na estrutura sintática, pelo que postula, ao contrário dos autores anteriores, que a informação estrutural deve ser analisada num nível distinto do da interface sintaxe-semântica (cf. Heusinger 2002 e Barbosa 2005). Desta forma, o falante é livre de determinar os limites e a organização interna da estrutura informacional, visto que as unidades informacionais apresentam uma estrutura própria. Ainda segundo Halliday (1967:200), a estrutura informacional é realizada fonologicamente pela distribuição do texto em grupos tonais. O núcleo destes

grupos é composto por acentos tonais, que têm como função marcar os focos informacionais da frase. Neste sentido, a estrutura informacional é composta pela distribuição das unidades informativas (correspondentes à estrutura temática) e pela organização interna de cada uma dessas unidades, designando-se cada uma dessas estruturas, respetivamente, por estrutura temática e 'givenness'. A primeira corresponde à ordenação linear das unidades informacionais em tema-rema, enquanto a segunda refere aquilo de que se fala nessas unidades.

Do ponto de vista de Lambrecht (1994:5), a estrutura informacional é uma componente da gramática da frase, na qual proposições como as representações concetuais dos estados de coisas

> are paired with lexicogrammatical structures in accordance with the mental states of interlocutors who use and interpret these structures as units of information in given discourse contexts.

Este autor refere-se à organização da frase como estrutura focal, ou seja, uma descrição estrutural anotada, que integra a forma fonológica e a forma semântica, e na qual os constituintes como foco e tópico são marcados. Além disso, a estrutura informacional preocupa-se ainda com fenómenos psicológicos, como as hipóteses que o falante coloca sobre os conhecimentos do ouvinte (Lambrecht 1994:3), tendo como função considerar as condições e os contextos discursivos em que as unidades de informação analisadas são produzidas (Lambrecht 1994:5). Assim, para Lambrecht (1994:6), o estudo da estrutura informacional envolve ainda três categorias:

i. pressuposição e asserção, que se relacionam com a divisão das proposições em porções que o falante assume serem ou não conhecidas pelo ouvinte;
ii. identificação e ativação dos referentes, relacionadas com as hipótese que o falante coloca sobre o estatuto das representações mentais presentes na mente do ouvinte;
iii. tópico e foco, que se relacionam com a avaliação que o falante faz da previsibilidade e da não previsibilidade das relações entre as proposições e os seus elementos nas diferentes situações discursivas.

A estrutura informacional pode manifestar-se através da prosódia, de marcadores gramaticais especiais, de constituintes sintáticos e da sua posição na frase, e de certas escolhas lexicais. Deste modo, a estrutura informacional constitui uma parte da gramática da frase que combina caraterísticas semânticas, pragmáticas, prosódicas e sintáticas.

Por sua vez, Prevost (1996) define a estrutura informacional de uma forma similar:

> Information Structure refers to the organization of information within an utterance. In particular, it defines how the information conveyed by a sentence is related to the knowledge of the interlocutors and the structure of their discourse. Sentences conveying the same propositional content in different contexts need not share the same information structure. That is, information structure refers to how the semantic content of an utterance is packaged, and amounts to instructions for updating the models of the discourse participants.

Wallace Chafe (1976) assume uma vertente de natureza cognitiva na sua abordagem da análise do modo como se processa o fluxo informacional, introduzindo a noção de empacotamento ('packaging') da informação, que é um elemento relevante para o entendimento da EI.

A este nível, Chafe salienta:

> Language functions effectively only if the speaker takes account of such states in the mind of the person he is talking to. It is only, for example, when the speaker adjusts what he says to what he assumes the addressee is thinking of at that moment that is message will be readily assimilated by the addressee.

Neste contexto, o falante tem de atender primeiramente à forma como a mensagem é enviada e só, depois, à mensagem propriamente dita[1], situando-se a análise mais ao nível da descrição da forma como um falante acomoda o seu discurso aos estados temporários da memória do seu interlocutor e menos ao da memória de longo prazo, embora isso também se verifique sob certas condições. O empacotamento da informação está, na

[1] Apesar da produtividade da proposta de Chafe, Krifka aponta alguns problemas decorrentes da sua definição de empacotamento da informação, o que tem consequências para a forma como a informação é organizada e veiculada: "One problem with Chafe's approach is that there are aspects of optimization of the message that, on the one hand, respond to the temporary state of the addressee, but on the other also affect the message itself, and hence cannot be treated as pure packaging" (2007:14).

perspetiva de Chafe, associado a um conjunto de elementos, que têm a sua especificidade, mas contribuem para o fluxo informacional do discurso: 'givenness', contraste, definitude, sujeitos e tópicos.

Do exposto, concluímos, de acordo com kruijff-Korbayová & Steedam (2003) que é possível associar ao conceito de estrutura informacional duas distinções básicas, que, sendo diferenciadas, permitem, na sua globalidade, uma melhor compreensão do fenómeno da estrutura da informação no discurso. De facto, os autores afirmam:

> They all draw at least one of these distinctions: (i) A "topic/comment" or "theme/rheme" distinction between the part of the utterance that relates it to the discourse purpose, and the part that advances the discourse; (ii) a "background/contrast" or "given/new" distinction, between parts of the utterance (…) which contribute to distinguishing its actual content from alternatives the context makes available" (kruijff-Korbayová & Steedam (2003: 251).

Nas secções seguintes, procede-se a uma breve reflexão sobre cada uma das dicotomias aqui enunciadas.

## 3. Os conceitos de tópico e foco[2]

A primeira proposta de organização da informação na frase parte de Hermann Paul e corresponde à divisão sujeito psicológico/predicado psicológico. Enquanto o sujeito psicológico diz respeito à "quantidade de ideias existente no consciente do que fala, do que pensa […]", o predicado psicológico corresponde à quantidade restante de ideias presentes na frase que se associa ao sujeito psicológico (Paul [1980] 1983:1333).

Firbas (1964), por sua vez, apresenta a relação tema/rema como uma relação hierárquica. Deste modo, a frase seria constituída por um elemento com o mais baixo

[2] Para uma distinção mais aprofundada, ver Barbosa 2005. Relativamente à dicotomia foco/pressuposição – a última apresentada no trabalho de Barbosa (2005:345) –, salientam-se as investigações de Chomsky (1971, referido por Barbosa 2005:345) e de Jackendoff (1972, referido por Barbosa 2005:345).

grau de dinamismo comunicativo, que corresponderia ao tema, e por um elemento com o mais alto grau de dinamismo comunicativo, que corresponderia ao rema.

No que diz respeito à divisão tópico/comentário destacam-se as investigações de Daneš (1970:134), que considera que o tema (ou tópico) reflete o assunto que se aborda na frase e o rema (ou comentário) contém as informações relativas ao assunto referido no tema, correspondendo, respetivamente o tema/tópico aos elementos conhecidos, que constituem o ponto e partida, e o rema/comentário, aos elementos novos, constitutivos do fulcro do enunciado (cf. Barbosa 2005).

É também a partir das abordagens de Daneš, segundo Barbosa (2005:344), que surge a divisão tópico/foco, considerada pela Moderna Escola de Praga. Aos elementos constitutivos do enunciado a que Daneš chama "ponto de partida" e "núcleo" do enunciado, a Escola de Praga designa por tópico e foco (Barbosa 2005:344).

Tal como os investigadores da Escola de Praga, Halliday (1985:39) recorre à terminologia tema/rema, afirmando que o tema "is one element in a particular structural configuration which, taken as a whole, organizes the clause as a message" e, por isso, a mensagem é composta por um tema combinado com um rema (Halliday 1985:39).

### 3.1. O conceito de tópico

Lambrecht, cujo trabalho seguimos de perto para a caracterização de tópico, considera que o tópico de uma frase "is the thing which the proposition expressed by the sentence is ABOUT" (1994:118). Desta forma, um referente pode ser interpretado como tópico de uma proposição se, num determinado contexto discursivo, essa mesma proposição expressar informação relevante que permita ao interlocutor aumentar o seu conhecimento em relação ao referente.

O tópico da frase estabelece-se a partir da relação entre um referente e uma proposição. Segundo Lambrecht (1994:150), o tópico encontra-se relacionado com a pressuposição pragmática, dado que

> since the topic is the already established «matter of current concern» about which new information is ADDED in an utterance, for a proposition to be construable as

> being about a topic referent this referent must evidently be part of the pragmatic presupposition, i.e. it must already be «under discussion» or otherwise available from the context.

Para Duarte (2003:118), o tópico apresenta uma função cognitiva que tem como finalidade

> selecionar e ativar um elemento existente na memória passiva do alocutário/leitor/ouvinte, transferindo-o para uma memória ativa em que possa ser combinado com novos elementos cognitivos introduzidos pelo comentário.

Neste sentido, os referentes dos tópicos devem estar acessíveis ao locutor, podendo ser encontrados no discurso anterior ou no contexto discursivo em que o texto é produzido e interpretado (Duarte 2003:118).

Lambrecht destaca dois tipos de tópico, tendo em conta a forma como é apresentado o constituinte que designam. Para além da noção de tópico já referida, em que a proposição em questão é construída como dizendo respeito a um determinado referente, refere ainda outro tipo a que chama expressão do tópico, e que corresponde à expressão linguística que designa o referente tópico da frase. Deste modo "a constituent is a topic expression if the proposition expressed by the clause with which it is associated is pragmatically construed as being the referent of this constituent" (Lambrecht 1994:131). A expressão do tópico através de uma expressão linguística deve designar um referente discursivo, pelo que a entidade referida na proposição deve fazer parte do universo de discurso dos interlocutores, de modo a poder ser designada pela expressão do tópico.

No que diz respeito à manifestação linguística dos tópicos na frase, Lambrecht salienta que

> while a topic expression always necessarily designates a topic referente, a referent which is topical in a discourse is not necessarily coded as a topic expression in a given sentence or clause. (1994:130)

Em suma, os tópicos devem estar inseridos na pressuposição, ou seja, devem ser elementos constituintes da pressuposição pragmática invocada pela frase (Lambrecht

1994:154). Neste sentido, pressupõe-se a existência dos referentes das expressões dos tópicos.

Lambrecht (1994:144-147) destaca que em algumas frases o elemento marcado como tópico não corresponde ao sujeito. Desta forma, o autor refere dois tipos de construções que podem ilustrar este facto: as frases téticas, que são caraterizadas pela ausência de uma relação tópica entre a proposição e o argumento que desempenha a função de tópico na construção tópico-comentário, e as construções topicalizadas, nas quais um constituinte que não apresenta a função de sujeito é marcado como uma expressão tópica através do seu posicionamento no início da frase (lugar ocupado, normalmente, pelo sujeito tópico).

Ainda segundo o autor, a topicalização das frases tem como funções marcar o referente de um sintagma nominal como um tipo de tópico particular na proposição, na qual desempenha o papel de argumento, e marcar a proposição como veiculando informação sobre o referente desse tópico. Existem também frases que apresentam dois ou mais tópicos. Estas construções não só expressam informação sobre os referentes dos tópicos, como também transmitem informação sobre a relação que estes últimos estabelecem como argumentos na proposição. É por esta razão que Lambrecht (1994:148) afirma que "the proposition can be said to be ABOUT this relation is because the existence of some relation between two [or more] topics is already established before the sentence is uttered" (1994:148).

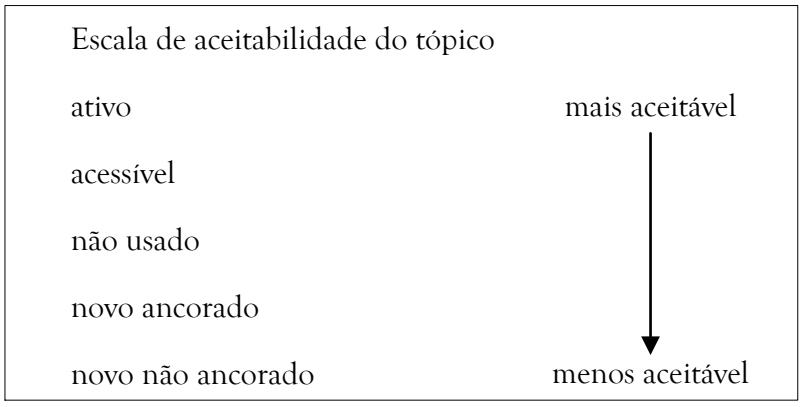
Para que um referente seja interpretado como tópico de uma proposição, de modo a que essa proposição possa ser interpretada como apresentando informação relevante sobre o tópico, o referente do tópico deve possuir certas propriedades de ativação. No entanto, "the fact that a particular referent has the activation properties required for a topic function in a sentence does not entail that it must be coded as a topic" (1994:163) dado que a acessibilidade cognitiva é uma condição necessária mas não suficiente para o uso de certas construções, como é o caso da topicalização.

Neste sentido, o falante toma uma decisão comunicativa ao efetuar a seleção do tópico da frase, ou seja, estabelece uma espécie de ponto de partida para a entidade sobre a qual pretende veicular informação. Contudo, esta decisão é condicionada

pelas hipóteses colocadas pelo falante relativamente ao estatuto do referente do tópico na mente do ouvinte, no momento em que a unidade informacional é produzida (Lambrecht 1994:163).

De modo a que um enunciado possa expressar informação sobre um tópico, o falante deve ter em conta que, no contexto de enunciação, considera esse enunciado relevante no que diz respeito ao seu tópico. Assim, "for a statement to be relevant with respect to a topic, this topic itself must be of CURRENT INTEREST" (Lambrecht 1994:164), isto é, deve estar estabelecido no discurso ou estar relacionado com outro elemento referido anteriormente.

Como forma de minimizar as dificuldades de interpretação causadas pela identificação do estado dos referentes dos tópicos na mente dos ouvintes, Lambrecht postula uma escala de aceitabilidade, transcrita no esquema 1, a partir da qual se torna possível medir "[...] the degree of pragmatic well formedness of a sentence containing a topic expression by the topic referent [...]" (Lambrecht 1994:165).




Esquema 1: Escala de acessibilidade do tópico (Lambrecht, 1994: 165)

Tendo em conta esta escala, Lambrecht refere que as frases mais aceitáveis do ponto de vista cognitivo e que são facilmente processadas apresentam os referentes dos tópicos que se encontram ativados no discurso, ocupando, por isso, um lugar mais "alto" na escala de aceitabilidade. O esforço mental necessário para o processamento deste tipo de frases torna-se menor, pois o ouvinte não necessita de realizar a tarefa de encontrar o referente do tópico através do seu armazenamento na memória a longo prazo ou da realização de inferências que permitam a sua

identificação. Estes referentes são designados por "preferred topics" (Lambrecht 1994:165).

As expressões tópicas mais difíceis de interpretar, ainda que sejam aceitáveis, apresentam referentes acessíveis, em relação aos quais,

> [...] the mental effort necessary to interpret the proposition which expresses the new information about the topic must be performed simultaneously with another processing task, the task of remembering, inferring, or otherwise determining the referent of the topic expression (Lambrecht 1994:166).

Nos casos em que a informação nova presente na frase é expressa através de um referente tópico não usado no discurso (ou seja, inativo), o esforço mental requerido é relativamente alto, dado que, para além de processar a informação proposicional sobre um determinado tópico, o ouvinte deve determinar o referente do tópico que não se encontra disponível no discurso prévio. Os tópicos não aceitáveis correspondem aos referentes que não são identificados pelo ouvinte no momento em que a informação é transmitida. Nestas situações, se o ouvinte não está apto para identificar mentalmente o referente do tópico, não consegue, também, determinar as suas condições de verdade.

A conceção de tópico aqui apresentada subsume simultaneamente aspetos que tipicamente são tratados de forma complementar, na medida em que podem ser associados ao conceito de tópico de frase, que constitui a expressão que exprime alguma coisa acerca do referente da frase, sujeita a condicionalismos de natureza sintática, semântica e prosódica, e tópico discursivo, definível em função discurso em construção. Reinhart estabelece de forma clara essa dupla interpretação, ao afirmar que

> Sentence topics must correspond to an expression in the sentence (the topic expression). Discourse topics are topics of larger units and they can be more abstract, though they do not have to be (1981: 54).

### 3.2. O conceito de foco

Embora este conceito não esteja no escopo deste trabalho, parece-nos pertinente considera-lo por constituir um elemento central na organização do fluxo informacional, juntamente com o tópico, e permitir uma análise do tópico em contraposição com as suas propriedades.

A definição do termo foco difere consoante o autor e a perspetiva de análise adotada, dado que os investigadores combinam noções e caraterísticas sintáticas, semânticas e pragmáticas: "there is no general agreement in the literature about what is that defines «focus», and it is difficult to point out a single basic (pre-theoretic) intuition that lies behind this notion [...] " (Kadmon 2000:252-253).

Rooth (1996:271) apresenta o foco como uma caraterística presente na frase, que possui uma interpretação ao nível semântico e pragmático e uma interpretação ao nível fonético e fonológico.

Para Lambrecht (1994), o foco corresponde à "parte da informação que o locutor assume não ser partilhada por si e pelo alocutário [...]", que se relaciona com o conceito de pressuposição, ou seja, com "a parte da informação que o locutor assume ser partilhada por si e pelo alocutário" (Barbosa 2005:340). A este respeito, Joaquim Barbosa afirma, assim, que o foco está relacionado com a "ideia de novidade, de informação nova", ou seja com "a informação que contrasta de algum modo com porções da informação antecedente, pré-existente, no contexto de enunciação (background)" (2005:341). Desta forma, o foco apresenta efeitos semânticos e pragmáticos, manifestando-se na sintaxe.

Segundo Halliday (1967b:204, referido por Winkler 1997:116), "what is focal is «new» information, not in the sense that it cannot have been previously mentioned, although it is often the case that it has not been, but in the sense that the speaker presents it as not being recoverable from the preceding discourse". O conceito de informação conhecida é, deste modo, tido como sendo recuperável quer anaforicamente, quer através da situação. Tendo em conta também as considerações de Halliday, nas quais a informação é considerada nova para o ouvinte, dado que é apresentada pelo falante

como sendo não recuperável no contexto (textual e situacional) precedente, Lambrecht carateriza o foco como um elemento que não é previsível nem pragmaticamente recuperável na frase (1994:207), dado que a informação da frase é transmitida através da relação que se estabelece entre os diversos elementos e as proposições – designada por relação focal em que estão inseridos e não a partir da atribuição das propriedades "novo" e "velho" aos conteúdos.

Bolinger (1954:152) apresenta o foco como "information point of the sentence", no qual o acento prosódico marca o elemento na frase onde se verifica a maior concentração de informação, que o ouvinte não é capaz de inferir sem indicações do falante.

Krifka (1991:33-34) toma como referência os conceitos e os critérios de Jackendoff (1972, referido por Krifka 1991:33-34) mencionando que o foco é uma caraterística atribuída a um nó da representação sintática. Para Krifka (1991:33) a caraterística da marcação do foco na frase pode estar associada a um elemento que funcione como operador do foco (por exemplo, os quantificadores). Esta associação é designada por "bound focus" (Krifka 1991:33).

Segundo Lambrecht (1994:218), a marcação do foco é um mecanismo formal que permite assinalar a relação entre um determinado constituinte pragmático e uma proposição. O objetivo da marcação do foco não é marcar um constituinte como sendo novo no contexto da enunciação, mas sim assinalar a relação que existe entre um dado elemento da proposição e a proposição propriamente dita. O autor refere, ainda, que um constituinte que não se encontra em foco faz parte da pressuposição, sendo que o domínio sintático através do qual se expressa o foco da pressuposição pragmática é designado por domínio do foco.

O domínio do foco é composto por constituintes que, quando acrescentados à pressuposição, produzem asserções (Lambrecht 1994:215). Como exemplos destes constituintes destacam-se os predicados ou argumentos e as proposições completas, que se inserem nas categorias frásicas (Lambrecht 1994:215).

O domínio do foco não pode ser expresso por categorias lexicais, dado que a estrutura informacional relaciona-se com a construção pragmática das relações entre entidades e estados de coisas nas várias situações discursivas (Lambrecht 1994:215). As

entidades e os estados de coisas são expressos sintaticamente através de categorias frásicas (Lambrecht 1994:215).

Com base nas caraterísticas que atribui para o foco, Lambrecht (1994) distingue a relação existente entre o tópico e a proposição e a relação que se estabelece entre o elemento em foco e a proposição. A primeira relação é previsível e recuperável pelo ouvinte, enquanto a segunda é tida como sendo uma relação não previsível e não recuperável pelo ouvinte, no momento da enunciação. Desta forma,

> the focus relation relates the pragmatically non-recoverable to the recoverable component of a proposition and thereby creates a new state of information in the mind of the addressee (Lambrecht 1994:218).

Relativamente às implicações semânticas do foco, Krifka (1991:34) destaca que a marcação de um dado elemento como sendo o foco origina a divisão da representação semântica da frase em duas partes: a parte que se encontra marcada como foco e a parte complementar que não se encontra em foco (Krifka 1991:34).

## 4. Os conceitos de dado e novo

A informação criada na mente dos interlocutores permite a partilha de conhecimentos entre ambos.

Chafe distingue o fluxo informacional subjacente a essa partilha da seguinte forma:

> Given (or old information) is that knowledge which the speaker assumes to be in the consciousness of the addressee at the time of the utterance. So-called new information is what speaker assumes he is introducing into the addressee's consciousness by what he says (1976:30).

