

A prosódia no discurso espontâneo

*Sónia Delfina Amaral¹, Rosa Lídia Coimbra^{1,2},
António J. S. Teixeira^{3,4}*

¹) Centro de Línguas e Culturas, Universidade de Aveiro

²) Departamento de Línguas e Culturas, Universidade de Aveiro

³) Instituto de Engenharia Electrónica e Telemática de Aveiro (IEETA)

⁴) Departamento de Electrónica e Telecomunicações, Universidade de Aveiro

1. Introdução

Com a utilização crescente de sistemas de diálogo em interacção Humano-Máquina aumenta a necessidade de melhorar a naturalidade das tecnologias envolvidas. No que concerne à síntese de voz, mais ligada ao presente trabalho, uma das áreas mais activas consiste na modelação da prosódia. No entanto, grande parte dos trabalhos utilizam informação que não vai além do horizonte temporal da frase. Torna-se necessário estudar e modelar a interdependência da prosódia e estrutura dos diálogos, capitalizando nos estudos da área do discurso, como os clássicos de Austin e Searle.

Do lado dos estudos experimentais da Fonética na área da Prosódia este género de pesquisa permitirá estender os estudos já existentes e em curso, muitas vezes dedicados a produções de fala demasiado controladas.

O interesse duplo em contribuir para a melhoria de sistemas automáticos de processamento da nossa língua e, ao mesmo tempo, contribuir para um aprofundamento dos conhecimentos base na área da Linguística advém da inserção deste trabalho no grupo interdisciplinar da área da Linguística e da Engenharia em formação na Universidade de Aveiro.

2. Objectivos

Este trabalho tem por objectivo investigar a inter-relação entre constituintes do discurso – por exemplo, os actos de fala – e a prosódia. O presente trabalho poderá ser um contributo para o estudo do discurso oral pretendendo dar continuidade a uma série de estudos já realizados no campo da prosódia e do discurso em Português. Em relação aos estudos da prosódia em curso na Universidade de Aveiro, esta pesquisa constitui a primeira incursão na análise de fala (quase) espontânea. Em relação ao discurso, permite passar da anteriormente estudada estrutura textual para o lado do discurso oral.

A questão primordial foi a de analisar que elementos prosódicos os falantes usam para marcar a estrutura do discurso e quais desses elementos são identificados como

relevantes nesse processo. Pretendeu-se, ainda, investigar possíveis alterações/variações na prosódia causadas pela estrutura do discurso.

3. Metodologia e corpus

Para a prossecução dos objectivos definidos, foram inventariadas e avaliadas várias possibilidades de obtenção de discurso espontâneo, tendo a escolha recaído na técnica de MapTask, seguindo o exemplo de vários estudos realizados para outras línguas. Nesta primeira fase, consideraram-se apenas discursos com dois intervenientes.