Lambrecht (1994) postula que o conceito de informação deve ser distinguido da noção de significado, pois o significado de uma frase encontra-se expresso nas palavras individuais ou nas relações que se estabelecem entre elas, enquanto a informação é transmitida através de proposições e está dependente dos estados mentais dos

interlocutores. Através da transmissão de informação sobre uma determinada situação ou sobre um determinado estado de coisas, o falante influencia as representações que o ouvinte tem do mundo. Estas representações são compostas por um conjunto de proposições que o ouvinte conhece ou acredita ser verdadeiro no momento da enunciação. O conjunto destas proposições é identificado como sendo o conhecimento que o ouvinte possui e que consiste em "to have a mental picture of its «denotatum» not in the sense of «to know its truth»" (Lambrecht 1994:44). Neste sentido, o falante, ao partilhar com o ouvinte informação sobre uma dada situação, inicia uma mudança no estado de conhecimentos do ouvinte, facto que se prende com a adição de proposições ao seu conhecimento.

Para Dahl (1976), o estado do conhecimento do ouvinte no início do discurso corresponde à informação já conhecida (dada), enquanto a mudança no estado de conhecimentos do ouvinte (provocada pelas proposições partilhadas) corresponde à informação nova. Desta forma, "the addressee receives «new information» in the sense that he comes to know or believe more about the world than he did before" (Dahl 1976:38).

No entanto, Lambrecht (1994:44) destaca que a mudança do estado de conhecimentos do ouvinte operada pelo falante através da adição de novas proposições afeta apenas uma parte da representação mental que o ouvinte possui. A informação nova que o ouvinte recebe é adicionada ao conjunto de conhecimentos já existente na sua mente. A transmissão da informação requer, portanto, uma constante atualização das hipóteses que o falante coloca em relação ao estado de conhecimentos do ouvinte. Deste modo, a informação só é transmitida com sucesso se as hipóteses colocadas pelo falante em relação ao estado de conhecimentos do ouvinte estiverem corretas, ou seja, se a informação veiculada ainda não se encontrar alojada na mente do ouvinte (Lambrecht 1994:46). A esta condição Strawson (1964) chama Princípio da Suposição da Ignorância. Contudo, dado que o ouvinte nunca se encontra num estado total de ignorância, pois está constantemente a receber informação nova que é acrescentada aos seus conhecimentos já existentes, Strawson (1964) apresenta igualmente o Princípio da

Suposição do conhecimento, que refere que o grau de informatividade das proposições está dependente do conhecimento que o ouvinte possui.

Dahl (1976:38) considera, também, que a informação recebida pelo ouvinte pode estar correta ou incorreta, ou seja, pode ser classificada como sendo verdadeira ou falsa. Contudo, as representações mentais de eventos, estados ou situações presentes na mente do ouvinte, cuja informação é realizada através de proposições, não podem ser caraterizadas no que diz respeito aos seus valores de verdade. Na verdade, o estudo da estrutura informacional preocupa-se centralmente com os estados das representações presentes na mente do ouvinte, e não com as condições de verdade que lhes estão subjacentes (Lambrecht 1994:45).

Assim, na construção dos enunciados, o ouvinte apresenta os conteúdos através de formas lexicais ou gramaticais diferentes, conforme estes se encontrem ou não armazenados na mente do falante, pelo que a transmissão da informação resulta da relação entre os factos que são considerados novos para o ouvinte e os factos que este apresenta como conhecidos, sendo que o último tipo de informação pode encontrar-se implícito nas proposições (Lambrecht 1994:51)[3].

Do exposto, podem salientar-se dois elementos de análise complementar: por um lado, a distinção entre dado e novo no fluxo informacional discursivo e, por outro, o modo como essa articulação se produz, tendo em consideração o estatuto dos referentes discursivos e a sua expressão linguística.

[3] Desta forma, Lambrecht (1994:52-53) substitui os termos *informação nova* e *informação já conhecida* por *asserção pragmática* e *pressuposição pragmática*, respetivamente. A pressuposição pragmática corresponde, assim, ao conjunto de proposições invocadas pela frase, que o falante assume serem conhecidas pelo ouvinte no momento da informação; a asserção pragmática diz respeito à proposição que o ouvinte deve armazenar – e, portanto, ficar a conhecer – no seu registo discursivo depois da enunciação da frase. A realização de uma asserção implica a expressão de uma proposição pragmaticamente estruturada, ou seja, uma proposição que reflita o assunto denotado e as hipóteses que o falante coloca sobre o estado mental do ouvinte no momento da enunciação, através da indicação dos conteúdos tidos como informação já conhecida e dos que constam como informação nova. Neste sentido, Lambrecht (1994:65) considera que as frases apresentam estruturas pressuposicionais, cuja função é representar as pressuposições do falante. Estas estruturas pressuposicionais devem estar, também, de acordo com as situações pressuposicionais (Lambrecht 1994:65). Em suma, "presuppositional structures are them grammatical conditions on the appropriate use of words and constructions in given discourse situations" (Lambrecht 1994:65).

Neste âmbito, começamos por considerar a bipartição do fluxo informacional em termos distintos do que designamos como tópico frásico. Embora haja um vasto conjunto de trabalhos dedicado ao seu estudo, privilegiamos neste trabalho a proposta de Ellen Prince (1981, 1992) pelas seguintes razões: i) os postulados teóricos desta autora continuam a ser referência nas propostas atuais que procuram alargar o âmbito do estudo sobre o fluxo informacional do discurso em termos do dado e do novo; ii) a sua proposta ultrapassa a inadequação de uma bipolaridade informacional associada tradicionalmente à bipartição dado-novo; iii) a sua proposta tipológica é funcionalmente central na delimitação de etiquetas para efeitos de anotação semântica de um *corpus* discursivo complexo. Isso não significa, no entanto, que não se recorra aos contributos de outros autores no sentido de complementar o quadro teórico traçado e a proposta de etiquetagem realizada na segunda parte deste trabalho.

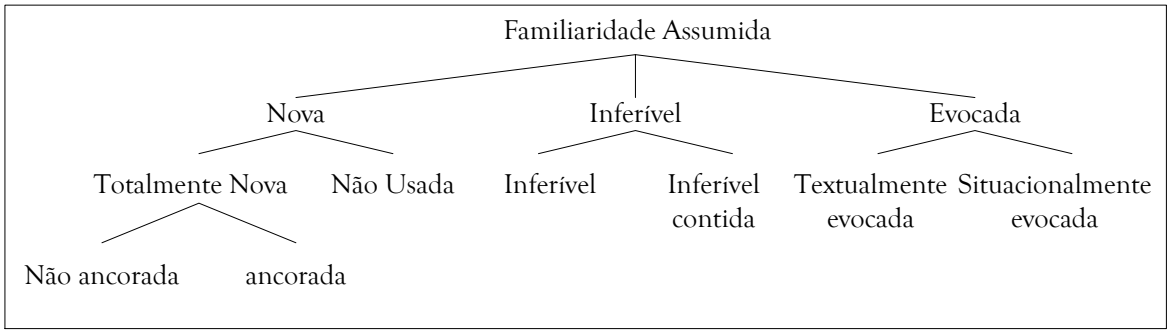
Prince (1981) parte dos postulados de Chafe (1976), que amplia para a formulação de uma proposta tipológica estruturada e menos restritiva, de modo a dar conta das variáveis ocorrentes no discurso ao nível da forma como se processa a organização em termos do estatuto das entidades discursivas integradas no fluxo discursivo.

A autora começa por apontar três níveis de 'givenness', que se fundam, respetivamente, nos princípios de predizibilidade/recuperabibilidade (Kuno 1972), saliência (Clark & Haviland 1977) e conhecimento partilhado, os quais, embora distintos, se interligam da seguinte forma:

> If a speaker assumes that the hearer can predict some particular item or items will occur in some particular position within the sentence, then the speaker must assume that it is appropriate that the hearer have some particular thing in his/her consciousness. And, if the speaker assumes that the hearer has some particular thing in his/her consciousness, then the speaker must assume that the hearer has some assumption or can draw some inference (1981: 231).

Assumindo que este modo de explicar a forma como se processa o dado se realiza no âmbito da relação falante/ouvinte, Prince questiona-se sobre os problemas levantados pela sua consideração na perspetiva ouvinte/falante, nomeadamente no que diz respeito ao conhecimento partilhado.

No sentido de resolver essas questões, propõe uma taxonomia que se funda no postulado da familiaridade presumida, através da qual introduz uma nova categoria de entidades no quadro do modelo discursivo, que é a das entidades inferíveis, que junta à categoria dos referentes discursivos novos e evocados. Desta forma, Prince classifica as entidades discursivas de acordo com o seu estatuto no texto/discurso, na taxonomia apresentada no esquema 2.




Esquema 2: Taxonomia da informação dada-nova (Prince 1981: 237)[4]

De acordo com a proposta de Prince (1981: 235-237), as entidades referidas assumem os seguintes valores:

i. Entidade nova: é introduzida pelo locutor e não pode ser recuperada no contexto antecedente. As entidades consideradas novas podem ser divididas em dois tipos:

   - completamente nova – quando o locutor cria uma nova entidade discursiva, podendo ser

     - ancorada – quando o sintagma nominal que representa a entidade se encontra ligado a uma outra entidade por meio de outro sintagma nominal

     - não ancorada – quando o sintagma nominal que representa a entidade não se encontra ligado a uma outra entidade por meio de outro sintagma nominal

[4] Seguimos a tradução proposta por Zamponi (2003).

- não usada – quando o interlocutor tem presente uma determinada entidade no seu esquema mental e apenas tem de a transpor para o modelo discursivo.

ii. Entidades inferíveis – quando o interlocutor consegue inferir uma entidade a partir de outras já existentes no contexto, podendo ser de dois tipos:
    - inferível propriamente ditas – quando o locutor consegue identificar o referente a partir de outras entidades presentes no discurso;
    - inferível contida – quando o referente que é inferido encontra-se contido no próprio sintagma nominal que o representa

iii. Entidades evocadas – os referentes evocados são aqueles que já se encontram presentes no modelo discursivo dos falantes, que podem ser
    - textualmente evocada – através do texto propriamente dito;
    - situacionalmente evocada – através da situação, isto é, do contexto extralinguístico.

A atribuição a uma entidade discursiva de cada um destes estatutos tem consequências na forma como é organizada e segmentada a informação veiculada, pelo que é de grande relevância para a análise do discurso.

No sentido de ultrapassar a inadequação associada a uma bipartição simples do discurso em termos de dado/novo, Prince (1992) reformula a sua proposta taxonómica, acrescentando como parâmetro de análise relevante duas categorias que se podem cruzar, a saber: dado/novo e dado para o ouvinte/novo para o ouvinte. Desta forma, salienta que o que é novo no discurso não é necessariamente novo para o ouvinte, ou seja, uma entidade nova no discurso pode não ser nova para o ouvinte. A esta distinção, acrescenta ainda uma outra, que consiste no facto de considerar que o grau de familiaridade em relação ao ouvinte ou ao discurso de uma determinada entidade pode ser vista como relativa ou absoluta.

A organização da estrutura informacional de um dado discurso processa-se em função da articulação da forma como as entidades nele configuradas são integradas, convivendo para tal informação de natureza linguística, pragmática e cognitiva.

## 5. Tipos de tópicos no discurso

Os tópicos têm vindo a ser definidos como correspondendo à informação já conhecida pelos sujeitos falantes (Reinhart 1982 referido por Braver 2007:2; Reinhart 1981:73). Na maioria dos casos, os tópicos representam informação já conhecida (Reinhart 1981: 73). No entanto, Reinhart (1981: 73) explica que sintagmas nominais com referentes idênticos podem desempenhar diferentes papéis informacionais, isto é, podem estar ou não disponíveis na mente dos falantes. Segundo esta autora (1982:4) os tópicos não podem ser definidos apenas como correspondendo à informação já conhecida no discurso, dado que é possível introduzir entidades que se encontram mencionadas no discurso prévio, mas que não correspondam ao tópico do discurso corrente.

As conexões existentes entre frases permitem introduzir um referente no discurso, tornando-o, de seguida num constituinte topicalizado através de, por exemplo, relações anafóricas (Reinhart 1981:75-76). Contudo, apesar dos enunciados se poderem apresentar conectados entre si, muitas vezes através de conectores semânticos, os seus tópicos não têm obrigatoriamente de se encontrar ligados referencialmente (Reinhart 1981:76).

Reinhart (1982, referido por Braver 2007:1) menciona a existência de dois tipos de tópico: os tópicos frásicos e os tópicos discursivos. Os tópicos frásicos correspondem geralmente a uma expressão na frase, enquanto os tópicos discursivos podem ser de natureza mais abstracta e correspondem a unidades maiores (Reinhart 1982, referido por Braver 2007:1).

Dos dois tipos de tópico apresentados, a autora (1982, referida por Braver 2007:1) destaca como objeto da sua investigação os tópicos frásicos, sendo que estes podem ser definidos através da estrutura linguística (posição na frase, estrutura gramatical e entoação) e de termos psicológicos (as intenções e os interesses dos falantes (Reinhart 1982, referida por Braver 2007:1)).

A distinção entre estes dois tipos de tópicos encontra-se também em Krifka (2001:1) que define os tópicos discursivos como sendo "what a part of a sentence is about" e os tópicos frásicos como "what is predicated about an entity in the sentence".

Uma forma de marcação do tópico é a deslocação à esquerda do constituinte com função de tópico. Para Reinhart (1981:73) esta forma de construção dos enunciados pode estar relacionada com a marcação do acento prosódico ou, segundo Keenan e Schieffelin (1976, referido por Reinhart 1981:77) pode originar a introdução de novos tópicos. Do ponto de vista de Runge (2008:2), esta construção frásica pode também servir para introduzir uma determinada entidade – que pode ou não já ter sido referida no discurso prévio – no discurso.

Reinhart (1981:64) aponta como principal caraterística das construções com deslocação à esquerda de constituintes o facto de estas só poderem ser usadas como tópico não podendo, assim, receber a acentuação prosódica caraterística da marcação do foco.

A análise de construções com deslocação à esquerda de constituintes encontra-se, também, no trabalho de Ward e Birner (2001:29-31). Estes autores (2001:13-18; 29-31) distinguem as construções denominadas preposing, em que o constituinte ocorre à esquerda e a informação está relacionada com o discurso prévio através de vários tipos de relações, como por exemplo tipo/subtipo, entidade/atributo, parte/todo (Ward e Birner 2001:13), das construções com deslocação à esquerda de constituintes, nas quais a posição do constituinte marcada é ocupada por um pronome correferencial (Ward e Birner 2001:29) e a informação veiculada pode não corresponder ao discurso prévio (Ward e Birner 2001:30).

Desta forma, Prince (1997, referido por Ward e Birner 2001:30-31) destaca três tipos de construções com deslocação à esquerda: "simplifying LDs", que envolvem entidades novas para o discurso e que ocupam uma posição sintática desfavorável para o papel que representam na conversação; "triggering a poset inference" e "amnestying na island violation", que envolvem informação referida no discurso prévio.

Prince (1991:9) refere, ainda que, apesar das construções com deslocação á esquerda de constituintes estarem associadas à produção de tópicos contrastivos, o conceito de contraste encontra-se relacionado com a ideia de saliência e com a escolha de um determinado elemento de um conjunto de alternativas. Neste sentido, a autora (1991:9) defende que o contraste "arises when alternate members of some salient set are evoked

and when there is felt to be a salient opposition in what is predicated of them", não sendo, por isso, um resultado da posição do constituinte com função de tópico.

Através da leitura bibliográfica efetuada, foi possível perceber autores como Frascarelli e Hinterhölzl (2002), Calhoun, Nissim, Steedman e Brenier (2005) e Prince e Walker (1993) observaram o comportamento dos tópicos no discurso e, tendo também em conta o seu contributo para a interpretação da conversação, definiram alguns tipos de tópico. Desta forma, encontram-se assim quatro tipos de tópico, que podem ser considerados como principais ou como a base de uma distinção da função dos constituintes marcados como tópico no enunciado: continuing topic (Givón 1983, referido por Frascarelli e Hinterhölzl 2002:2), familiar topic (Chafe 1987, referido por Frascarelli e Hinterhölzl 2002:2), shifting topic (Givón 1983, referido por Frascarelli e Hinterhölzl 2002:3) e contrastive topic (Kuno, 1976, Büring 1999, referido por Frascarelli e Hinterhölzl 2002:2).

O continuing topic verifica-se quando "a familiar topic is textually given and d-linked with a pre-established aboutness topic" (Givón 1983, referido por Frascarelli e Hinterhölzl 2002:2). Este tipo de tópico tem como principal função manter o mesmo referente ao longo do discurso, de forma a atribuir-lhe alguma coerência.

(1) [...] nesse dia surgiu o primeiro sinal de que eu ia ter o parto [...]
(exemplo 2 do corpus)

O tipo de tópico familiar (Chafe 1987, referido por Frascarelli e Hinterhölzl 2002:2) pode ser considerado característico da interação verbal espontânea. Este tipo de tópico remete para a capacidade que os falantes possuem de identificar um referente que se encontra acessível no discurso, mesmo que este tenha sido referido antes da introdução de novos referentes. Deste modo, Chafe (1987, referido por Frascarelli e Hinterhölzl 2002:2) define o familiar topic como sendo "a given or accessible constituent [...]".

(2) [...] o resto que eu vos disse sobre a atuação de D. Pedro também sabemos que é verdade [...]
(exemplo 25 do corpus)

O shifting topic (Givón 1983, referido por Frascarelli e Hinterhölzl 2002:3) pode caracterizar todos os tópicos "that are newly introduced or newly changed to" (Givón 1983, referido por Frascarelli e Hinterhölzl 2002:3). Este tipo de tópico marca a constante mudança de tópico e a introdução de novos referentes discursivos (por vezes através da prosódia), podendo ser esperada a sua ocorrência também em contexto de discurso oral onde seja bastante evidente a interação entre os sujeitos falantes.

A mudança de tópico e/ou a introdução de novos referentes discursivos pode manter-se ao longo de um determinado período de tempo e pode ocorrer de forma momentânea, isto é, não se manter durante um determinado período de tempo ao longo do discurso. A manutenção de uma mesma entidade como tópico é designada por smooth-shifting topic (Walker e Prince 1993:2), enquanto a referência de forma momentânea a uma determinada entidade denomina-se rough-shifting topic (Walker e Prince 1993:2). Segundo Walker e Prince (1993:2) o smooth shifting topic "tends to occur when the speaker has started talking about a new entity and is doing so in a way that indicates s/he will continue talking about that entity", enquanto o rough-shifting topic verifica-se "when the speaker has started talking about a new entity but is only doing so momentarily" (Walker e Prince 1993:2).

(3) [...] sobre Inês não sei se vocês conhecem alguma coisa [...]
(exemplo 7 do corpus)

(4) [...] para a organização da sílaba nós vamos ter que voltar a falar nisto [...]
(exemplo 34 do corpus)

O tipo de tópico contrastive (Kuno, 1976, Büring 1999, referido por Frascarelli e Hinterhölzl 2002:2) pode ser tido como mais frequente em alguns contextos (por exemplo no contexto do corpus que irá ser analisado: sala de aula). Segundo Kuno (1976, referido por Frascarelli e Hinterhölzl 2002:2) e Büring (1999 referido por Frascarelli e Hinterhölzl 2002:2) o contrastive topic ocorre quando se verifica a existência de um elemento "that induces alternatives which have no impact on the focus value and creates oppositional pairs with respect to other topics". O contraste operado por este tipo de tópico pode verificar-se, também a partir da selecção de um dado elemento de um conjunto de alternativas. Calhoun, Nissim, Steedman e Brenier 2005:5 justificam estas

duas formas de contraste ao afirmarem que o falante "intends to contrast the word with a previous one which was (a) a current topic; (b) semantically related to the contrastive word, such that they belong to a natural set".

(5) [...] um na rua era quando o Lemos estava cá em baixo à janela [...]
(exemplo 200 do corpus)

(6) [...] em que falava daqueles que da lei da morte se vão libertando [...]
(exemplo 6 do corpus)

As terminologias utilizadas pelos investigadores dos diferentes tipos de tópico podem depender da forma como é considerado o estatuto das entidades discursivas, isto é, se se encontram ou não acessíveis no próprio discurso ou no universo discursivo dos falantes.

## 6. O estatuto informacional dos referentes do discurso

O discurso é construído a partir da referência a diversas entidades que devem ser reconhecidas pelos interlocutores, de modo a produzir a coerência da conversação.

Segundo Lambrecht (1994), os referentes discursivos podem ser representados através de entidades ou de proposições. Neste último caso, o estatuto de referente discursivo só é atribuído à proposição se o falante considerar que o seu conteúdo é reconhecido pelo ouvinte. No que diz respeito às representações sintáticas, os referentes discursivos encontram-se expressos como argumentos podendo, assim, ser representados através de sintagmas nominais, diferentes tipos de orações subordinadas e de sintagmas adverbiais que se referem às circunstâncias da predicação Por sua vez, os referentes derivados do conteúdo proposicional podem ser representados através de orações subordinadas ou de pronomes e, finalmente, os referentes discursivos não se encontram expressos em segmentos classificados sintaticamente como predicados, dado que esta categoria apenas atribui propriedades às entidades que compõem os argumentos da frase.

Relativamente ao estudo da estrutura informacional, este autor menciona duas categorias que permitem a identificação e o reconhecimento de uma representação discursiva por parte do ouvinte. A primeira categoria diz respeito à identificação do referente e relaciona-se com " a speaker's assessment of wether a discourse representation

of a particular referent is already stored in the hearer's mind or not" (Lambrecht 1994:76). Neste sentido, esta categoria cognitiva leva a que o falante, ao produzir uma asserção que envolve uma entidade que ainda não se encontra representada na mente do ouvinte, sinta a necessidade de criar uma descrição linguística que represente essa entidade e que possa ser mencionada anaforicamente no discurso subsequente. Para Lambrecht (1994:77) a criação deste novo referente discursivo pode ser vista como o estabelecimento de um novo ficheiro referencial armazenado no registo discursivo, ao qual podem ser acrescentados novas entidades no decorrer do discurso.