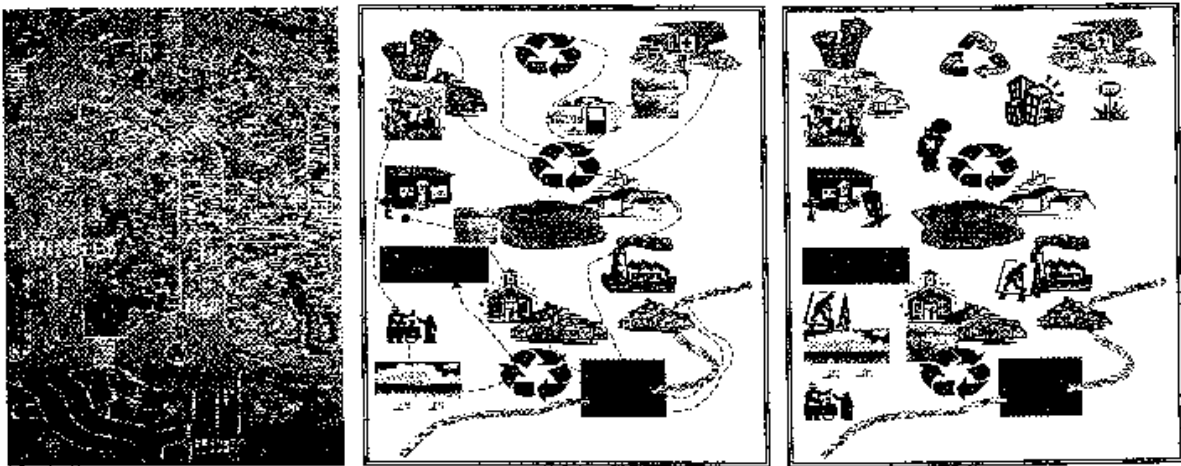


Figura 1: Mapas utilizados na tarefa MapTask

Optámos, também, por testar a metodologia num pré-inquérito. Este consistiu na utilização de um mapa tapete com planta de uma cidade para actividades lúdicas de crianças (figura 1, à esquerda), e sendo solicitado a um dos informantes que desse indicações de locais a visitar partindo de um mapa miniatura com um percurso previamente definido pelos autores. Neste pré-inquérito o material mais espontâneo foi conseguido aquando da introdução de obstáculos inesperados no percurso pelos investigadores. Com base nesta experiência, foram criados mapas de raiz, em papel, contendo elementos facilmente identificáveis (praia, casas, fábricas, pontes, lago, etc.). Foram fornecidos mapas ligeiramente diferentes (ver figura 1, imagens ao centro e à direita) aos intervenientes do inquérito final.

Produziu-se, deste modo, um corpus de diálogo seguindo o modelo do corpus HCRC do MapTask (Anderson et al. 1991), por se aproximar mais da conversação espontânea.

4. Informantes e gravação

Das várias possibilidades de obtenção de discurso espontâneo optou-se pela utilização de informantes pré-adolescentes. Desta forma, a influência de automatismos provenientes da escrita e da escolaridade será minimizada. A utilização de idades inferiores tornaria a gravação muito mais difícil.

No total, foram gravados diálogos com duas crianças, de 6 e 8 anos, e três pré-adolescentes, dois com 11 e um com 13 anos, todos falantes do Português.

As gravações originaram um corpus experimental decorrentes da gravação de um pré-teste com dois dos locutores, o de 8 e o de 13 anos, utilizando o mapa tapete e uma cópia do mesmo em papel com um percurso assinalado, a fim de testar a constituição linguística/discursiva do corpus e sua adequação na metodologia proposta ao estudo da prosódia.

Na gravação do corpus final, foram gravados quatro locutores, divididos em dois pares e prestando-se a duas gravações cada par, em que se invertiam os papéis *giver* – *follower*. Nas gravações deste corpus final foram utilizados os mapas em papel. O primeiro par era formado pelo informante de 8 e um informante de 6 anos; o segundo par era formado por dois informantes de 11 anos.

As gravações foram realizadas usando DAT no par de idade superior, e directamente para o disco duro do computador, usando o sistema de gravação KAY CSL 4400 da Kay Elemetrics, no outro caso. As gravações do pré-teste e do primeiro par de informantes do corpus final tiveram lugar no Laboratório de Fonética da Universidade de Aveiro; as restantes em sala de aula. Os intervenientes foram colocados frente a frente, com um painel separador ocultando os respectivos mapas, sendo o microfone colocado na posição intermédia a uma altura similar à da boca de ambos. A gravação de uma sessão foi efectuada na totalidade e para um único ficheiro, evitando-se a perda de qualquer informação e respectiva posição temporal.

5. Anotação

Para a anotação dos actos de fala e, partindo das formulações teóricas de Austin e Searle, recorreremos aos estudos de Sinclair e Coulthard (1975) baseados na análise do discurso na sala de aula e aos estudos de Carletta et al. (1997) adaptados para o MapTask. Estas anotações incluem a estrutura do diálogo em movimentos, jogos e transacções. Cada movimento do diálogo consiste em uma ou mais palavras, silêncios ou ruídos e é etiquetado com o seu tipo de movimento. Esta codificação é feita para cada informante.

Com base nesta informação o discurso poderá ser segmentado em transacções por forma a facilitar a adição de anotações mais detalhadas (prosódia, focus, sintaxe). Para tal, foi utilizada aplicação informática AGTK TABLE TRANS (versão 1.2). Este programa permite visualizar todas as formas de onda, em tempo real, bem como operar sobre determinados segmentos seleccionados, ouvir, apagar, anotar, etc. Assim, em diversos níveis, procedeu-se à anotação do movimento, do jogo e da transacção – um esquema de anotação da estrutura do diálogo, segundo a teoria adoptada, baseada na Conversational Game Analysis Theory (Carletta et al. 1997), e partindo do seu princípio basilar de que “the participants only communicate when they need to perform some goal”.

Na metodologia de análise foram considerados desde aspectos prosódicos dos actos de fala indirectos (ordens usando perguntas, perguntas usando asserções), a

prosódia como focalização do discurso, as questões/informações usando enunciados sucintos e a prosódia do aparte.

No estudo dos movimentos (*dialogue moves*), foi considerado, de acordo com o modelo teórico adoptado, que um jogo de conversação é um jogo de enunciados: começa por uma enunciação, abrange todos os enunciados até que a finalidade do jogo seja cumprida ou não. São 12 os movimentos tipo de conversação considerados: 6 movimentos de iniciação (*initiating moves*): *Instruct, Explain, Check, Align, Query-yn, Query-w*; 5 movimentos de resposta (*response moves*): *Acknowledge, Reply-y, Reply-n, Reply-w, Clarify*; e um movimento de pré-iniciação (*Ready move*).

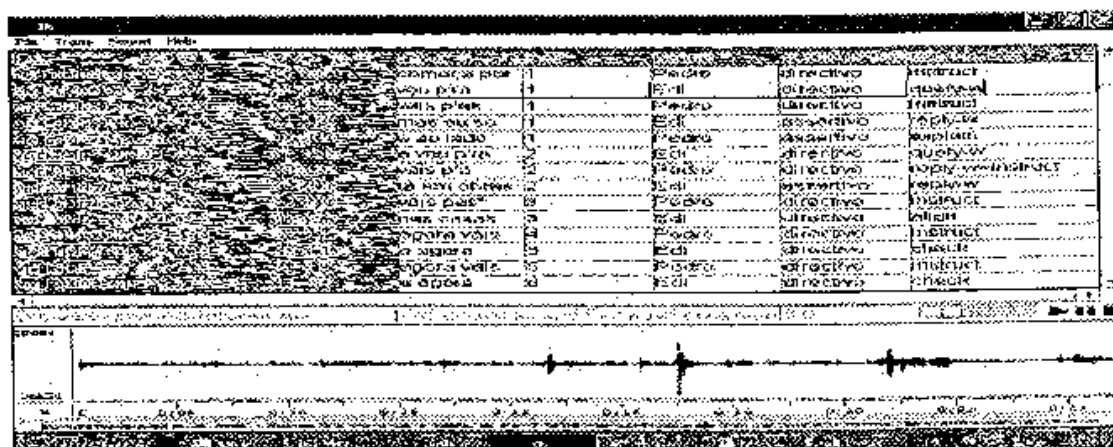


Figura 2: Exemplo de janela do AGTK TABLE TRANS

A aplicação usada permite-nos anotar ao mesmo tempo quem fala, as frases, as transacções, os actos de fala, etc. (ver figura 2). Uma vez que se trata de discurso espontâneo, aqui é possível trabalhar todo o diálogo sob as várias estruturas. Primeiro procedeu-se à segmentação do sinal, identificando o falante, depois procedeu-se à transcrição das frases usando o manual de anotação CSLU Labeling Guide, e ainda a informação relacionada com a análise do discurso: transacções, actos de fala, moves, etc.

6. Análise e apresentação de alguns resultados

Apresentamos, de seguida, alguns resultados obtidos, que permitem caracterizar sob o ponto de vista prosódico algumas situações discursivas particulares.

6.1. Frases complexas

Em primeiro lugar, foram analisadas frases complexas, interrompidas por uma breve intervenção do interlocutor, de que aqui apresentamos um exemplo (ver figura 3). Neste, e em outros analisados no corpus, verificou-se que, prosodicamente, o falante usou o F0 para indicar continuação. Depois da interrupção, retomou o enunciado na mesma altura em que o tinha suspenso, ficando, deste modo, claro que se tratava de uma continuação.

Um dos participantes, neste caso o *giver*, com movimentos de *Instruct*, vai estruturando o seu discurso com recurso a frases complexas, mas retomando a segunda oração – e <pausa> *podes ir p'o estádio* – no mesmo ponto em que tinha ficado no fim da primeira oração – *então contornas a rotunda*. Tratando-se de uma frase complexa, de duas orações, o *follower* coloca o seu movimento de *check* precisamente no final da primeira oração, informandô o interlocutor (*giver*) de que a instrução foi por ele recebida.