Tendo em conta o modo como se processa a identificação de um referente discursivo, Lambrecht (1994:77) considera que esta categoria se traduz na capacidade que o ouvinte possui de selecionar, de entre outros que se encontram armazenados no seu registo discursivo, um determinado referente (que pode ser designado por uma expressão linguística particular) e identifica-lo com o referente presente na mente do falante. Deste modo, um referente identificável é aquele em que a representação partilhada já existe na mente do falante e do ouvinte no momento da enunciação, enquanto um referente não identificável é aquele em que a representação apenas existe na mente do falante (Lambrecht 1994:77-78).

Do ponto de vista do estudo da estrutura informacional, a caraterística relevante de um referente identificável é o facto de o falante pressupor que uma determinada representação já se encontra presente na mente do ouvinte e que, por isso, pode ser invocada no discurso (Lambrecht 1994:78). A identificação de um referente pode ser marcada através de sintagmas nominais como *A Ana*, *O Presidente da República*, *O Sol*[5], dado que estas expressões apresentam propriedades referenciais específicas que os distinguem dos outros referentes, tendo apenas em comum o facto de serem considerados únicos no universo discursivo. Os referentes expressos através de sintagmas nominais são facilmente identificados pelo ouvinte sendo, por isso, considerados mais acessíveis do ponto de vista pragmático. Para além dos referentes designados por sintagmas nominais, encontram-se também, entre outras expressões que designam referentes únicos, as descrições genéricas, quer definidas quer indefinidas. Desta forma, ao identificar a classe

[5] Exemplos adaptados de Lambrecht (1994:87).

de entidades designada por uma expressão linguística, o ouvinte identifica o referente único dessa expressão (cf. Lambrecht 1994:88).

Os sintagmas que referem conjuntos de entidades em vez de indivíduos podem também ser considerados identificáveis, devido ao facto de se destacarem no universo pragmático do falante e do ouvinte. Como exemplos deste tipo de sintagmas salientam-se *as crianças* e *a senhora da limpeza* [6](Lambrecht 1994:88), que constituem elementos facilmente extraídos das classes, pois no universo discursivo dos interlocutores existe um elemento mais saliente que é, geralmente, designado pelo sintagma nominal definido.

O facto de os referentes discursivos ficarem armazenados na memória dos interlocutores, podendo ser inseridos no contexto do discurso, permite a identificação desses referentes quer representem entidades únicas ou entidades cuja unicidade é dada pelo conhecimento partilhado entre os interlocutores. Contudo, este princípio não se aplica nos casos em que os referentes se apresentam expressos de forma deítica ou anafórica (Lambrecht 1994:88)[7].

A segunda categoria proposta por Lambrecht para o reconhecimento dos elementos referidos no discurso é designada por ativação. O processo de ativação consiste, essencialmente, em saber reconhecer que um determinado referente apresenta diferentes estados mentais (Lambrecht 1994:93).

Como já foi referido, esta categoria relaciona-se com as proposições pressupostas invocadas pelas frases (Lambrecht 1994:77). Neste sentido, o ouvinte deve conhecer o

[6] Exemplos traduzidos de Lambrecht 1994:88).

[7] Relativamente à referência deítica, os elementos são identificáveis por se encontrarem visíveis ou salientes no plano discursivo. Estes referentes podem, ainda, ser ancorados a um dos interlocutores como parte da sua individualidade (Lambrecht 1994:88).
A manifestação dos referentes deíticos no último caso deve-se à ocorrência de uma expressão possessiva que remete para a presença do interlocutor ou de elementos relacionados com o mesmo (Lambrecht 1994:88).No que diz respeito à referência anafórica, Lambrecht (1994:89) menciona que a identificação do referente é garantida, dado que este já se encontra referido no discurso anterior. Deste modo, "the identifiability status of a referent is normally preserved throughout a discourse, and from one discourse to another, unless the speaker assumes that the addressee has forgotten the existence of the referent" (Lambrecht 1994:89).

conjunto de proposições pressupostas relevante e aceder facilmente a estas proposições e aos elementos que as compõem, de modo a conseguir processar as proposições transmitidas pelo enunciado (Lambrecht 1994:93).

Os referentes podem ocorrer com diferentes estados de ativação.

Grosz & Sidner (2009:4) afirmam que "access to the relevant information at each stage of sentence production is associated with *accessibility* statuses of the corresponding units". Os estados de ativação dos referentes são determinados por fatores psicológicos como a consciência e a diferença entre a memória a longo prazo e a memória a curto prazo dos falantes.

Chafe (1987) apresenta três estados de ativação: ativo, semi-ativo (também denominado acessível) e inativo. Um referente encontra-se no estado ativo quando se apresenta como foco na consciência dos interlocutores, podendo ser ativado em qualquer momento. O estado semi-ativo ou acessível verifica-se quando o referente se encontra na consciência periférica dos interlocutores, não ocupando a posição de foco. Os referentes podem ser tidos como semi-ativos quando se produz uma deativação de um estado anterior, que consistiu tipicamente em assumir o estatuto ativo numa fase anterior do discurso ou porque fazem parte do conjunto de expetativas associadas a um esquema, ocorre quando algumas (ou todas) das expetativas que compõem o discurso são inseridas no estado semi-ativo tornando-se, assim, mais disponíveis do que se permanecessem no estado inativo [8] (Chafe 1987:22ff, referido por Lambrecht 1994:99). No estado inativo, o referente encontra-se presente na memória a longo prazo dos interlocutores, não estando em foco nem ativo na consciência periférica (Chafe 1987:22ff, referido por Lambrecht 1994:94).

Os estados ativo e inativo apresentam uma caraterística fundamental que os distingue: enquanto no estado ativo a ativação de um dado referente termina quando um novo referente é ativado, no estado inativo o estatuto cognitivo do referente mantém-se por um período de tempo mais longo, independentemente dos elementos que se encontram sob a atenção dos interlocutores (Lambrecht 1994:94). A acessibilidade (isto é, o estado semi-ativo) de um referente pode ser atribuida a fatores como: a desativação de um estado

[8] Segundo Chafe (1987), um esquema.

anterior (como por exemplo do estado inativo), as inferências realizadas a partir de um esquema e a presença do mesmo referente no contexto externo ao discurso. Lambrecht (1994:100) estabelece uma correspondência entre os fatores que determinam a acessibilidade do referente e as seguintes categorias: acessível a partir do texto, acessível a partir da realização de inferências e acessível a partir da situação. Assim, um referente é acessível a partir do texto quando um estado anterior é desativado. Quando o referente é ativado através da realização de inferências de um outro elemento acessível presente no discurso, a sua acessibilidade é dada pelo recurso a essas inferências; um elemento cuja referência se obtém através da situação enunciativa é classificado como sendo acessível a partir da situação (Lambrecht 1994:100).

No entanto, o facto de os referentes pertencerem ao tipo inferencial ou ao tipo situacional não significa que estejam presentes na consciência do ouvinte (Lambrecht 1994:104). Relativamente à marcação dos estados de ativação verifica-se que um referente ativo não apresenta acento prosódico e pode ocorrer como uma forma pronominal que corresponde à expressão linguística a que se refere (Lambrecht 1994:96). Os referentes inativos apresentam uma expressão referencial acentuada, cujo conteúdo pode ser representado através de itens lexicais, mas não por formas pronominais, com exceção dos pronomes deíticos (Lambrecht 1994:96). Assim, apesar de os referentes ativos não poderem ser marcados como tal através do acento prosódico ou de uma forma pronominal sem causar ambiguidade, os referentes inativos não apresentam formas prosódicas ou morfológicas não ambíguas[9].

Neste contexto, é relevante incluir na reflexão sobre a integração dos referentes no universo discursivo as propostas de Ariel (1988, 1990) e Gundel, Hedberg & Zacharski (1993), que especificam que a representação de um dado referente na memória se produz, respetivamente, em função do seu grau de acessibilidade ou do seu estatuto cognitivo, incumbindo às expressões referenciais sinalizar esse grau de acessibilidade. As tabelas 1 e 2 são retiradas de Silva (2006: 89-90).

[9] Com base nestas considerações, Lambrecht (1994:104-105) propõe que a acessibilidade seja vista como um processo de ativação e não como o estado do referente na mente dos interlocutores. No processo de ativação o falante espera que o ouvinte tenha capacidade para estabelecer as inferências necessárias que lhe permitam interpretar o referente, a partir de estruturas pressuposicionais particulares (Lambrecht 1994:105).

| Escala de Acessibilidade (“Accessibility Marking Scale”) |
|---|
| Marcação do grau de acessibilidade: + acessível > – acessível |
| elipse > reflexivos > marcadores de concordância > pronomes clíticos > pronomes átonos > pronomes tónicos > pronomes tónicos + gesto > demonstrativo de proximidade (+ NP) > demonstrativo de distanciamento (+ NP) > demonstrativo de proximidade (+ NP) + modificador > demonstrativo de distanciamento (+ NP) + modificador > nome próprio > apelido > descrição definida breve > descrição definida longa > nome completo > nome completo + modificador |

Tabela 1: escala de acessibilidade de Ariel

Hierarquia do Dado (“Giveness Hierarchy”)

| Estatuto cognitivo | Entidade linguística | Função | Exemplo |
|---|---|---|---|
| em foco | eu, tu, ele... | entidade activada, localizada no foco da atenção do locutor | Eu não consegui dormir a noite passada |
| Activado | isto, isso<br>este, esse N | a entidade está activada e foi já objecto de menção no contexto linguístico ou encontra-se no contexto extralinguístico imediato | Esse cão ao pé do portão ladrou toda a noite |
| Familiar | aquele N | o alocutário possui já uma representação mental do referente | Aquele cão ladrou toda a noite |
| unicamente identificável | o N | o alocutário recupera ou constrói uma representação única do referente designado pelo locutor | O cão ladrou toda a noite |
| referencial | este N (indefinido) | o locutor pretende referir uma entidade particular, esperando que o alocutário construa a sua representação mental | Este cão ladrou toda a noite |
| identificável quanto ao tipo de entidade | um N | o interlocutor pode identificar o tipo de entidade descrita | Um cão ladrou toda a noite |

Tabela 2: Hierarquia do dado de Gundel, Hedberg & Zacharski

Em ambas as escalas de acessibilidade do referente,

> o grau de informação disponibilizada por estas expressões é proporcionalmente inverso ao grau de acessibilidade do referente na memória discursiva, isto é, quanto mais acessível for o referente textual na memória operacional do falante, maior é o estado da sua activação e, por conseguinte, menor a especificação das informações contidas na expressão indexical (Silva 2006: 91).

Neste âmbito, o tópico pode assumir diferentes papéis no decorrer do discurso, consoante as intenções e os interesses dos sujeitos falantes.

No caso do discurso oral espontâneo é esperada a existência de uma maior oscilação no que diz respeito às funções que os vários tópicos podem desempenhar, devido às características deste tipo de discurso (espontaneidade, produção dos enunciados no momento exato da enunciação, influência direta do contexto envolvente, etc.).

Desta forma, no capítulo 2 pretender-se-á apresentar uma lista de critérios semânticos que possam ilustrar as várias funções dos tópicos e analisar alguns tópicos apresentados em vários segmentos de discurso oral espontâneo e preparado do ponto de vista das funções semânticas que cada um deles pode representar no contexto do discurso oral.

# CAPITULO 2

# PROPOSTA DE ANOTAÇÃO SEMÂNTICA DE UM DISCURSO ORAL

Na sequência da reflexão teórica realizada na primeira parte deste trabalho, estabelecemos um conjunto de princípios de análise descritiva conducentes à determinação de um quadro operatório para a anotação semântica de um *corpus* de discurso oral, especificamente no que se refere à determinação dos tópicos discursivos ocorrentes. Nesse âmbito, definimos sumariamente o conceito de estrutura informacional, associando-o à possibilidade de analisar a organização informacional do discurso, a nível frásico ou a um nível superior ao da frase, procurando interligar os fatores linguísticos, pragmáticos e cognitivos que lhe estão subjacentes. Daí decorreu a apresentação de alguns tipos de tópicos comummente usados para a descrição da forma como se estrutura a informação discursiva, que servem de base à proposta de uma tipologia específica para a anotação realizada.

Neste capítulo, que tem como tema central uma proposta de etiquetagem da anotação do corpus do ponto de vista da sua análise semântico-discursiva, procedemos: i) à caracterização sumária do *corpus* usado; ii) à metodologia seguida; iii) à explicitação do conjunto de etiquetas; iv) à exemplificação da aplicação das etiquetas definidas ao *corpus*. Na sequência deste percurso, apresentamos e discutimos os resultados obtidos, no sentido de avaliar a metodologia proposta e de refletir sobre a articulação entre a descrição teórica e o trabalho de anotação.

## 1. Apresentação do *corpus*

O *corpus* analisado constitui uma amostra de 116 exemplos transcritos do corpus CPE-FACES – Corpus de Português Europeu Falado por Adolescentes em Contexto Escolar.

Trata-se de um corpus de discurso espontâneo e planeado gravado em escolas secundárias portuguesas, com o objetivo inicial de "reunir um conjunto de textos falados com características situacionais diferenciadas, cuja ocorrência se assume como regular no contexto escolar português, em que os alunos participam no papel de falantes" (Mata, 1995: 349). Contém um total de 15h de gravações, incluindo o discurso de 2 professores e 25 alunos, e está ortograficamente transcrito.

No quadro do projeto em que se enquadra este trabalho, COPAS, o *corpus* referido é um dos cinco *corpora* considerados para o trabalho de investigação proposto, tendo sido o primeiro a ser objeto de um trabalho de anotação sintática, prosódica e semântica. Trata-se, na sua essência, de um trabalho exploratório, com duas finalidades básicas: i) estabilizar uma proposta de etiquetagem; ii) aferir da possibilidade de correlação entre as camadas de anotação consideradas.

Esse facto motivou a seleção de 116 exemplos do *corpus* CPE-Faces, aleatoriamente extraídos, contendo ocorrências de discurso espontâneo e planeado, bem como discurso de professores e alunos, que serviu como amostragem para a proposta de anotação e o estabelecimento das camadas de análise a considerar.

## 2. Metodologia

Tendo em consideração o conjunto de tópicos de análise previstos no projeto, foi tomada a decisão de concentrar o trabalho conjunto de aferição de critérios e de parâmetros nos tópicos, localizados na periferia esquerda da frase.

## 2.1. Metodologia geral

Tomado como objeto de análise das três disciplinas envolvidas neste fenómeno, seguiu-se a seguinte metodologia de trabalho:

i. análise preliminar do *corpus* para observação empírica dos dados;
ii. revisão da literatura sobre a questão;
iii. delimitação dos parâmetros de análise comuns para a anotação;
*iv.* proposta de uma taxonomia passível de distinguir as diferentes ocorrências do *corpus*;
v. anotação da amostra de acordo com a proposta taxonómica realizada;
vi. discussão e avaliação dos resultados;
vii. reformulação da proposta inicial;
viii. nova anotação da amostra;
ix. discussão dos resultados obtidos na anotação da amostra.

## 2.2. Método de análise dos exemplos do corpus

O trabalho de anotação foi feito de forma sequencial através do recurso ao programa Praat, seguindo-se a seguinte orientação: i) anotação prosódica; ii) anotação sintática; iii) anotação semântica. Esta sequencialidade decorreu de uma opção prática de organização do trabalho, depois de estabelecidas em conjunto as camadas a considerar e as diferentes etiquetas a incluir.

Os exemplos que constituem o corpus foram analisados, do ponto de vista de todas as áreas envolvidas, com recurso ao programa informático Praat. Neste sentido, selecionaram-se ambos os ficheiros em formato de áudio e em formato textgrid, modificando as características do som em *clone time domain* (figura 1).

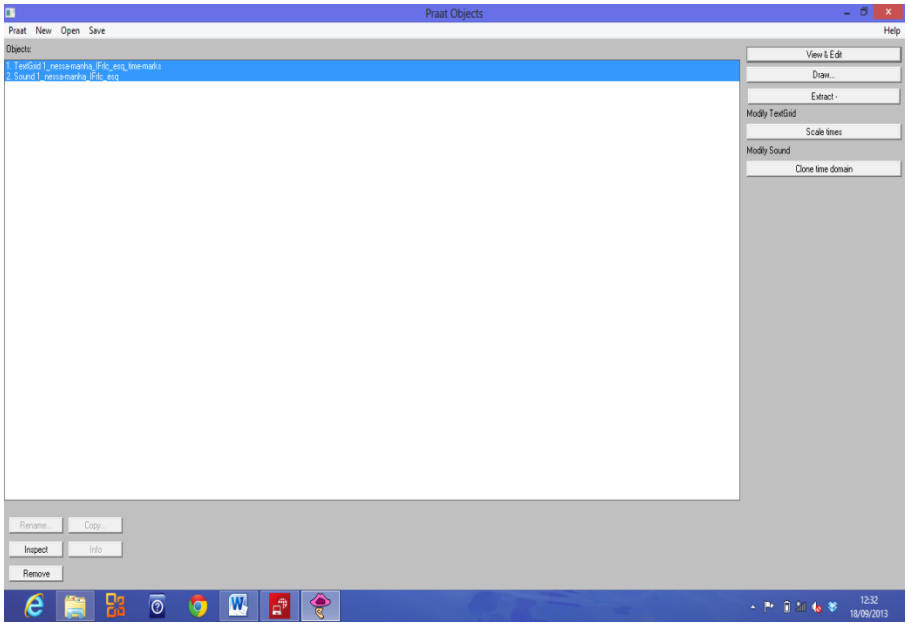



Figura 1[10]: Iniciação do programa Praat

Para visualizar a transcrição e o ficheiro de áudio que correspondem ao exemplo em análise é necessário utilizar a opção *view and edit* (figura 2).

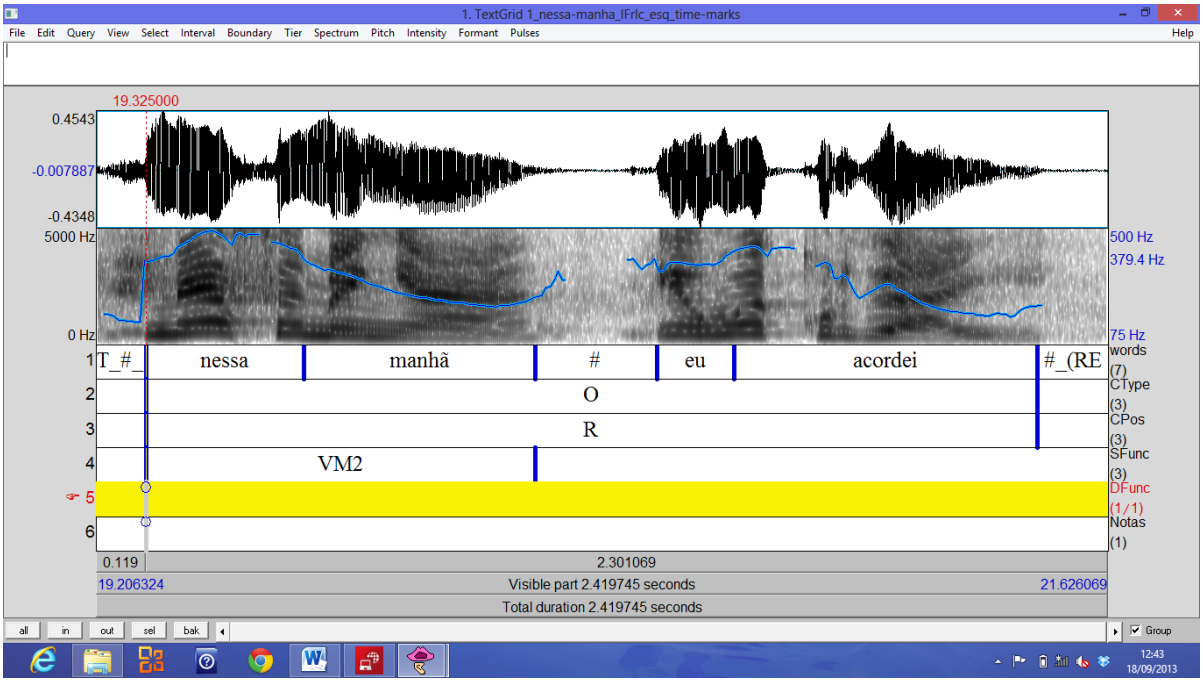



Figura 2: Visualização da transcrição e do ficheiro de áudio.

[10] Tendo em consideração a especificidade de cada um dos capítulos no contexto deste trabalho, optámos por numerar os exemplos, quadros e esquemas e gráficos de forma independente.

As fiadas apresentadas pelo Praat encontram-se numeradas de um a seis, sendo que a primeira corresponde à transcrição do exemplo, a segunda corresponde ao tipo de construção, a terceira ao tipo de construção prosódica, a quarta corresponde à função sintática, a quinta à função semântica e a sexta pode ser utlizada para o registo de algumas notas que o anotador considere relevantes.

Após a visualização da transcrição e do ficheiro de som correspondentes ao exemplo que se encontra a ser analisado delimitam-se as fronteiras do constituinte que ser quer classificar. No caso concreto deste trabalho, o objetivo é marcar a função semântica (fiada 5) do constituinte que desempenha a função de tópico na periferia esquerda da frase.

A delimitação das fronteiras semânticas do constituinte que desempenha a função de tópico seguiu a mesma delimitação realizada pela sintaxe na fiada 4 (figura 3).