Os indicadores formais são muito úteis na segmentação da frase complexa, quando as orações que eles marcam não estão invertidas nem intercaladas dentro de outras orações. Neste caso – *então contornas a rotunda e podes ir p'o estádio* – não há inversão, a segunda oração só pode estar nesta posição, uma vez que se trata da tarefa do *MapTask*, o locutor dá a instrução – *e podes ir p'o estádio* – depois de ter dito para o alocutário contornar a rotunda, de outra forma o trajecto correcto não seria traçado. Portanto, a questão da criação de estados que se vão modificando gradativamente com o avanço de cada palavra, até chegar ao estado final, o recurso ao uso de marcadores formais de coordenação e subordinação pressupõe uma curva entoacional ascendente/descendente da oração aos seus componentes menores. Neste exemplo, o processamento continua a ser da esquerda para a direita: identificando o marcador de coordenação e que levará a frase até ao fim desta oração.

e p'a selva	9	Daniel
não porque tenho obras	9	André
então contornas a rotunda	9	Daniel
sim	9	André
e <paú> podes ir p'o estádio	9	Daniel
não	9	André
<bs> então regressa á escola <bs>	9	Daniel
posso	9	André

No entanto, não se verifica uma curva entoacional descendente, mas sim ascendente (fig. 3), uma vez que as orações estão separadas pelo vocábulo *sim*, produzido pelo interlocutor. Podemos ter ainda a noção de que esta interrupção do *follower* pode representar uma pausa – um indicador de separação.

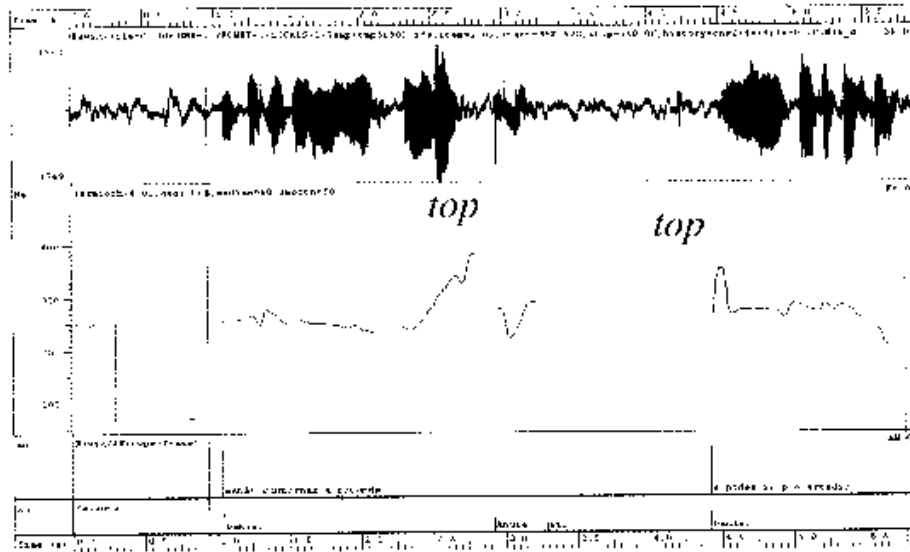


Figura 3: Exemplo do enunciado de uma frase complexa com interrupção pelo interlocutor

O que podemos verificar é que, mesmo com interrupção do follower, o giver mantém a estrutura do discurso como se a interrupção *sim* não existisse. Com a subida da curva entoacional verifica-se uma marcação de como o locutor ainda não acabou a sua tomada de palavra. O *sim* apenas aparece como indicador de que o trajecto no mapa está a ser seguido com êxito. É talvez a coordenação que está na origem desse reconhecimento. O *podes* designa o “tu” – interlocutor cuja existência foi postulada pela primeira frase *contornas a rotunda*. Trata-se aqui de um único acto de enunciação que corresponde a uma única intenção (explícita): contornar a rotunda e seguir para o estádio. Há uma relação entre o fenómeno de interrupção por parte do interlocutor e o da coordenação.

No exemplo acima transcrito, a frase complexa não apresenta independência entre as orações que a compõem, o que facilita a segmentação. A coordenativa e apesar de vincada, não tem o poder de anular o que foi dito antes dentro da frase, mesmo que a segunda oração apresente uma nova instrução.

As orações da frase têm autonomia sintáctica e morfológica, no entanto não podemos dizer o mesmo quanto à autonomia prosódica. O facto da primeira oração terminar em subida e a segunda oração começar também em subida, mostra uma dependência ao nível da prosódia.

6.2. O aparte

Uma segunda situação estudada foi o caso do aparte no discurso. Do nosso corpus conseguimos apenas recolher um exemplo (figura 4), mas que mostra o seu efeito na organização e estrutura do discurso oral.

Numa situação de diálogo como a do MapTask, “quase-espontâneo”, o aparte surge dentro de um contexto e contribui para a interpretação desse mesmo contexto por parte do alocutário.