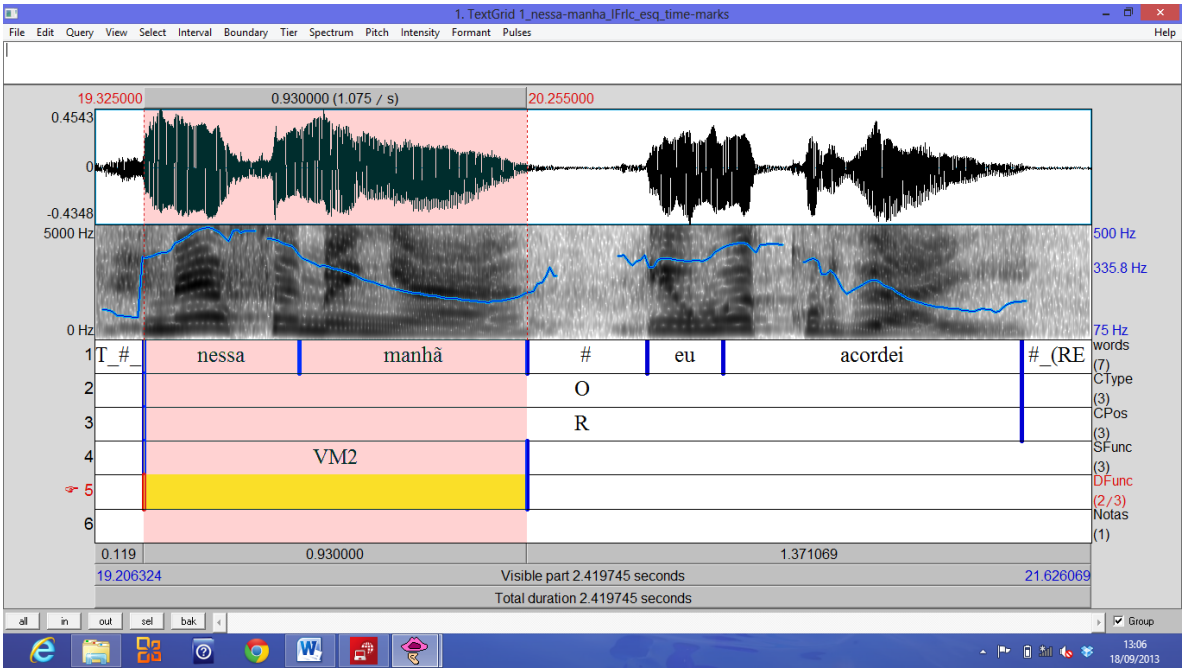



Figura 3: Delimitação das fronteiras do constituinte marcado como tópico

Para se classificar o constituinte delimitado é necessário recorrer à audição do ficheiro de som bem como ao contexto, isto é, de modo a atribuir a classificação correta deve-se ouvir o exemplo de forma a perceber se existem marcas de marcação prosódica e recorrer ao contexto antecedente para identificar as marcas que permitem a escolha de um determinado tipo de tópico (figura 4).

A visualização do contexto pode ser feita através de ficheiros em formado *word*, onde se encontram transcritas todas as conversações.

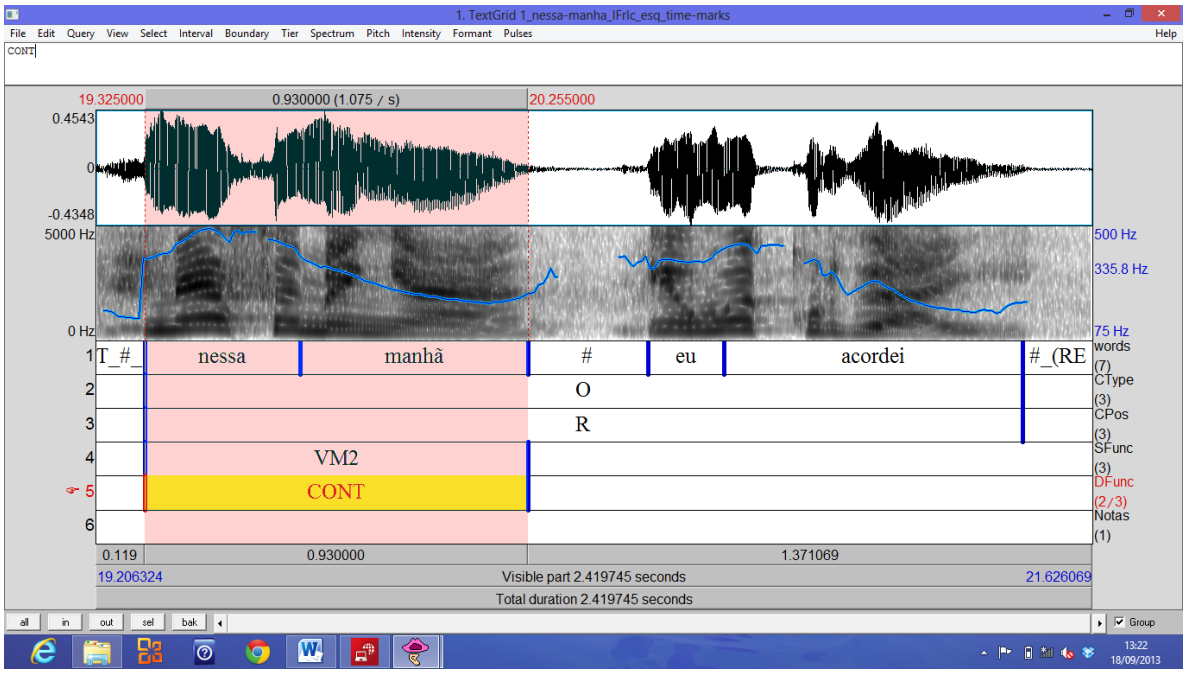



Figura 4: Classificação do constituinte marcado como tópico

## 3. Taxonomia de tópicos para a anotação discursiva

A classificação gerada para a anotação dos exemplos baseia-se nos trabalhos de Givón (1983, referido por Frascarelli e Hinterhölzl_2002:2), Chafe (1987, referido por Frascarelli e Hinterhölzl 2002:2), Kuno (1976 referido por Frascarelli e Hinterhölzl 2002:2), Büring (1999, referido por Frascarelli e Hinterhölzl 2002:2), de Calhoun, Nissim, Steedman e Brenier (2005) e de Walker e Prince (1993).

A consideração de diferentes conceitos propostas por vários autores deve-se ao facto de se tentar adequar a metodologia de classificação não só à Língua Portuguesa, como também ao tipo de exemplos analisados.

Após o esclarecimento relativamente ao papel que a informação desempenha ao longo do discurso e ao estatuto que as entidades discursivas que fazem parte dela podem assumir no decorrer do texto, bem como da descrição de diferentes tipos de tópicos na primeira parte deste trabalho, selecionaram-se os tipos de tópicos considerados mais

produtivos que permitem delimitar e classificar a função dos diferentes tópicos identificados no *corpus*.

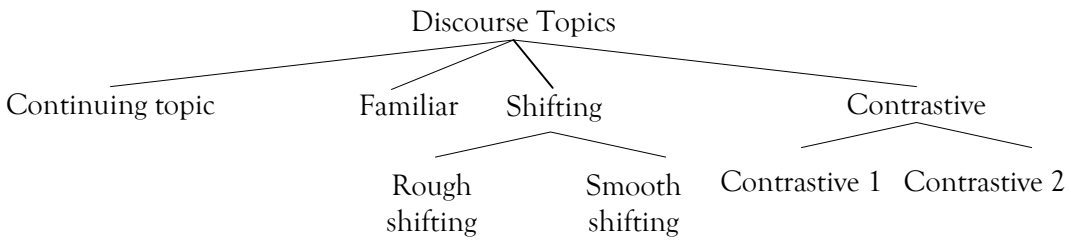
Esta proposta, no entanto, não esteve presente na análise preliminar dos exemplos. Os vários tipos de tópico considerados ao longo deste trabalho bem como a forma de os aplicar foram sendo redefinidos com o auxílio de novas leituras bibliográficas e da própria análise dos exemplos. Note-se que no decorrer da elaboração do conjunto de etiquetas semânticas tentou-se criar uma terminologia que não fosse nem demasiado reduzida, nem demasiado extensa.

Uma das propostas consideradas apresentou bastantes semelhanças com a que é apresentada neste trabalho, dado que considerava apenas os quatro tipos de tópicos tidos aqui como principais (*continuing topic*, *familiar topic*, *shifting topic* e *contrastive topic*). No entanto, esta proposta revelou-se insuficiente na medida em alguns dos exemplos analisados combinados com o contexto em que ocorrem suscitaram dúvidas na sua anotação, ou seja, no tipo de tópico a atribuir.

Neste sentido, a proposta apresentada neste trabalho tem em consideração as diferenças que podem existir entre tópicos que pertencem ao mesmo tipo (como é o caso do *shifting topic* e do *contrastive topic*), o que facilita a distinção e consequente anotação dos exemplos do corpus.

De seguida, apresentam-se os tópicos selecionados e que servem de etiqueta aos segmentos do texto anotado identificados como sendo tópico discursivo, ilustrando-se a a sua aplicação com recurso a vários exemplos do *corpus*.

O esquema 1 sintetiza a taxonomia proposta.




### *3.1. Continuing topic* ⇨ CONT

O termo *continuing topic* (Givón 1983, referido por Frascarelli e Hinterhölzl_2002:2) diz respeito à continuidade de um referente no discurso, isto é, à forma que o locutor utiliza para referir uma mesma entidade discursiva ao longo da conversação. Neste sentido, Givón (1983, referido por Frascarelli e Hinterhölzl_2002:2) considera que este tipo de tópico ocorre quando um determinado tópico é referido no discurso prévio (1) - (4). O *continuing topic* verifica-se também nas relações anafóricas existentes no discurso.

**(1) [...] <u>nessa manhã</u> eu acordei [...]**

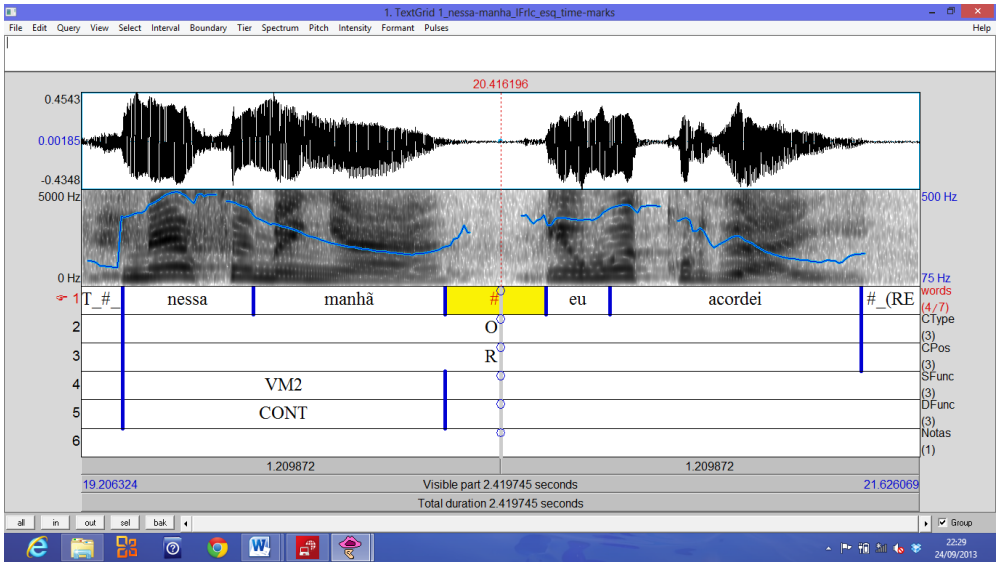



Figura 5: análise do exemplo (1) na plataforma informática Praat

**(2) [...] <u>nesse dia</u> surgiu o primeiro sinal de que eu ia ter o parto [...]**

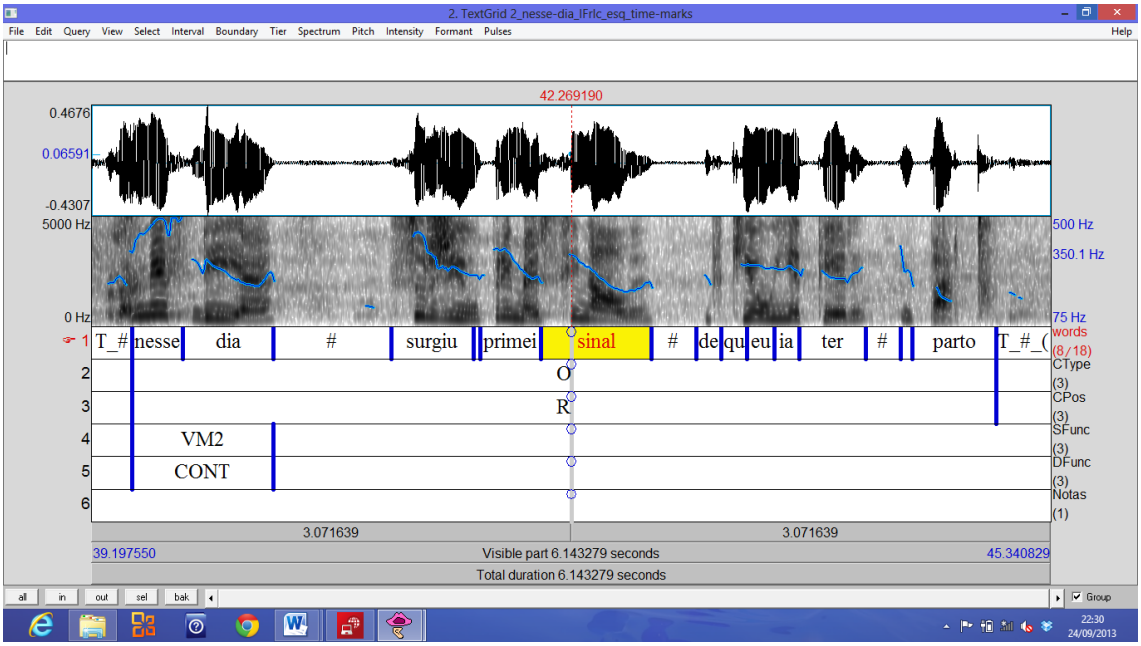



Figura 6: análise do exemplo (2) na plataforma informática Praat

(3) [...] <u>dessa relação</u> **nasceram três filhos** [...]

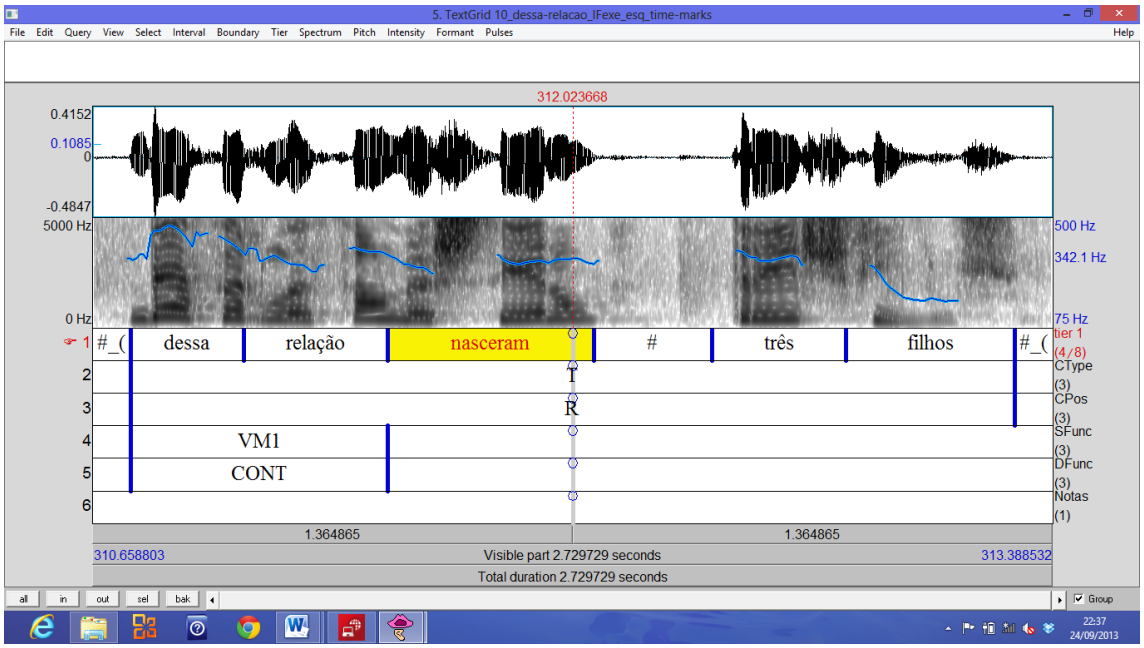



Figura 7: análise do exemplo (3) na plataforma informática Praat

(4) [...] <u>essas viagens que se chamavam infinitas entre aspas</u> **os homens pela sua ambição de ter mais e mais iam para India** [...]

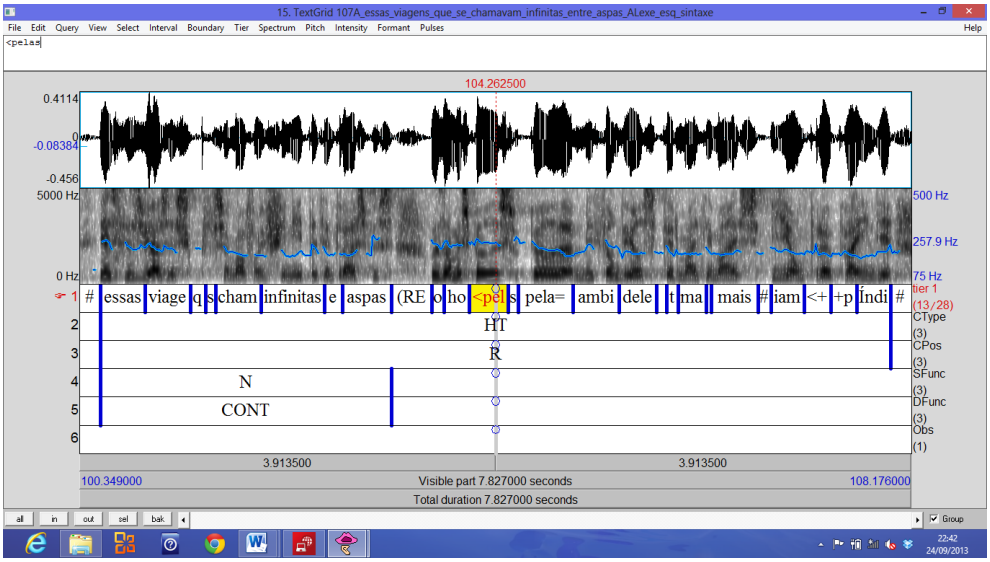



Figura 8: análise do exemplo (4) na plataforma informática Praat

### *3.2. Familiar topic ⇨ F*

O termo *familiar topic* (Chafe 1987, referido por Frascarelli e Hinterhölzl 2002:2) revela algumas semelhanças com o *continuing topic* (Givón 1983, referido por Frascarelli e Hinterhölzl 2002:2), no sentido em que o interlocutor reconhece um determinado referente. Deste modo, um tópico pode ser classificado como familiar quando um dado referente se encontra acessível no discurso, isto é, quando o interlocutor consegue identificar facilmente um referente discursivo (5) – (9).