Gumperz (1982) fala da existência de “pistas de contextualização”, sinais verbais ou não verbais que os interlocutores utilizam para fazer relacionar o que está a ser dito num dado momento e num dado lugar com o conhecimento adquirido por experiências anteriores. O aparte pode funcionar como uma dessas pistas que opera a nível prosódico. O ritmo com que é proferido mantém o envolvimento necessário e avalia o que se pretende significar. O aparte é um sinal de reformulação, dando lugar a uma explicação suplementar. É uma estratégia interaccional, um tipo de sequência para gerir os actos conversacionais.

No exemplo que se segue, o locutor, neste caso o Daniel, prosseguia a tarefa do MapTask com movimentos de *Instruct* e foi interrompido pelo alocutário, o André, com um aparte, revelador do conhecimento do contexto que iria especializar a interpretação e compreensão do que foi dito. Foi feita uma interpretação do enunciado pelo alocutário que necessitava de uma explicação rápida por parte do locutor, para que a tarefa continuasse.

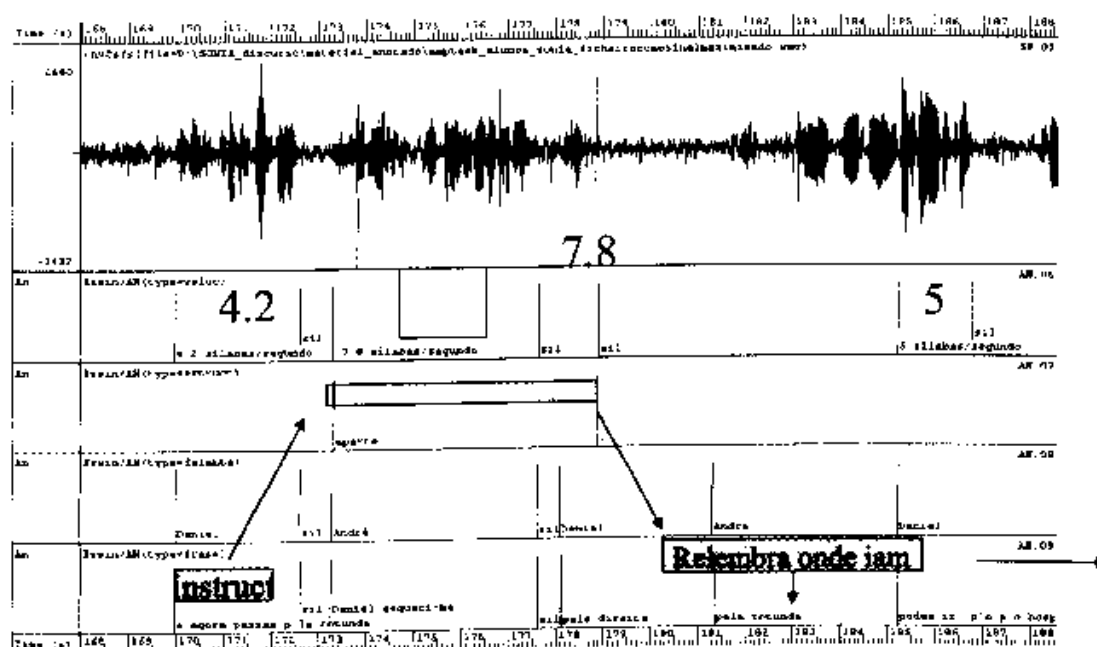


Figura 4: Exemplo de aparte

Giver: *agora passas p'la rotunda* [*Instruct*]

Follower: *Daniel eu esqueci-me de fazer uma pergunta contorna-se a rotunda pela esquerda ou pela direita* [*Query-W*] [*APARTE*]

Giver: *pela direita* [*Reply-W*]

Follower: *pela rotunda* [*Ack*]

No aparte *Daniel, esqueci-me de fazer uma pergunta contorna-se a rotunda pela esquerda ou pela direita*, verifica-se um maior número de sílabas por segundo, uma vez que a sua intenção era que o locutor respondesse também rapidamente à sua pergunta. Depois o alocutário relembra onde iam. O aparte vai, por isso, conferir ritmo ao enunciado. O aspecto prosódico do ritmo é, portanto, utilizado para distinguir o aparte do discurso em que este se insere, destacando-o prosodicamente. Uma vez que a

duração das sílabas é nitidamente menor no aparte, essa escolha evidencia o seu carácter marginal em relação ao discurso principal. Ou seja, ao falar com um débito acelerado no aparte, o falante está a dar transmitir a ideia de que está de facto a interromper a tarefa, mas que rapidamente a retomará.