(5) **[...] o resto que eu vos disse sobre a atuação de Dom Pedro também sabemos que é verdade[...]**

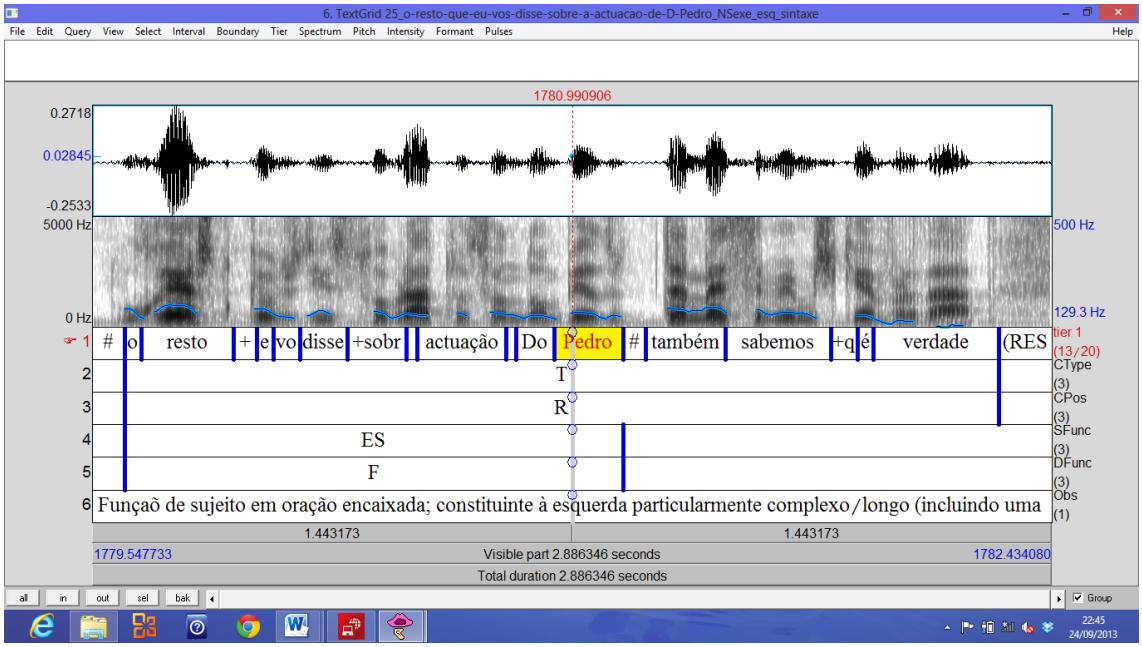



Figura 9: análise do exemplo (5) na plataforma informática Praat

**(6) [...] mas antes de começarmos a trabalhar os Lusíadas e ver de que modo é que <u>Camões também ele</u> foi um renascentista e um humanista [...]**

Figura 10: análise do exemplo (6) na plataforma informática Praat

**(7) [...] <u>com uma antena</u> a dar-lhe cacetadas a ver se ele saia [...]**

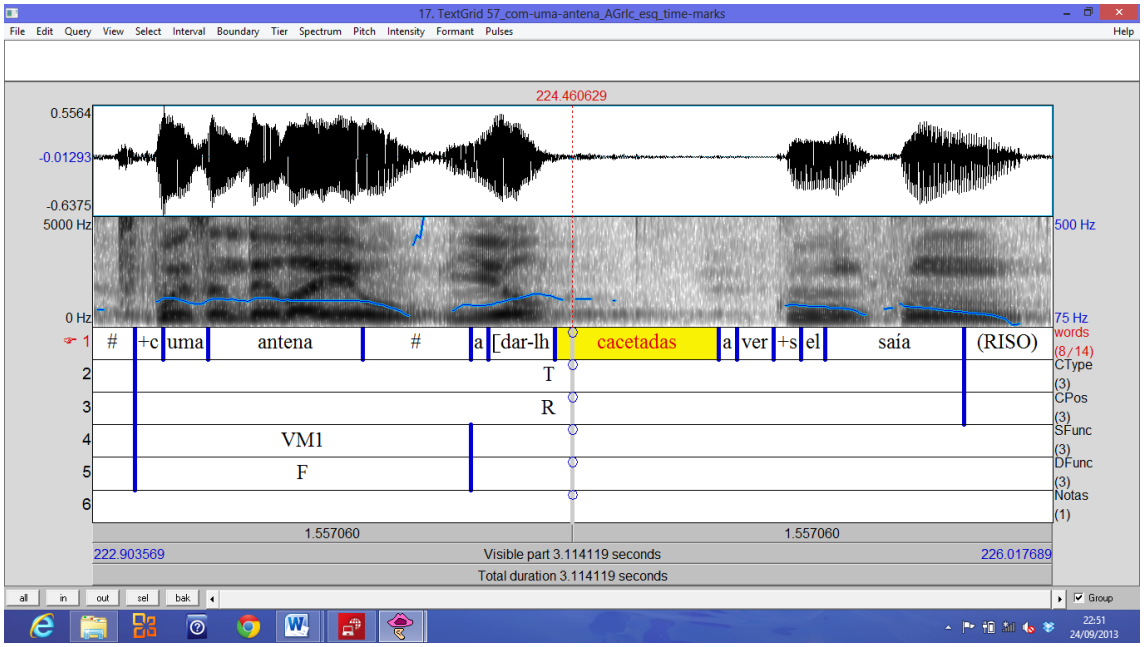



Figura 11: análise do exemplo (7) na plataforma informática Praat

(8) [...] nós antes disso se não acontecesse nada de mal a gente tinha pensado dar uma volta pelo Chiado mas como aconteceu isso [...]

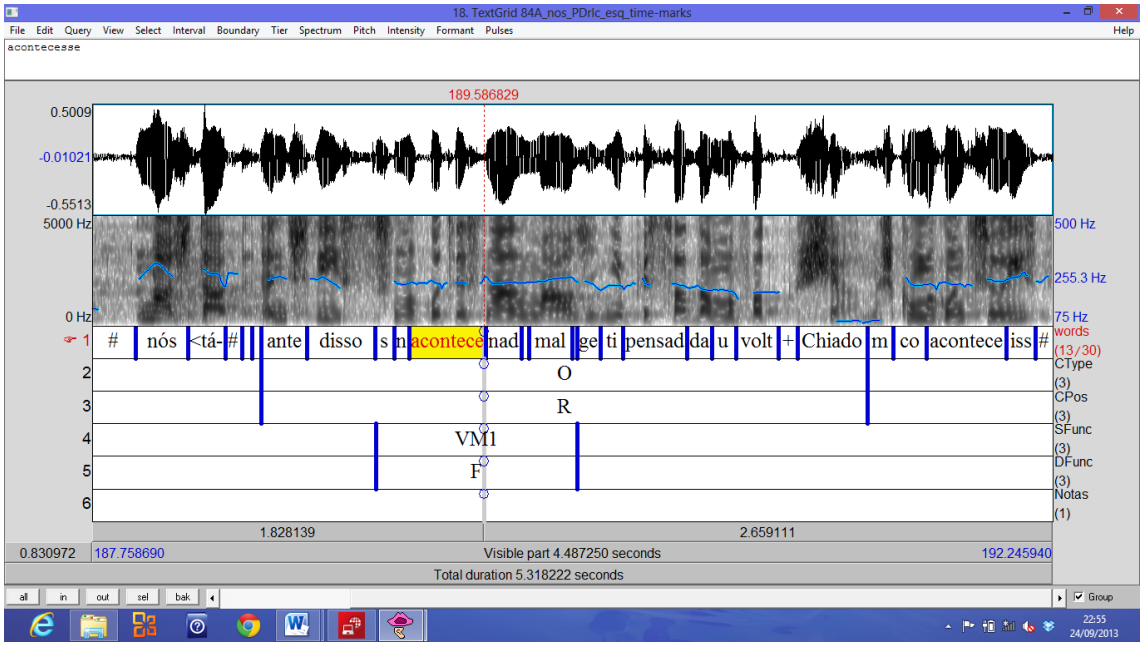



Figura 12: análise do exemplo (8) na plataforma informática Praat

### *3.3. Shifting topic ⇨ SHIFT*

O *shifting topic* (Givon 1983, referido por Frascarelli e Hinterhölzl 2002:3) ocorre quando se verifica uma mudança de tópico ou a introdução de um novo tópico, ou seja, verifica-se quando o locutor abandona (momentânea ou definitivamente) um determinado referente e introduz outro que não se encontra relacionado com o primeiro

Segundo Walker e Prince (1993:2) este tipo de tópico pode ser dividido em dois subtipos: *rough-shifting topic* e *smooth-shifting topic*.

#### *3.3.1. Rough-shifting topic ⇨ SHIFT(RSHIFT)*

Ocorre quando o falante refere de forma momentânea uma nova entidade no discurso (9) - (12) (Walker e Prince 1993:2).

**(9) [...] <u>Dom Sancho quarto</u> entre filhos que teve teve alguns ilegítimos [...]**

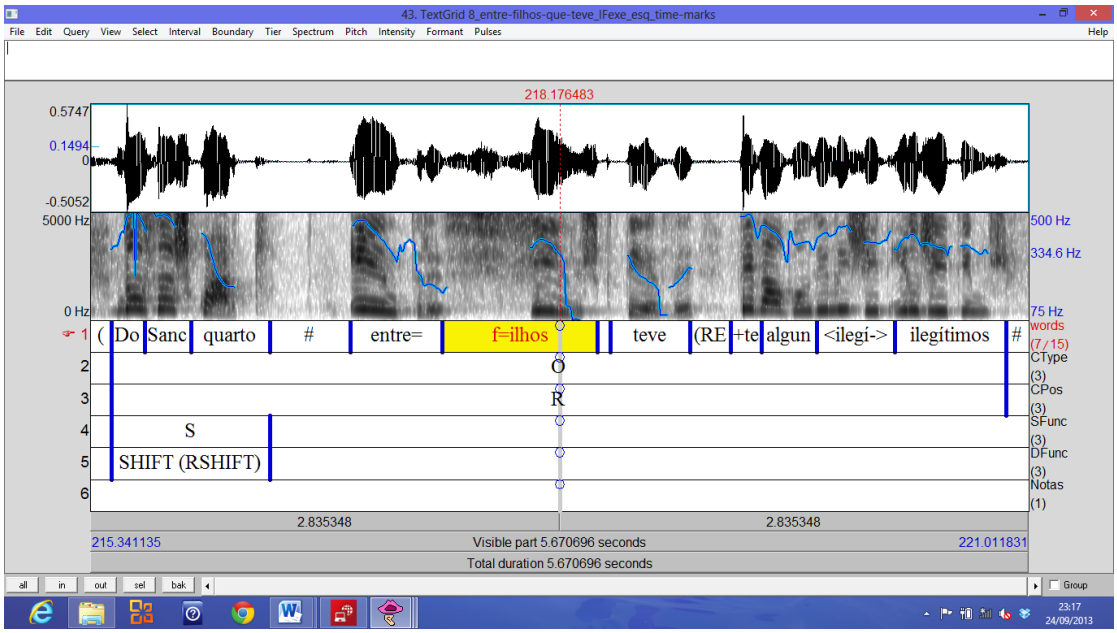



Figura 13: análise do exemplo (9) na plataforma informática Praat

**(10) [...] <u>politica</u> sabemos ainda pouca coisa [...]**

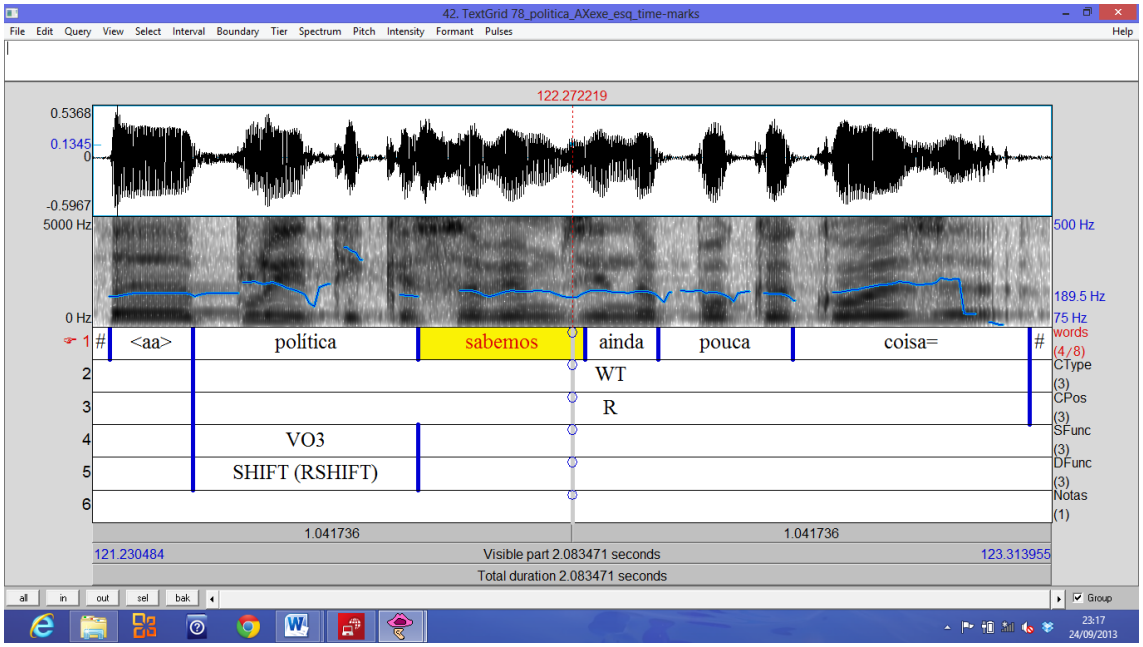



Figura 14: análise do exemplo (10) na plataforma informática Praat

(11) [...] depois do <u>marido</u> não temos nada [...]

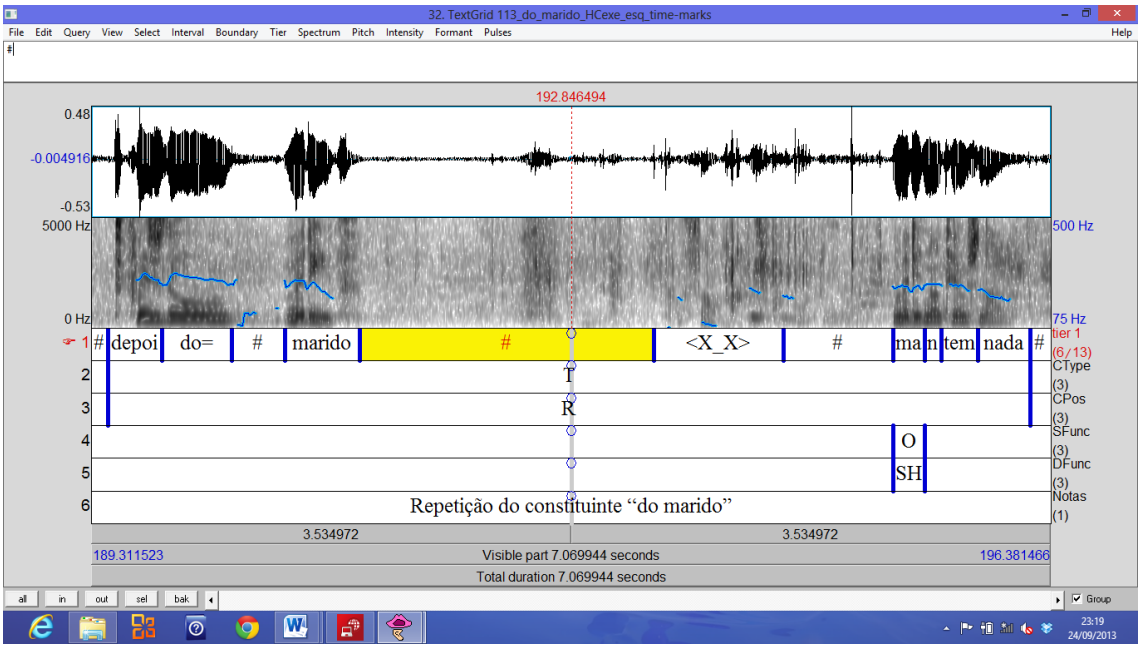



Figura 15: análise do exemplo (11) na plataforma informática Praat

(12) [...] <u>os professores</u> também não tenho nada contra raras exceções [...]

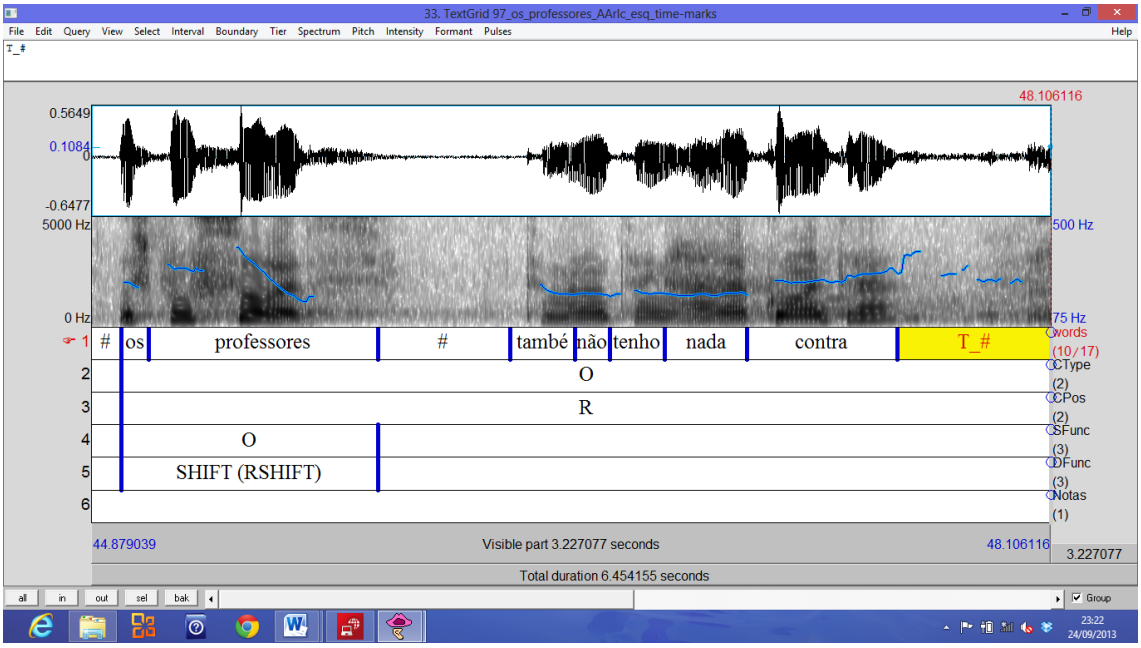



Figura 16: análise do exemplo (12) na plataforma informática Praat

### *3.3.2. Smooth-shifting topic⇨ SHIFT(SSHIFT)*

Verifica-se quando o falante introduz uma nova entidade no discurso de forma a indicar que possui intenções de continuar a referir essa mesma entidade (13) - (16).

(13) [...] <u>em Castela</u> a certa altura houve um problema [...]

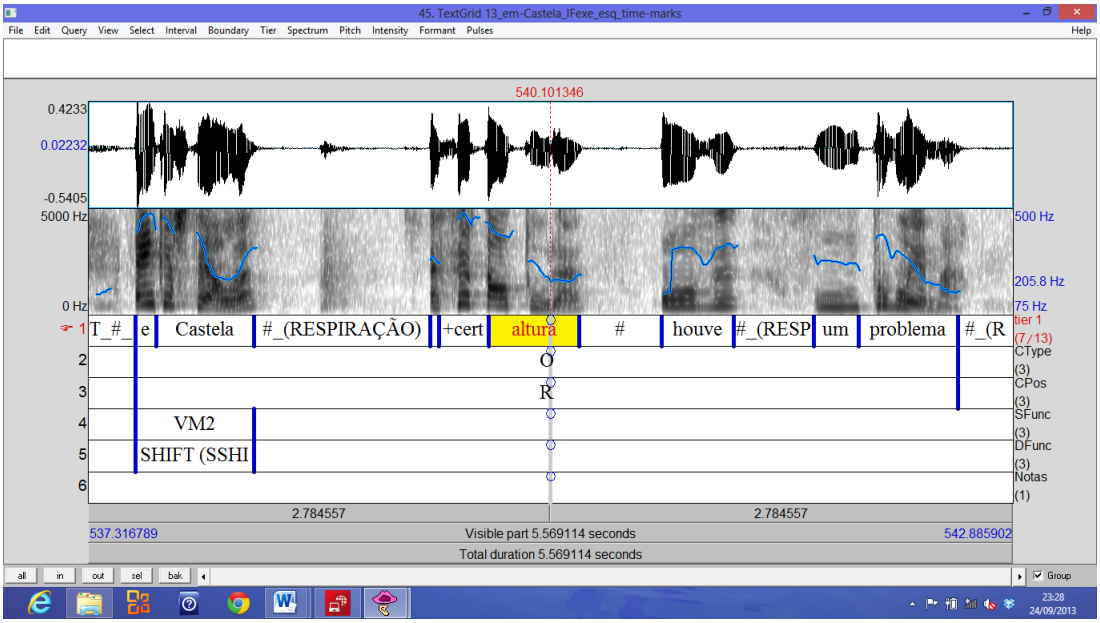



Figura 17: análise do exemplo (13) na plataforma informática Praat

**(14) [...] <u>desporto</u> é lá bastante praticado o baseball [...]**

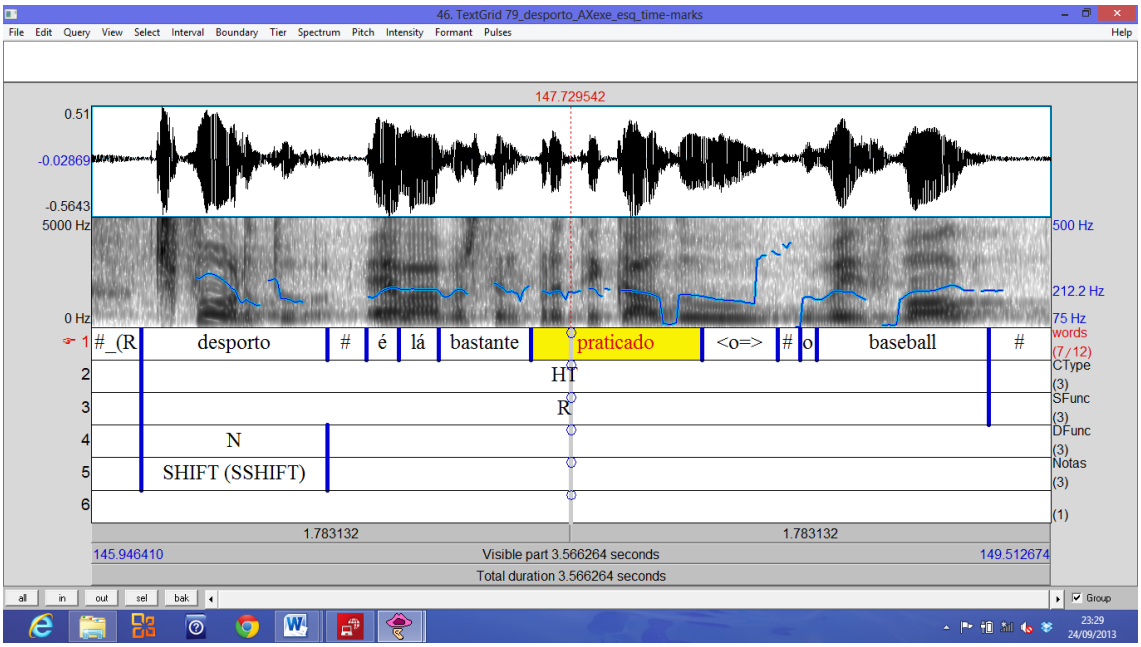



Figura 18: análise do exemplo (14) na plataforma informática Praat

**(15) [...] <u>nós</u> sobre o marido não temos muitas informações sobre a moça temos só duas [...]**

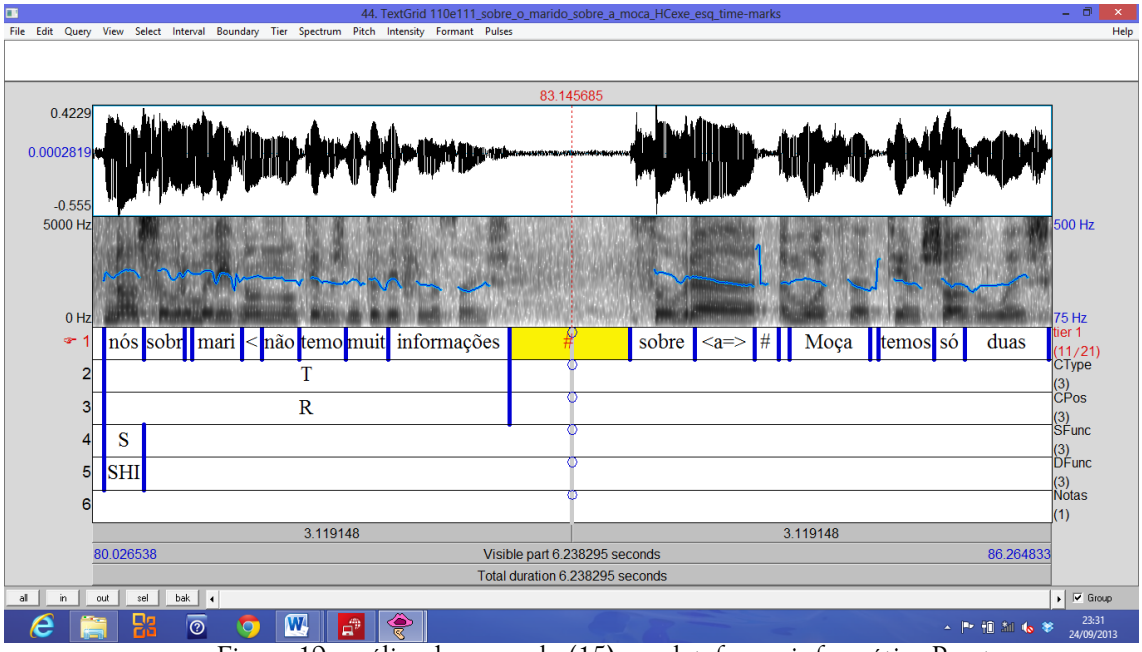



Figura 19: análise do exemplo (15) na plataforma informática Praat

(16) [...] <u>do Lemos</u> é uma pessoa que é mestre em artes poéticas contos instrumentos de corda [...]

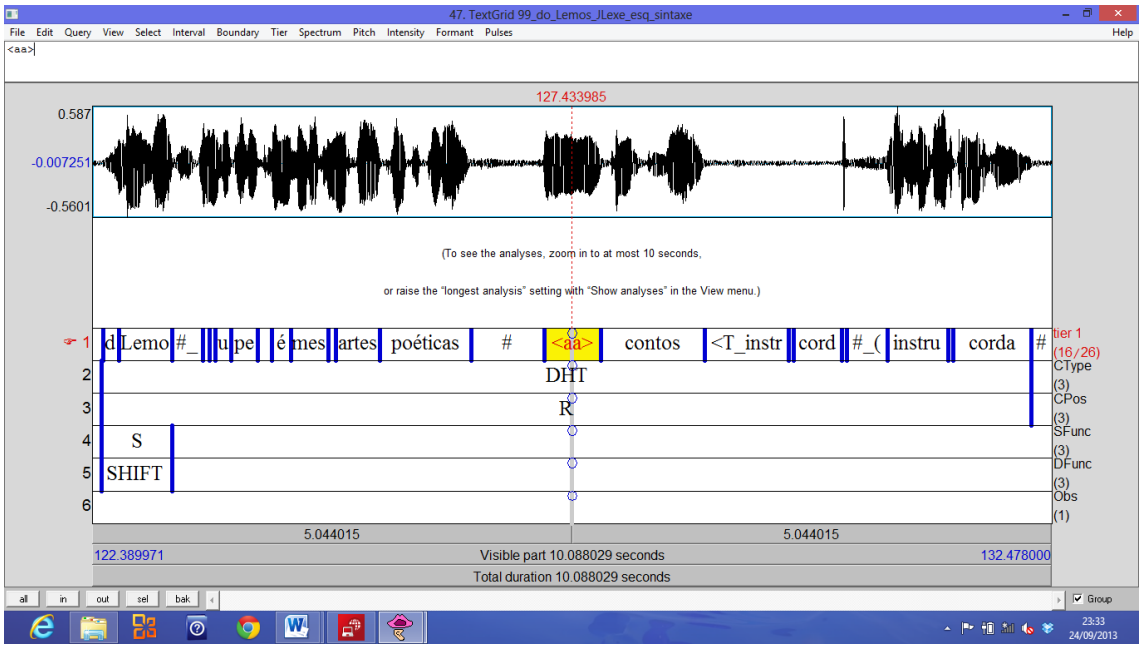



Figura 20: análise do exemplo (16) na plataforma informática Praat

### *3.4. Contrastive topic ⇨ C*

O último termo considerado denomina-se *contrastive topic* (Kuno, 1976, Büring 1999, referido por Frascarelli e Hinterhölzl 2002:2), dado que implica a existência de um contraste entre dois ou mais elementos presentes no discurso. Este tipo de tópico ocorre quando se verifica a presença de um valor contrastivo ou a existência de um elemento que induz um conjunto de alternativas (Kuno, 1976, Büring 1999, referido por Frascarelli e Hinterhölzl 2002:2).

Tal como o *shifting topic* (Givón 1983, referido por Frascarelli e Hinterhölzl 2002:3), o *contrastive topic* (Kuno, 1976, Büring 1999, referido por Frascarelli e Hinterhölzl 2002:2) também pode ser dividido em dois subtipos, tendo em conta os elementos e a forma como o contraste é realizado.

#### *3.4.1. Contrastive 1⇨ C(C1)*

O contraste é realizado através da oposição entre dois elementos, por exemplo x ou y (17) – (19) (Kuno 1976 e Büring 1999 referidos por Frascarelli e Hinterhölzl 2002:2).

**(17) [...] <u>para mim</u> é um herói mas para você não deve ser [...]**

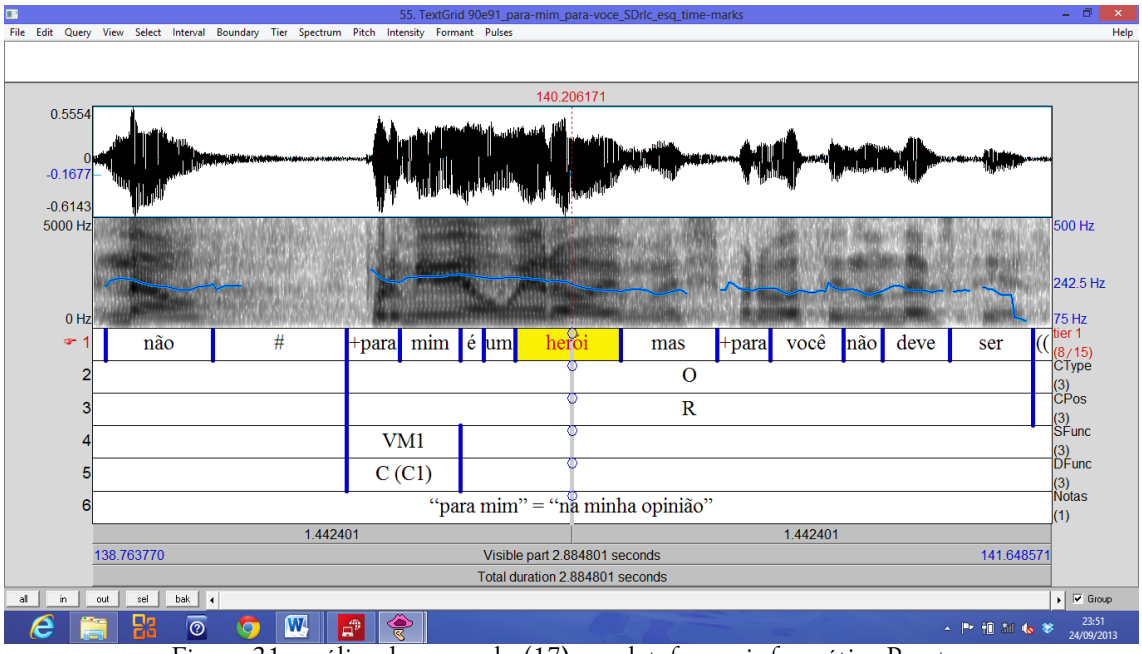


Figura 21: análise do exemplo (17) na plataforma informática Praat

(18) [...] <u>o que fizeram a mim</u> também já fizemos a ele [...]

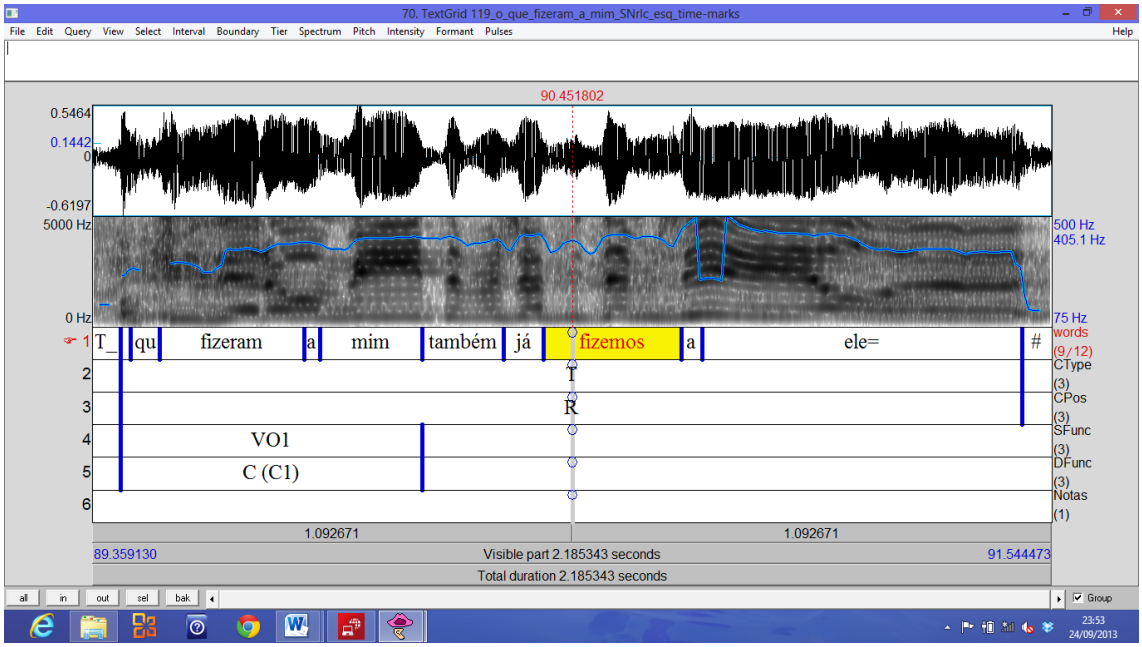



Figura 22: análise do exemplo (18) na plataforma informática Praat

(19) [...] <u>tarefa cinco</u> ninguém mais tem [...]

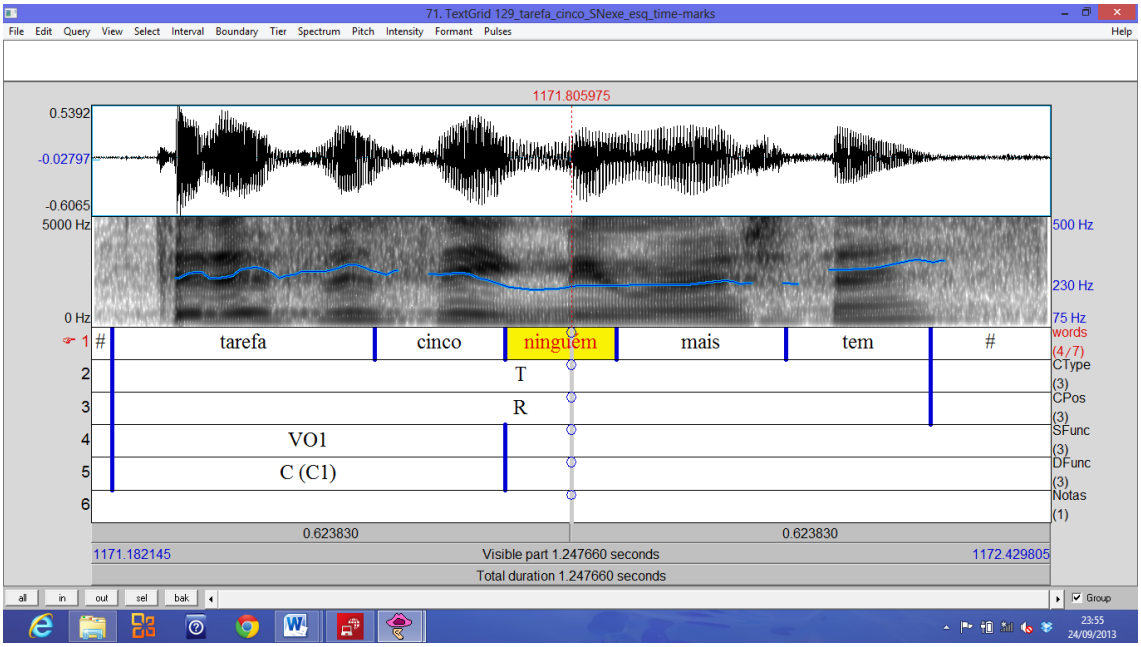



Figura 23: análise do exemplo (19) na plataforma informática Praat

### *3.4.2. Contrastive 2⇨ C(C2)*

O contraste é realizado a partir da escolha de um determinado elemento de um conjunto de alternativas presente no contexto ou no conhecimento do interlocutor (55), (20) – (22) (Calhoun, Nissim, Steedman e Brenier 2005:5).

(20) [...] e <u>entre as damas que a acompanhavam todas elas de boas famílias</u> vinha Dona Inês de Castro [...]

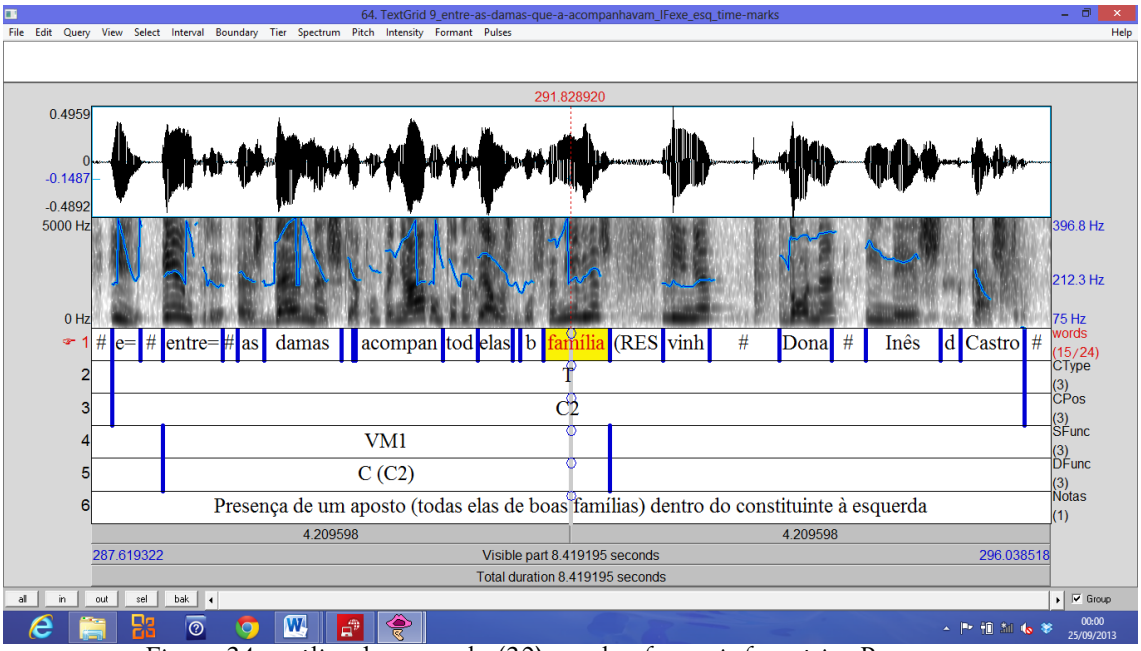


Figura 24: análise do exemplo (20) na plataforma informática Praat

(21) [...] Dom Sancho quarto <u>entre filhos que teve</u> teve alguns ilegítimos [...]

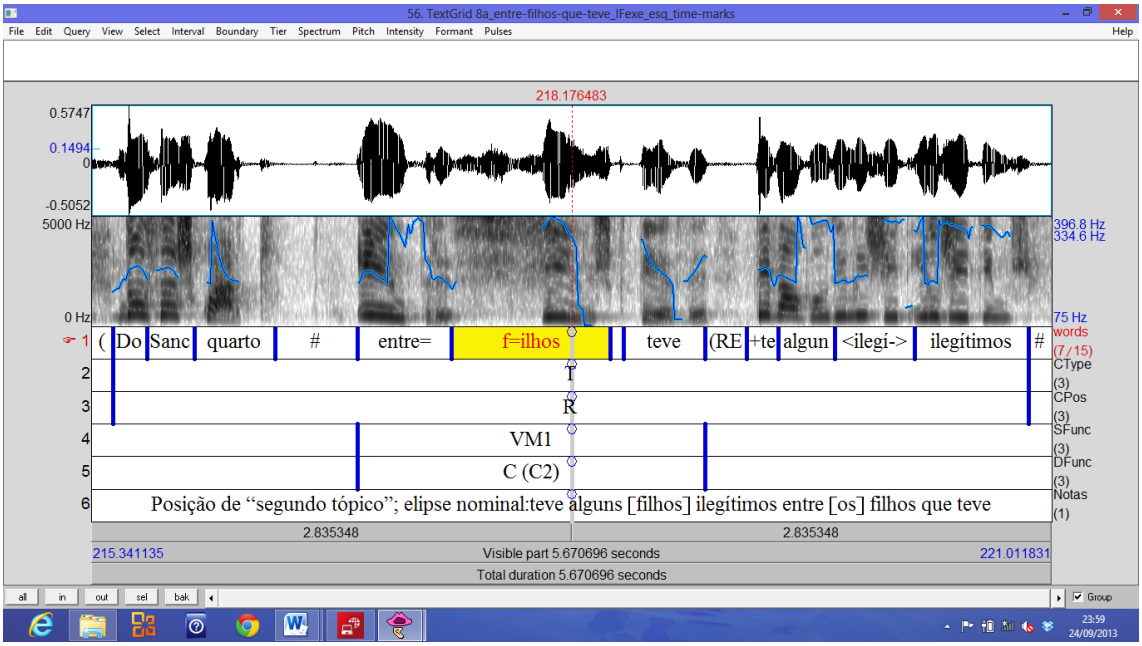



Figura 25: análise do exemplo (21) na plataforma informática Praat

(22) [...] <u>para ti</u> quais são as diferentes classes sociais [...]

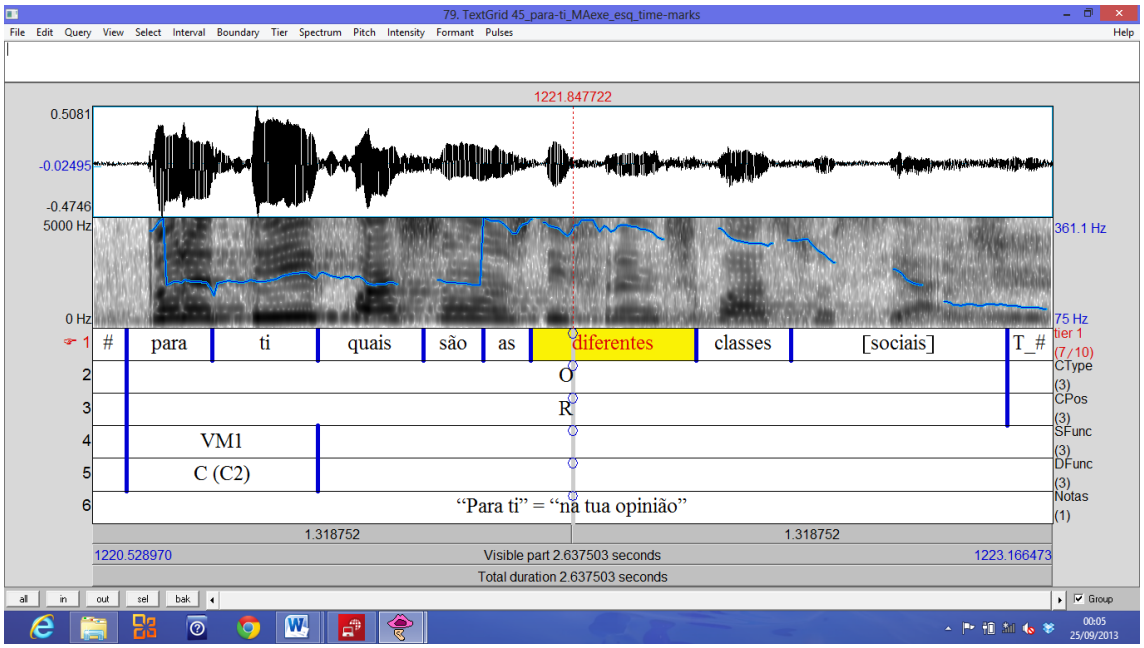



Figura 26: análise do exemplo (22) na plataforma informática Praat

(23) [...] outro <u>na cama</u> era quando estava lá o Castelhano [...]