6.3. Enunciados da mesma frase

O corpus permitiu, ainda, a análise da enunciação da mesma frase em pontos diversos da interacção, uma vez que uma mesma tarefa poderia ser solicitada ao longo do percurso.

Exemplificando, e partindo da comparação entre os dois enunciados de uma mesma frase (figura 5), podemos verificar que, tratando-se de uma frase do tipo declarativo, ainda que sob a forma indirecta de uma ordem, o sinal é semelhante, sem grandes alterações do F0, ou seja, revelam uma curva prosódica idêntica. A existência de contexto vai determinar o objectivo da repetição do enunciado da mesma frase em alturas distintas da tarefa.

A repetição da mesma frase em momentos distintos da interacção prende-se, portanto, com a prossecução da tarefa do mapa – chegar a um local.

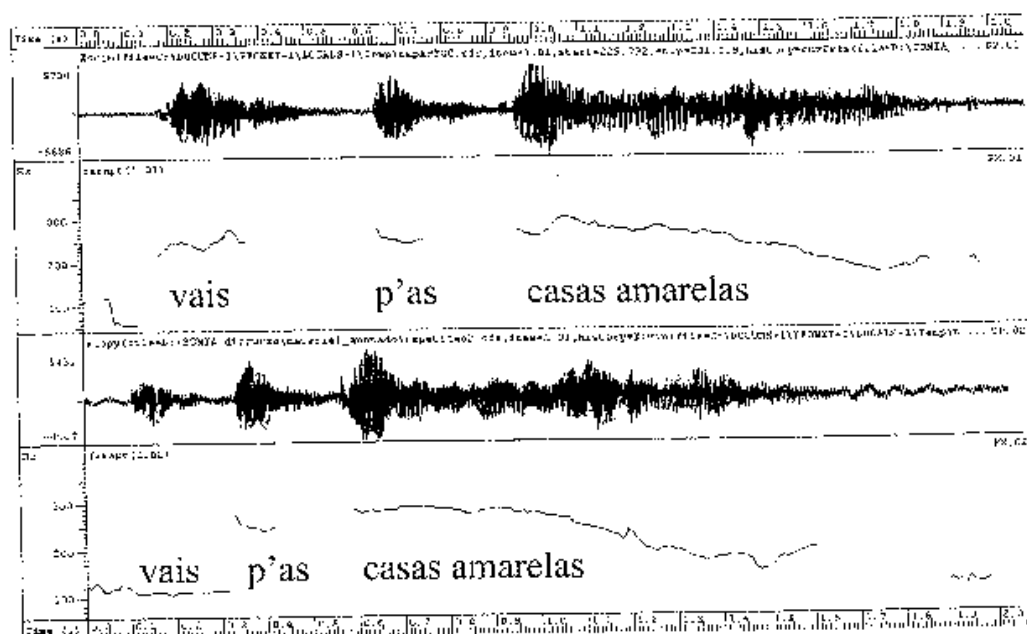


Figura 4: Dois enunciados da frase *vais p'as casas amarelas*

Outro tipo de repetição que pretendíamos estudar era a repetição imediatamente realizada após a elocução, para enfatizar, corrigir e reforçar o acto conversacional como marcador da estruturação discursiva. Pela reduzida dimensão do corpus não tivemos acesso a qualquer realização deste fenómeno.

6.4. Os enunciados sucintos (*Check*)

É na interacção dialógica que os falantes mais utilizam enunciados sucintos, de frases elípticas, já que o contexto situacional comum aos interlocutores, bem como o contexto linguístico precedente, permitem omitir uma parte da informação frásica, limitando-se a superfície textual à apresentação da informação nova. Também surgem enunciados sucintos nas expressões fáticas, de manifestação de continuidade discursiva, para confirmar ou infirmar uma informação dada, etc.

No contexto do MapTask, os interlocutores utilizam pistas de forma a ser possível manter o envolvimento necessário e avaliar o que se pretende significar. Desta forma, a existência de contexto vem facilitar a interpretação necessária, para que se faça sentido. Tais pistas podem operar ao nível da prosódia.