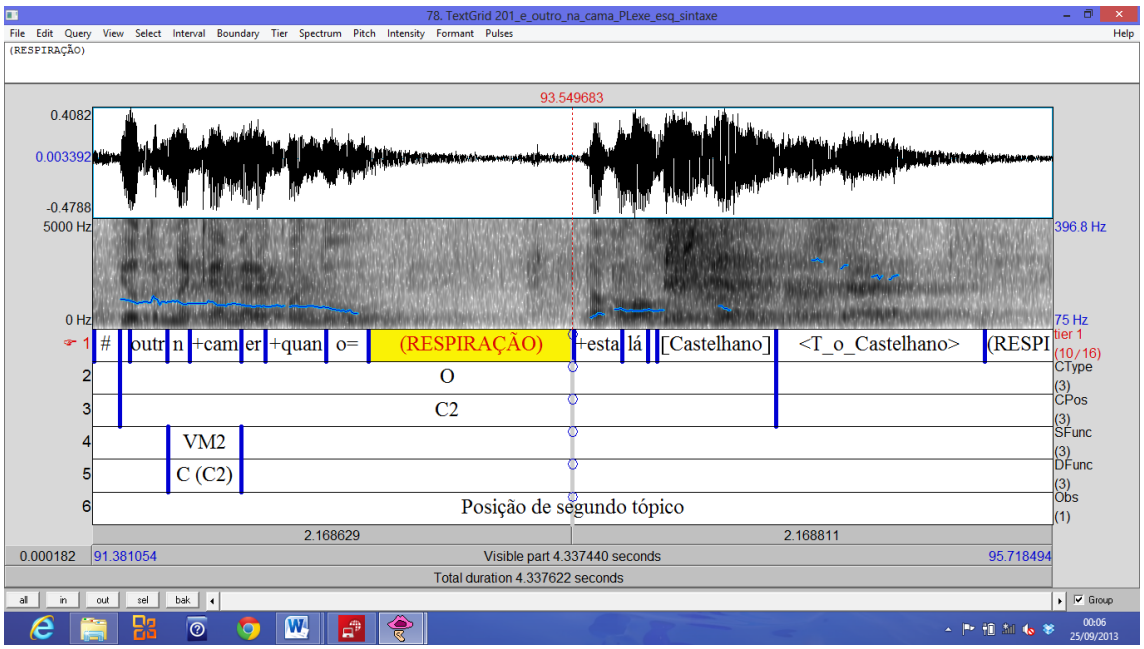



Figura 27: análise do exemplo (23) na plataforma informática Praat

## *4.* Apresentação de discussão dos resultados

Apresentam-se de seguida os resultados obtidos na anotação dos 116 exemplos extraídos do corpus, passando-se de seguida à discussão desses resultados.

### *4.1.* Apresentação dos resultados

Os resultados das ocorrências dos diferentes tipos de tópicos analisados são os contantes do gráfico 1.

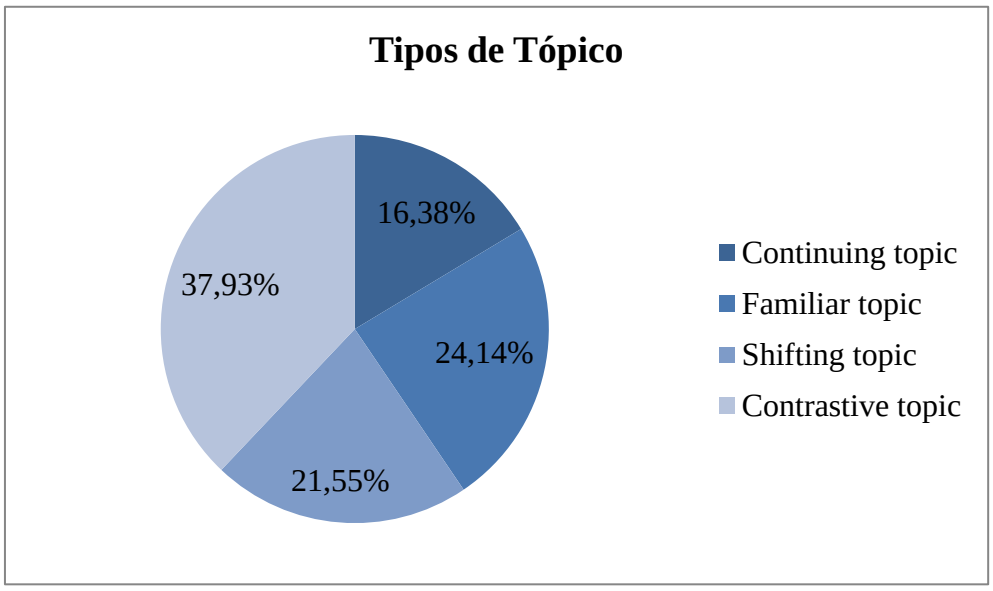


Gráfico 1: ocorrência de tópicos na amostra anotada

No conjunto dos dados analisados, foram encontradas:

- 19 ocorrências de constituintes marcados como *continuing topic*, o que corresponde a aproximadamente 16,38% do total dos exemplos analisados;
- 28 exemplos forma marcados como *familiar topic*, o que corresponde a aproximadamente 24,14%;
- 25 ocorrências foram registadas como *shifting topic*, correspondentes a 21,55% do total dos exemplos analisados;
- 44 exemplos foram marcados como contrastive topic, o correspondente a 37,93% do *corpus*.

Tendo em conta os diferentes subtipos de tópico, foram registadas 14 ocorrências de smooth-shifting topic e 11 ocorrências de rough-shifting topic. As primeiras correspondem a aproximadamente 12, 07% do total de exemplos analisados, enquanto as segundas correspondem a 9,48%, como mostra o gráfico 2.

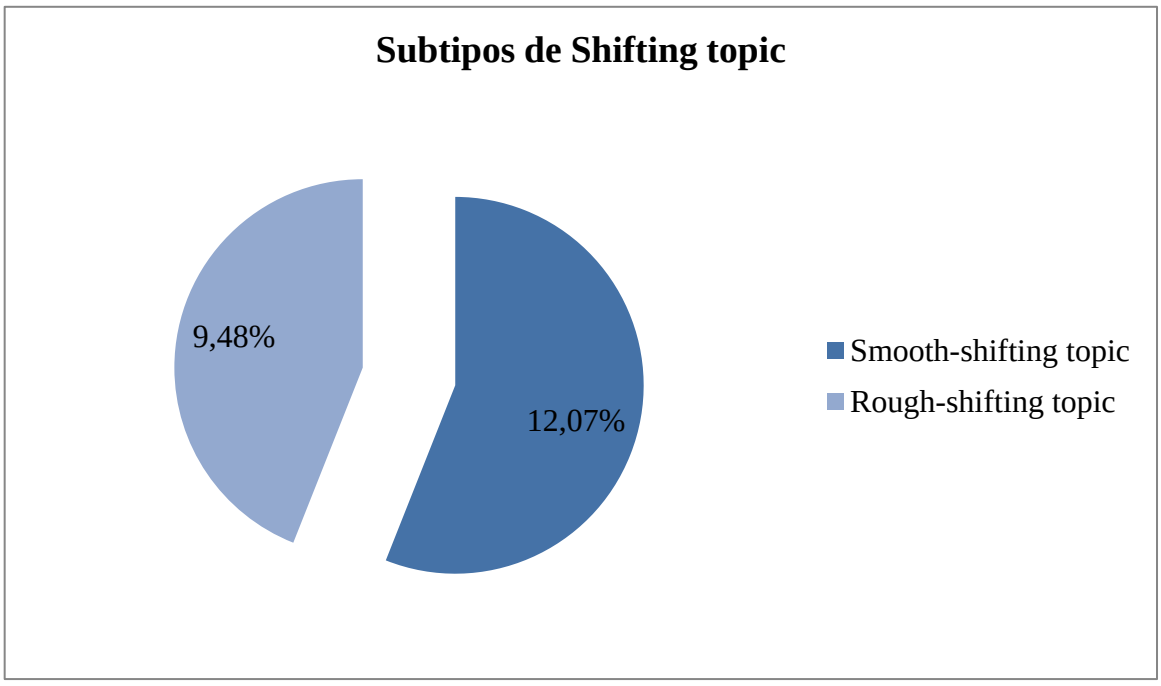



Gráfico 2: ocorrência de subtipos de shifting topic no *corpus*

Como exemplos de contrastive topic 1 foram classificados 33 excertos discursivos, o que corresponde a 28,44% do total do corpus estudado. Os exemplos anotados como contrastive topic 2 registaram 11 ocorrências, que correspondem a cerca de 9,48% do total do corpus. O gráfico 3 ilustra estes resultados.

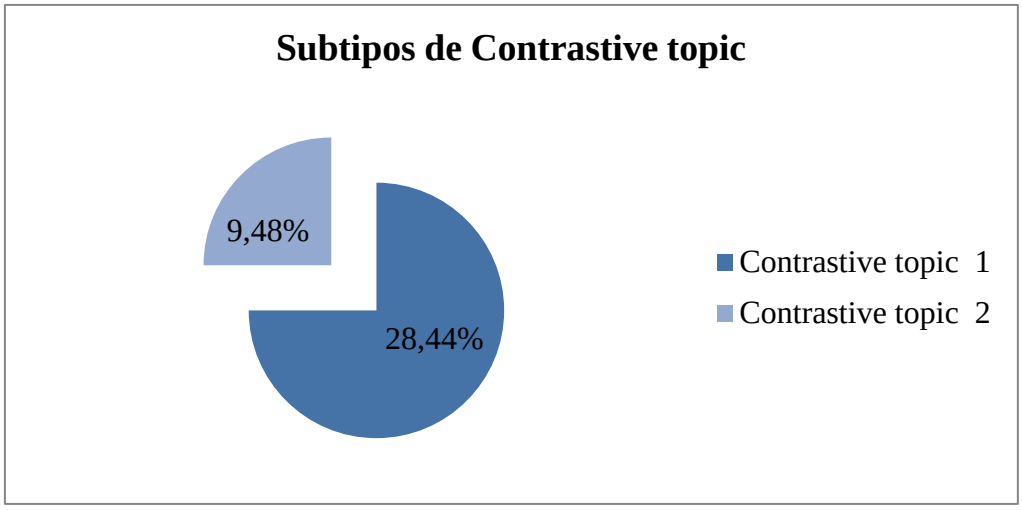



Gráfico 3: Ocorrência de subtipos de *contrastive topic* no *corpus*.

No que se refere à marcação linguística associada a cada um dos tipos de tópico discriminados, apresentamos, no quadro 1, uma síntese de alguns dos marcadores mais relevantes. Como vimos no primeiro capítulo, há tipicamente uma relação entre a expressão do tópico e a estratégia escolhida pelo locutor para a sua expressão linguística, pelo que é possível encontrar determinadas regularidades ou categorias associadas a cada um desses tópicos. Note-se, no entanto, que no caso em que se verifica mudança de tópico, o contexto discursivo surge como fator determinante para a sua marcação.

| **Tópicos discursivos e marcadores linguísticos** | | |
|---|---|---|
| **Tipos de Tópico** | **Triggers linguísticos** | **Exemplos** |
| **Continuing topic** | Sintagmas preposicionais<br>Relações anafóricas | Nessa manhã [...]/ Nesse dia [...]<br>Isso tudo [...]/ Isto [...] |
| **Familiar topic** | Sintagmas nominais | O resto [...]/ Camões também ele [...] |

| | | |
|---|---|---|
| **Smooth-shifting topic** | Contexto | [...]posteriormente teriam de ser lembradas vezes sem conta, bom, sobre Inês [...]/ [...] vamos ao trabalho, eu só quero que vocês se lembrem de uma expressãozinha [...] |
| **Rough-shifting topic** | Contexto | [...] narrativas em que este amor é lembrado, quanto ao beija-mão [...]/ [...] essa precisão, trinta segundos é fantástica, quem é que falta, ninguém e as faltas de atraso marco-as todas [...] |
| **Contrastive 1** | Conjunções<br>Sintagmas preposicionais | [...] se se amavam ou não [...] / [...] têm expressão ou não[...]<br>Ao colo [...] e na mão [...]/ um na rua [...] outro na cama [...] |
| **Contrastive 2** | Sintagmas preposicionais | [...] e entre as damas que a acompanhavam [...]/[...] para ti [...] |

Quadro 1: Tópicos discursivos e alguns marcadores linguísticos

### *4.2.* Breve discussão dos resultados

Dado o número reduzido de exemplos ainda analisado, não é possível estabelecer qualquer generalização acerca das percentagens obtidas na análise desta amostra. No entanto, é possível fazer o levantamento de algumas hipóteses que possam explicar, pelo menos, em parte esses resultados. De uma forma global, consideramos que esses resultados obtidos resultam do tipo de interação no qual é produzido o discurso, visto que tem um papel relevante na orientação, organização e estrutura discursivas, condicionando as opções discursivas e, por conseguinte, linguísticas dos interlocutores.

Através da análise e da classificação dos exemplos do corpus torna-se possível afirmar que, dos quatro tipos de tópico principais (*continuing*, *familiar*, *shifting* e *contrastive*), os excertos discursivos mais verificados correspondem ao *contrastive topic* (com um total de 44 ocorrências). Tais resultados podem dever-se ao facto de o *corpus* ser constituído por um vasto conjunto de conversações espontâneas registadas em contexto de exposição de conteúdos em que se pretende a interação com os interlocutores. Estas conversações

foram recolhidas em contexto de sala de aula, o que motiva a interação entre os interlocutores a partir da apresentação de alguns contrastes que lhes permitam fazer as escolhas consideradas mais adequadas. Esta interação origina, também, uma mudança frequente de tópico, daí que os exemplos classificados como *shifting* tenham obtido 25 ocorrências, enquanto os exemplos anotados como *continuing topic* tenham sido os menos registados no *corpus*, contando apenas com 19 ocorrências. A mudança de tópico pode dever-se ao facto de por ser um grupo que envolve vários interlocutores e por ser uma conversa espontânea o tópico se vá alterando, causando o aparecimento de novos tópicos.

Os exemplos anotados como *familiar topic* registam mais ocorrências (28 ocorrências) do que os exemplos classificados como *continuing topic*. Esta diferença pode ser justificada mais uma vez pelo contexto, dado que, por vezes, verifica-se a introdução de outros tópicos no meio do discurso e a retoma dos anteriores mais à frente.

Dos subtipos apresentados para o *shifting topic*, verifica-se que os exemplos anotados como *smooth-shifting* apresentam mais ocorrências (14 ocorrências) do que os exemplos classificados como *rough-shifting* (11 ocorrências). Apesar da constante mudança de tópico motivada pela interação entre os interlocutores, a introdução de um novo tópico no discurso leva a que esse determinado tópico se mantenha ativo durante um certo período de tempo.

Esta diferença também se verifica no que diz respeito aos subtipos apresentados para o *contrastive topic*. Neste caso, o subtipo *contrastive 1* regista mais ocorrências (33 ocorrências) do que o subtipo *contrastive 2* (que regista apenas 11 ocorrências, tal como o subtipo *rough-shifting topic*). Como justificação para esta disparidade de valores entre os dois subtipos de *contrastive topic*, enuncia-se a hipótese de que, no contexto de sala de aula como o que é representado pelo corpus aqui estudado, seja mais comum a realização de um contraste em que duas (ou mais) entidades se encontrem em oposição. O próprio contexto (exposição de conteúdos escolares em sala de aula) pode também justificar os resultados verificados para o *contrastive topic* 2, na medida em que o papel do professor é lançar escolhas/alternativas aos alunos.

A classificação dos vários segmentos discursivos como diferentes tipos de tópico revela que, no discurso oral, os locutores recorrem a diversas formas que lhes permitam conduzir os interlocutores na interpretação da conversação.

A partir da atribuição das categorias alusivas aos diferentes tipos de tópico aos exemplos que constituem o corpus do projeto, é possível constatar a importância que o contexto adquire aquando da análise discursiva, dado que para se classificar um constituinte como sendo um determinado tipo de tópico é necessário recorrer sempre ao contexto, quer antecedente, quer precedente. A interação entre os interlocutores característica do contexto em que os dados do corpus foram produzidos e recolhidos, propicia a heterogeneidade dos resultados encontrados na análise. Por esta razão é, também, importante ter presente as características da situação em que os dados são produzidos (contexto de sala de aula, no caso do corpus estudado neste trabalho). Neste sentido, o próprio contexto condiciona a anotação dos segmentos discursivos, podendo suscitar dúvidas ao anotador durante o processo de classificação dos tópicos. Este fator originou a redefinição da terminologia de classificação dos tópicos ao longo deste trabalho, pois a aplicação das categorias semânticas aos exemplos do corpus colocou diversas dúvidas sobre o tipo de tópico a adotar.

## Considerações Finais

No momento de terminar esta fase preliminar de anotação do corpus CPE FACES, que consistiu na anotação de 116 exemplos, refletimos sobre o percurso efetuado no sentido de fazer o balanço do trabalho realizado, no sentido de avaliar o cumprimento dos objetivos propostos. Ao longo deste trabalho tivemos essencialmente duas preocupações fundamentais. Por um lado, procurámos conhecer alguma da literatura mais relevante sobre a forma como se organiza e estrutura a informação ao nível da frase, mas sobretudo do discurso, nomeadamente a que se ocupa de questões como estrutura discursiva, estrutura informacional, tópico, foco, dado, novo e referenciação.

A partir dessas referências, construímos um quadro teórico de base para a parte prática do trabalho, que consistiu na anotação de um *corpus,* com o objetivo de estabelecer uma ferramenta de análise operatória para a segmentação e caracterização de um conjunto de tópicos ocorrentes na periferia esquerda da frase. Esse trabalho, realizado em vários momentos e com a discussão, a diferentes tempos, com os colegas da área da semântica e da área da sintaxe, resultou basicamente num duplo resultado: i) foi estabelecida uma base comum de tratamento dos dados, atendendo à especificidade de cada domínio de análise; ii) foi testada a etiquetagem proposta e a metodologia seguida na anotação.

Esse processo, que foi sendo sujeito a sucessivos reajustamentos, foi positivo, na medida em que contribuiu, ainda que por enquanto em muito pequena medida, para a criação de uma ferramenta de análise de corpora do português no domínio da organização da informação. Trata-se de um resultado importante, visto que as propostas de anotação semântica do discurso são escassas ou mesmo quase nulas para o português.

O aspeto positivo que se depreende deste resultado constitui ao mesmo tempo uma das limitações que se colocam à sua realização, dado que isso implicou a falta de referências e a necessidade de mais testes e reformulações.

Acresce a este aspeto o facto de se trabalhar com discurso oral espontâneo e planeado, com as suas especificidades diversas, nomeadamente ao nível da organização da informação. A estrutura informacional de um texto contempla fatores linguísticos e não linguísticos. Neste sentido, ao analisar a estrutura informacional de um determinado tipo de discurso, o investigador deve ter em conta os processos sintáticos, semânticos e prosódicos que lhe estão subjacentes e também o meio extralinguístico que o rodeia e os processos psicológicos que conduzem à interpretação das unidades discursivas.

Sendo um conceito que faz parte da estrutura informacional do texto, o tópico pode ser analisado a partir de diferentes perspetivas: sintática, semântica e prosódica. Cada um destes pontos de vista apresenta características específicas (como a posição do constituinte marcado como tópico, no caso da sintaxe, a interpretação e o significado do referentes discursivo, no caso da semântica e a acentuação, no caso da prosódia). Como é possível verificar no decorrer desta dissertação, as várias entidades discursivas podem desempenhar diferentes papéis ao longo da conversação. Esta dinâmica advém do facto de a informação que é transmitida pelo locutor poder ser absorvida pelo interlocutor em diferentes formas, tendo em conta os seus conhecimentos sobre as próprias entidades linguísticas e sobre o mundo que o rodeia.

Do ponto de vista linguístico, o tópico e o foco são as duas grandes categorias linguísticas que estão subjacentes à organização da estrutura do enunciado e onde se agrupam as várias entidades que o constituem. Apesar de neste trabalho e também no projeto COPAS se focar apenas a marcação e anotação de diferentes tipos de tópico no discurso oral, não é possível excluir o conceito de foco, dado que estas duas noções são complementares, não se excluindo mutuamente. Por isso se prevê também o tratamento deste fenómeno no âmbito do projeto e dos corpora já referidos.

Assim, apesar dos resultados obtidos através da análise deste corpus, o tema da atribuição de diferentes categorias semânticas aos vários tópicos encontrados no discurso oral necessita ainda de uma investigação mais aprofundada. É necessário ampliar o corpus

e diversificar as fontes, para conseguir validar de forma segura as etiquetas propostas neste trabalho, mas também definir padrões discursivos e linguísticos para cada tipo de tópico e para o conjunto dos tópicos tratados. Nesse contexto, a articulação dos resultados obtidos com os resultados obtidos na anotação da mesma amostra pela sintaxe e a prosódia serão uma contribuição para o entendimento mais profundo de como efetivamente processamos a informação no discurso e de que forma isso reflete a nossa atuação como falantes.

## Referências bibliográficas


Ariel, M. (1988) - *Referring and and accessibility*. Journal of Linguistics, 24, 65-87.