O movimento *Check* pode ser considerado mais uma estratégia pragmática de controlo da situação ou ainda um mecanismo conversacional pelo qual também se atribui a palavra ao interlocutor (cf. Mateus et al., 2003, pp. 477-479 sobre a questão das interrogativas tag, cuja forma consiste na repetição do verbo da frase declarativa que a precede, servindo como pedido de informação).

Nos enunciados que se seguem, os interlocutores reconhecem os objectivos ilocutórios que levam a um determinado tipo de comportamento linguístico. A contextualização do MapTask vai permitir entrada de informação, recuperar informação anterior ou até tirar partido da informação disponibilizada. Nesta perspectiva, é a contextualização que explica que o alocutário (follower) possa reconhecer a intenção indirecta de um aparente pedido formulado pelo locutor (giver). O follower estrutura a informação recebida e, como forma de se certificar da acção partilhada para a contextualização, utiliza enunciados sucintos, que constituem a base suficiente para a cooperação na interacção entre os dois participantes na tarefa do MapTask. Neste caso, a interacção constitui uma acção conjunta.

onde é que eu começo Daniel	
hmm podés começar no lago	
no Lago	←→ <i>Check - após um Instruct</i>
sim <pau> pas[sas]* passas pela escola	
sim	
agora <pau> passas pela rotunda	
sim	
e agora passas pela ca[sã]* passas por a ponte	
sim	
e vais p'a casa verde	
sim já estou lá	
<bs> agora p'a casa <pau> p'a casa branca <bs>	
casa branca	←→ <i>Check- após um Instruct</i>
e agora passas pela rotunda	
daniel eu esqueci-me de fazer uma pergunta cc br	

Figura 5 – Transcrição ortográfica (AGTK Table trans) da interacção dialógica incluindo enunciados sucintos de tipo *Check*, seguindo-se a enunciados de tipo *Instruct*

No exemplo ilustrado nas figuras 5 e 6 – *podes começar no lago/ no lago* – temos um *Check* seguido de um *Instruct*. Se analisarmos a curva de entoação (Fig. 6), verificamos uma configuração ascendente no *Check*. Podemos dizer que existe um reforço prosódico da focalização. O sintagma “no lago”, sendo um complemento de lugar, transmite uma informação fundamental no MapTask, devido à importância dos lugares marcados no mapa para a concretização da tarefa. Essa importância explica a utilização de um movimento ascendente de F0 associado ao *Check* por parte do follower, dando, desse modo, a entender ao giver que o lugar foi atingido e que pode prosseguir com a instrução seguinte.

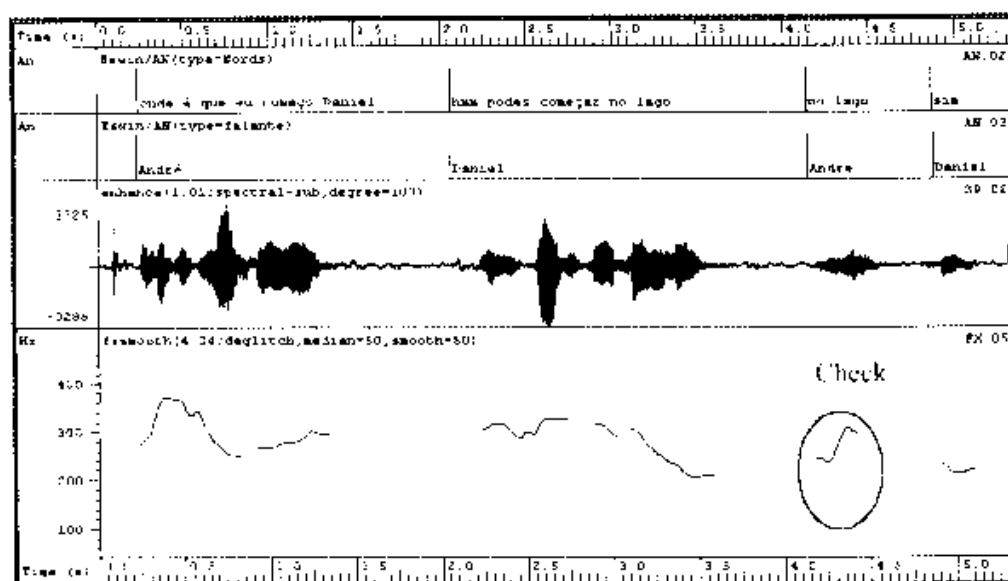


Figura 6 – contorno entoacional dos enunciados
 “onde é que eu começo Daniel / hmm podes começar no lago / no lago / sim”

Podemos dizer que os *Check* são instrumentos de coesão discursiva que asseguram a contextualização. O que se diz e a forma como algo é dito é, de algum modo, condicionado pelo que foi anteriormente dito e pela forma como foi dito. A simples observação do contexto linguístico em que os enunciados ocorrem mostra a importância do discurso que foi previamente produzido sobre os enunciados que se vêm a produzir.

Os *Check* são, pois, produções de sentido não convencionais que permanecem implícitas nos enunciados e são relevantes para a interpretação dos mesmos, sobretudo a nível prosódico.

7. Conclusão

Neste estudo, apresentámos alguns resultados decorrentes de uma pesquisa mais alargada (Amaral, 2005), para a qual foi constituído um corpus de interações dialógicas suscitadas por MapTask. Pelas análises efectuadas, de que aqui se

apresentaram alguns exemplos, pode-se concluir pela utilidade do corpus e da anotação já efectuada, e que viabilizou as primeiras explorações.

Podemos concluir da pertinência deste tipo de pesquisa para o estudo dos seguintes fenómenos: a utilização de F0 para indicar continuação de frase, mesmo com interrupções do interlocutor; a utilização do ritmo para estruturação do discurso (ex.: aparte); a repetição de enunciados em partes distintas do diálogo (*tasks*) com marcação prosódica; a utilização de enunciados sucintos para check.

De todos os exemplos analisados, constatámos, nos resultados obtidos, a funcionalidade das estratégias interaccionais nos diferentes registos observados, a acrescentar ao papel efectivo que estas desempenham no discurso oral.

Em situação de discurso, os actos de fala são portadores de significado que resulta da utilização de mecanismos que atribuem ao enunciado um significado pragmático ou comunicativo para a situação específica de interacção, neste caso o MapTask.

Apesar da idade dos interlocutores que fazem parte do corpus analisado, através dos resultados obtidos, podemos referir que houve cooperação mínima entre os participantes, contendo os princípios básicos que regulam o processo de interacção aqui estudado. O estabelecimento de toda a informação relevante processa-se a diversos níveis: ao nível da prosódia, nomeadamente pela organização temporal da fala e pela entoação adequada à situação; ao nível sintáctico-semântico, nomeadamente através da coesão entre enunciados, da presença de conectores de coordenação (as frases complexas) ou a não utilização desses conectores; ao nível interaccional, pelos tipos de cooperação entre objectivos comunicativos que regulam o dar a palavra e o tomar a palavra, a utilização de pausas, as estratégias pragmáticas de controlo da situação; ao nível da atenção, uma vez que são crianças e pré-adolescentes.

O trabalho desenvolvido permite-nos concluir sobre a importância do estudo do discurso oral espontâneo, no sentido de determinar marcas distintivas dos registos discursivos, produzidos em situação de diálogo orientado por participantes jovens e submetidos a objectivos específicos, de forma a encontrar regularidades na diversidade aparente dos factos prosódicos, nos quais integramos todos os processos relacionados com a estrutura do discurso oral espontâneo. É de referir que o próprio objecto de estudo – a produção de um corpus desta natureza – é um dos resultados mais relevantes deste trabalho.

8. Referências bibliográficas

- AMARAL, Sónia Delfina (2005) *A Prosódia no Discurso Espontâneo*. Dissertação de Mestrado apresentada à Universidade de Aveiro, no âmbito do Mestrado em Estudos Portugueses. (Entregue em Abril de 2005).
- ANDERSON, A. *et al.* (1991) The HCRC MapTask Corpus. *Language and Speech*, 34, pp. 351-366.
- AUSTIN (1986) *How to do things with words*. Oxford: Oxford University Press.
- CARLETTA, J. C. *et al.* (1997) The Reliability of a Dialogue Structure Coding Scheme. *Computational Linguistics*, 23(1), pp. 13-31.

- GUMPERZ, John J. (1982) *Discourse Strategies*. Cambridge: Cambridge: University Press.
- MATEUS, Maria Helena Mira *et al.* (2003) *Gramática da Língua Portuguesa*, 5ª ed. Lisboa: Caminho.
- PEREIRA, Isabel *et al.* (1992) *Estudos em Prosódia*. Lisboa, Edições Colibri.
- SEARLE, John R. (1979) *Expression and Meaning – Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.
- SEARLE, John R. (1969) *Speechs Acts*. Cambridge: Cambridge University Press.
- SINCLAIR, J. & Coulthard, M. (1975) *Towards an Analysis of Discourse*. Oxford: Oxford University Press.