Ariel, M. (1990) - *Accessing noun phrases antecedents*. London: Routledge

Asher, N. 1999. Discourse and the Focus/Background Distinction. In Bosch, P. & van der Sandt, R. (eds.) 1999. *Focus: Linguistic, Cognitive and Computational Perspectives*. Cambridge: Cambridge University Press, pp. 247-267

Barbosa, J. 2005. *Foco e Tópico: algumas questões terminológicas*. Porto: Universidade do Porto – Faculdade de Letras, pp. 339-351, disponível em http://aleph20.letras.up.pt/exlibris/aleph/a20_1/apache_media/C9GVN7857TET2T345V4XYRP2UF8RHV.pdf, acedido em 09-10-2012

Bolinger, D. 1954. English prosodic stress and Spanish sentence order. In *Hispania*, 37, pp. 152-156, referido por Lambrecht, K. 1994. Information structure and sentence form – topic, focus and the mental representations of discourse referents. In Andreon, S.R.; Bresnan, J.; Combrie, B.; Dressler, W.; Ewen, C.; Huddleston, R.; Lass, R.; Lightfoot, D.; Lyons, J.; Matthews, P. H.; Posner, R.; Romaine, S.; Smith, N.V.; Vincent, N. (eds) 1994. *Cambridge studies in linguistics*, 71, Cambridge: Cambridge University Press

Braver, A. 2007. Reinhart 1982- Pragmatics and Linguistics: an analysis of sentence topics, pp. 1-5, disponível em http://people.brandeis.edu/~smalamud/ling140/3-13-reinhart1982-AaronBraver.pdf, acedido em 28-09-2013

Büring, D. 1999. Topic. In Bosch, P. & van der Sandt, R. (eds.).1999. Focus. Linguistic, Cognitive and Computational Perspectives. Cambridge: Cambridge University Press, pp. 142.165. In Gutiérrez-Rexach, J. (ed.) 2003. *Semantics: Critical concepts in linguistics*, vol. 6. London: Routledge, pp. 84-109

Büring, D. 1999. Topic. In Bosch, P. & Van der Sand, R. (eds.). Focus – Linguistic Cognitive and Computational Perspectives. Cambridge: Cambridge University Press, pp. 142-165, referido por Frascarelli, M. & Hinterhölzl, R. 2002. Types of topics in german and Italian.

In Winkler, S. & Schwabe, K. (eds.). *On information structure, meaning and form*. Amsterdam/Philadelphia: John Benjamins, pp.1-29, retirado de https://www.dropbox.com/home/COPAS_Team/bibl, acedido em 15-04-2013

Calhoun, S.; Nissim, M.; Steedman, M. & Brenier, J. 2005. *A Framework for Annotating Information Structure in Discourse*, pp.1-8, disponível em http://groups.inf.ed.ac.uk/switchboard/infostruc.pdf, acedido em 02-08-2013

Clark, H. & Haviland, S. 1977. Comprehension and the given-new contract. In Freedle, R. (ed.). *Discourse Production and Comprehension*. Hillsdale, N. J.: Lawrence Erlbaum Associates, pp. 1-40, referido por Prince, E. F. 1981. *Toward a Taxonomy of Given-New Information*. Academic Pres, Inc., pp. 223-255, disponível em ftp://babel.ling.upenn.edu/facpapers/ellen_prince/givennew.pdf, acedido em 02-08-2013

Chafe, W. 1987. Cognitive constraints on information flow. In Tomlin (ed.), pp. 21-52, referido por Lambrecht, K. 1994. *Information structure and sentence form – topic, focus and the mental representations of discourse referents*. In Andreon, S.R.; Bresnan, J.; Combrie, B.; Dressler, W.; Ewen, C.; Huddleston, R.; Lass, R.; Lightfoot, D.; Lyons, J.; Matthews, P. H.; Posner, R.; Romaine, S.; Smith, N.V.; Vincent, N. (eds) 1994. *Cambridge studies in linguistics*, 71, Cambridge: Cambridge University Press

Chafe, W. 1987. Cognitive constrains on information flow. In Tomli, R. /ed.). Coherence and grounding in discourse. Amsterdam: Benjamins, pp. 21-51, referido por Frascarelli, M. & Hinterhölzl, R. 2002. Types of topics in german and Italian. In Winkler, S. & Schwabe, K. (eds.). *On information structure, meaning and form*. Amsterdam/Philadelphia: John Benjamins, pp.1-29, retirado de https://www.dropbox.com/home/COPAS_Team/bibl, acedido em 15-04-2013

Chafe, Wallace L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. L (ed.). Subject and Topic. Nova Iorque: Academic Press, 25-55.

Cowles, H. W.; Walenski, M. & Kluender, R. 2007. Linguistic and cognitive prominence in anaphor resolution: topic, contrastive focus and pronouns. In *Topoi*. 2007, vol. 26, pp. 3-18, disponível em http://link.springer.com/article/10.1007/s11245-006-9004-6#, acedido em 24-04-2013.

Dahl, O. 1976. What is new information?. In Erk Enkvist, N. & Kohonen, V. (eds.) *Approaches to word order*. Reports in Text Linguistics No. 72. Meddelanden fran Stiftelsens för Abo Akademi Forskningsinstitut, 8. Abo/Turku, referido por Lambrecht, K. 1994. Information structure and sentence form – topic, focus and the mental representations of discourse referents. Andreon, S.R.; Bresnan, J.; Combrie, B.; Dressler, W.; Ewen, C.; Huddleston, R.; Lass, R.; Lightfoot, D.; Lyons, J.; Matthews, P. H.; Posner, R.; Romaine, S.; Smith, N.V.; Vincent, N. (eds) 1994. *Cambridge studies in linguistics*, 71, Cambridge: Cambridge University Press

Degand, Lisbeth & Anne Catherine Simon (2009). On identifying basic discourse units in speech: theoretical and empirical issues. Discours, 4, disponível em http://discours.revues.org/5852 (acedido em 24 de março de 2013).

Diesing, M. 1989. Bare plural subjects, inflection, and the mapping to LF. In Bach, Kratzer & Partee (eds.) *Papers on Quantification*. NSF report, UMass, Amherst, referido por Partee, B. 1991. Topic, focus and quantification. In Partee, B. 1991 *Proceedings of Semantics and Linguistic Theory (SALT)*, 1. Ithaca, NY: CLC Publications, Cornell University, pp. 159-187. In Gutiérrez-Rexach, J. (ed.) 2003. *Semantics: Critical concepts in linguistics*, vol. 6. London: Routledge, pp. 7-32

Duarte, I. 2003. Aspetos linguísticos da organização textual. In Mateus, M. H. et ali. 2003. *Gramática da Língua Portuguesa*. Lisboa: Caminho, pp. 87-123

Erteschik-Shir, N. 1997. *The dynamics of focus structure*. Cambridge: Cambridge University Press, referido por Barbosa, J. 2005. *Foco e Tópico: algumas questões terminológicas*. Porto: Universidade do Porto – Faculdade de Letras, pp. 339-351, disponível em http://aleph20.letras.up.pt/exlibris/aleph/a20_1/apache_media/C9GVN7857TET2T345V4XYRP2UF8RHV.pdf, acedido em 09-10-2012

Féry, C. & Krifka, M. 2008. Information structure: notional distinctions, ways of expression. In van sterkenburg, P. (ed.) 2008. *Unity and diversity of languages*. Amesterdam: John Benjamins, pp. 123-136, disponível em http://amor.cms.hu-berlin.de/~h2816i3x/Publications/Fery_Krifka_InformationStructure.pdf, acedido em 28-09-2013

Frascarelli, M. & Hinterhölzl, R. 2002. Types of topics in german and Italian. In Winkler, S. & Schwabe, K. (eds.). *On information structure, meaning and form*. Amsterdam/Philadelphia: John Benjamins, pp.1-29, retirado de https://www.dropbox.com/home/COPAS_Team/bibl, acedido em 15-04-2013

Garrod, S. 1995. Distinguishing between Explicit and Implicit Focus during Text Comprehension. In Rickheit, G. & Habel, C. (eds.) 1995. *Focus and Coherence in Discourse Processing*. New York: Walter de Gruyter, pp. 3-16

Grosz, B. & Sidner, C. 1986. Attention, intentions, and the structure of discourse. In *Journal of Computational linguistics*, nº3, vol. 12. Cambridge, MA: MIT Press, pp. 175-204, retirado de http://delivery.acm.org/10.1145/20000/12458/p175-grosz.pdf?ip=37.28.229.148&acc=OPEN&key=1B55DF923F77674F55057ED4F3766CA0&CFID=226590416&CFTOKEN=29066923&__acm__=1371228335_c8d85d6334303a5ca3d999ca171a8074, acedido em 25-05-2013

Gundel, J. K., Hedberg, N. & Zacharski, R. (1993) - Cognitive status and the form of referring expressions in discourse. *Language*, 69 (2), 274-307.

Gundel, J. 1999. On Different Kinds of Focus. In Bosch, P. & van der Sandt, R. (eds.) 1999. *Focus: Linguistic, Cognitive and Computational Perspectives*. Cambridge: Cambridge University Press, pp. 293-305.

Gundel, J. K. s/data. *Topic, Focus, and the Grammar-Pragmatics Interface*, pp.1-16, disponível em http://ancienthebrewpoetry.typepad.com/ancient_hebrew_poetry/files/gundel.pdf, acedido em 02-08-2013

Gyuris, B. 2004. Two Types of Contrastive Topics? In Steube, A. (ed.) 2004. *Language, Context and Cognition: Information Structure Theoretical and Empirical Aspects*, vol.1. New York: Walter de Gruyter, pp. 75-96

Hajičová, E. & Sgall, P. 2004. Degrees of Contrast and the Topic-Focus Articulation. In Steube, A. (ed.) 2004. *Language, Context and Cognition: Information Structure Theoretical and Empirical Aspects*, vol.1. New York: Walter de Gruyter, pp. 1-12

Hajičová, E. 2012. On scalarity in information structure. In *Linguística Pragensiana*, 22, pp. 60-78, disponível em http://www.degruyter.com/view/j/lipr.2012.22.issue-2/v10017-012-0005-6/v10017-012-0005-6.xml, acedido em 28-09-2013

Halliday, M. 1967b. Notes on Transitivity and Theme in English (Part 1 & 2). In *Journal of Linguistics*, 3, pp. 37-81; 199-244, referido por Barbosa, J. 2005. *Foco e Tópico: algumas questões terminológicas*. Porto: Universidade do Porto – Faculdade de Letras, pp. 339-351, disponível em http://aleph20.letras.up.pt/exlibris/aleph/a20_1/apache_media/C9GVN7857TET2T345V4XYRP2UF8RHV.pdf, acedido em 09-10-2012

Halliday, M. A. K. 1967. Notes on transitivity and theme in English. Part 2. In *Journal of Linguistics*, 3, pp. 199-244, referido por Prince, E. F. 1981. *Toward a Taxonomy of Given-New Information*. Academic Pres, Inc., pp. 223-255, disponível em ftp://babel.ling.upenn.edu/facpapers/ellen_prince/givennew.pdf, acedido em 02-08-2013

Halliday, M.A.K. (1985). An Introduction to Functional Grammar. Londres: Edward Arnold

Kadmon, N. 2000. *Formal Pragmatics*. Oxford: Blackwell Publishers, referido por Barbosa, J. 2005. *Foco e Tópico: algumas questões terminológicas*. Porto: Universidade do Porto – Faculdade de Letras, pp. 339-351, disponível em http://aleph20.letras.up.pt/exlibris/aleph/a20_1/apache_media/C9GVN7857TET2T345V4XYRP2UF8RHV.pdf, acedido em 09-10-2012

Kadmon, N. 2001. *Formal Pragmatics*. Malden, MA. Oxford: Blackwell, referido por Gyuris, B. 2004. Two Types of Contrastive Topics? In Steube, A. (ed.) 2004. *Language, Context and Cognition: Information Structure Theoretical and Empirical Aspects*, vol.1. New York: Walter de Gruyter, pp. 75-96

Keenan-Ochs, E. & Schieffelin, B. 1976z. *Topics as a discourse notion*. In Li (1976), referido por Reinhart, T. 1981. Pragmatics and linguistics: an analysis of sentence topics. In *Philosophica*, 27, pp. 53-94, disponível em http://logica.ugent.be/philosophica/fulltexts/27-4.pdf, acedido em 28-09-2013

Kratzer, A. 1989b. Stage-Level and Indicidual-Level predicates. In Bach, Kratzer & Partee (eds.) *Papers on quantification*. NSF report, UMass, referido por Partee, B. 1991. Topic, focus and

quantification. In Partee, B. 1991 *Proceedings of Semantics and Linguistic Theory (SALT)*, 1. Ithaca, NY: CLC Publications, Cornell University, pp. 159-187. In Gutiérrez-Rexach, J. (ed.) 2003. *Semantics: Critical concepts in linguistics*, vol. 6. London: Routledge, pp. 7-32

Krifka, M. 2003. A compositional semantics for multiple focus constructions. Austin: University of Texas, referido por Partee, B. 1991. Topic, focus and quantification. In Partee, B. 1991 *Proceedings of Semantics and Linguistic Theory (SALT)*, 1. Ithaca, NY: CLC Publications, Cornell University, pp. 159-187. In Gutiérrez-Rexach, J. (ed.) 2003. *Semantics: Critical concepts in linguistics*, vol. 6. London: Routledge, pp. 7-32

Krifka, M. 1991. A compositional semantics for multiple focus constructions. In Krifka, M. 1991. *Proceedings of Semantics and Linguistic Theory (SALT)*, 1. Ithaca, NY: CLC Publications, Cornell University, pp. 127-158. In Gutiérrez-Rexach, J. (ed.) 2003. *Semantics: Critical concepts in linguistics*, vol. 6. London: Routledge, pp. 33-72

Krifka, M. 2001. *Topik und Fokus*. Wintersemester: Humboldt-Universität, pp.1-4, disponível em http://amor.cms.hu-berlin.de/~h2816i3x/Lehre/2001_HS_TopikFokus/TopikFokus8.pdf, acedido em 28-09-2013

Krifka, M. 2006. Basic notions of information structure. In Féry, C. and Krifka, M. (eds.), *Interdisciplinary Studies of Information Structure* 6, Potsdam. Também in *Acta Linguistica Hungarica* 55 (2008), pp. 243-276.. A translation in Georgian appeared in Language Logic Computation II, CLLS, Tbilisi State University, 2008, 7-50, disponível em http://amor.cms.hu-berlin.de/~h2816i3x/Publications/Krifka_InformationStructure.pdf, acedido em 28-09-2013

Kuno, S. 1976. Subject, theme, and the speaker's empathy – A reexamination of relativization phenomena. In Li, C. (ed.) Subject and Topic. New York: Academic Press, pp. 1-29, referido por Frascarelli, M. & Hinterhölzl, R. 2002. Types of topics in german and Italian. In Winkler, S. & Schwabe, K. (eds.). *On information structure, meaning and form*. Amsterdam/Philadelphia: John Benjamins, pp.1-29, retirado de https://www.dropbox.com/home/COPAS_Team/bibl, acedido em 15-04-2013

Kuno, S. 1978. Generative discourse analysis in America. In Dressler, W. (ed.) *Current Trends in Textlinguistics*. Berlim e New York: de Gruyter, pp. 275-294, referido por Prince, E. F. 1981. *Toward a Taxonomy of Given-New Information*. Academic Pres, Inc., pp. 223-255, disponível em ftp://babel.ling.upenn.edu/facpapers/ellen_prince/givennew.pdf, acedido em 02-08-2013

Lambrecht, K. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press, referido por Barbosa, J. 2005. *Foco e Tópico: algumas questões terminológicas*. Porto: Universidade do Porto – Faculdade de Letras, pp. 339-351, disponível em http://aleph20.letras.up.pt/exlibris/aleph/a20_1/apache_media/C9GVN7857TET2T345V4XYRP2UF8RHV.pdf, acedido em 09-10-2012

Lambrecht, K. 1994. Information structure and sentence form – topic, focus and the mental representations of discourse referents. In Andreon, S.R.; Bresnan, J.; Combrie, B.; Dressler, W.; Ewen, C.; Huddleston, R.; Lass, R.; Lightfoot, D.; Lyons, J.; Matthews, P. H.; Posner, R.; Romaine, S.; Smith, N.V.; Vincent, N. (eds) 1994. *Cambridge studies in linguistics*, 71, Cambridge: Cambridge University Press

Mata, Ana Isabel (1995). Apresentação do CPE FACES – Um 'corpus de Português Falado por Adolescentes em Contexto Escolar', para um Estudo de Prosódia dos Estilos de Fala. In Nascimento, M. Fernanda Bacelar, Celeste Rodrigues & José Bettencourt Gonçalves (orgs.). Actas do XI Encontro Nacional da Associação Portuguesa de Linguística, Lisboa: Colibri, 349-358.

Max, I. 2004. Information Structure – Two-dimensionally Explicated. In Steube, A. (ed.) 2004. *Language, Context and Cognition: Information Structure Theoretical and Empirical Aspects*, vol.1. New York: Walter de Gruyter, pp. 97-123

Myackykov, A.; Garrod, S. & Scheepers, C. 2009. Attention and syntax in sentence production: A critical review. In Revues.org. 2009. *Discours*, nº4, pp. 1-17, disponível em http://discours.revues.org/7594, acedido em 09-07-2013

Müsseler, J. 1995. Focussing and the Process of Pronominal Resolution. In Rickheit, G. & Habel, C. (eds.) 1995. *Focus and Coherence in Discourse Processing*. New York: Walter de Gruyter, pp. 53-74

Partee, B. 1991. Topic, focus and quantification. In Partee, B. 1991 *Proceedings of Semantics and Linguistic Theory (SALT)*, 1. Ithaca, NY: CLC Publications, Cornell University, pp. 159-187, In Gutiérrez-Rexach, J. (ed.) 2003. *Semantics: Critical concepts in linguistics*, vol. 6. London: Routledge, pp. 7-32

Partee, B.; Bach, E. Kratzer, A. 1987. Quantiffication: A cross-linguistic perspective, NSF proposal, UMass, Amherst, referido por Partee, B. 1991. Topic, focus and quantification. In Partee, B. 1991 *Proceedings of Semantics and Linguistic Theory (SALT)*, 1. Ithaca, NY: CLC Publications, Cornell University, pp. 159-187. In Gutiérrez-Rexach, J. (ed.) 2003. *Semantics: Critical concepts in linguistics*, vol. 6. London: Routledge, pp. 7-32

Prince, E. F. 1997. On the functions of left-dislocation in English discourse. in Kamio, A. (ed.) *Directions in functional linguistics*. Amsterdam and Philadelphia: John Benjamins, pp. 117-43, referido por Ward, G. & Birner, B.2001. *Discourse and information structure*. Evanston: Northwestern University, pp.1-42, disponível em http://www.ling.northwestern.edu/~ward/HDA01.pdf, acedido em 28-09-2013

Prince, E. F. 1981. *Toward a Taxonomy of Given-New Information*. Academic Pres, Inc., pp. 223-255, disponível em ftp://babel.ling.upenn.edu/facpapers/ellen_prince/givennew.pdf, acedido em 02-08-2013

Prince, E. F. 1991. *On the functions of left-dislocations in English discourse*. Pensnylvania: University of Pensnylvania, pp. 1-25, disponível em http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=91BECFEBFB3DB022576E7E8E325C48C7?doi=10.1.1.32.2672&rep=rep1&type=pdf, acedido em 28-09-2013

Reinhart, T. 1981. Pragmatics and linguistics: an analysis of sentence topics. In *Philosophica*, 27, pp. 53-94, disponível em http://logica.ugent.be/philosophica/fulltexts/27-4.pdf, acedido em 28-09-2013

Reinhart, T. 1981. Pragmatics and linguistics: an analysis os sentence topics. In *Philosophica*, 27, pp. 53-94, referido por Barbosa, J. 2005. *Foco e Tópico: algumas questões terminológicas*. Porto:

Universidade do Porto – Faculdade de Letras, pp. 339-351, disponível em http://aleph20.letras.up.pt/exlibris/aleph/a20_1/apache_media/C9GVN7857TET2T345V4XYRP2UF8RHV.pdf, acedido em 09-10-2012

Rooth, M. 1985. *Association with focus*. PhD dissertation. University of Massachusetts: Amherst, referido por Hajičová, E. & Sgall, P. 2004. Degrees of Contrast and the Topic-Focus Articulation. In Steube, A. (ed.) 2004. *Language, Context and Cognition: Information Structure Theoretical and Empirical Aspects*, vol.1. New York: Walter de Gruyter, pp. 1-12

Rooth, M. 1996. Focus. In Lappin & Shalom (org.). *The Handbook of Contemporary Semantic Theory*. Oxford: Blackwell Publishers, pp. 271-297, referido por Barbosa, J. 2005. *Foco e Tópico: algumas questões terminológicas*. Porto: Universidade do Porto – Faculdade de Letras, pp. 339-351, disponível em http://aleph20.letras.up.pt/exlibris/aleph/a20_1/apache_media/C9GVN7857TET2T345V4XYRP2UF8RHV.pdf, acedido em 09-10-2012

Runge, S. 2008. *Left- dislocations and their contribution to information structure*. SBL Annual Meeting, Nov. 22-25, pp- 1-17, disponível em http://www.ntdiscourse.org/docs/Runge_L-dis_SBL.pdf, acedido em 28-09-2013

Van Dijk, T. 2007. Cognitive context models and discourse. In Stamenow, M. (ed.). 1997. *Language Structure, Discourse and the Access to Consciousness*. Amesterdão: Benjamins, pp. 189-226, referido por Degand, L. & Simon, C. 2009. On identifying basic discourse units in speech: theoretical and empirical issues. In Revues.org. 2009. *Discours*, nº4, pp. 1-18, disponível em http://discours.revues.org/5852, acedido em 09-07-2013

Walker, M. A. & Prince, E. F. 2003. A Bilateral Approach to Givenness: A Hearer-Status and a Centering Algorithm. In Fretheim, T & Gundel, J. (eds.). Reference and referent accessibility. Amsterdam/Philadelphia: John Benjamins, pp. 291-306, disponível em http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.31.2238, acedido em 02-08-2013

Ward, G. & Birner, B.2001. *Discourse and information structure*. Evanston: Northwestern University, pp.1-42, disponível em http://www.ling.northwestern.edu/~ward/HDA01.pdf, acedido em 28-09-2013

Winkler, S. 1997. Focus structure in principle-based theory of grammar. In Winkler, S. 1997. *Focus and Secondary Predication*. New York: Mouton de Gruyter, pp. 93-124