

Classificação prosódica de marcadores discursivos

Vera Cabarrão^{1,2}, Helena Moniz^{1,2}, Jaime Ferreira¹, Fernando Batista^{1,3},
Isabel Trancoso^{1,4}, Ana Isabel Mata², Sérgio Curto¹

¹L2F, INESC-ID, ²FLUL/CLUL, ³ISCTE-IUL, ⁴IST – Universidade de Lisboa
{veracabarrao, jaimeferreira.90, scurto}@gmail.com, {Helena.Moniz, fernando.batista, isabel.trancoso}@inesc-id.pt, aim@letras.ulisboa.pt

Abstract

This work describes the discourse markers present in two corpora for European Portuguese, in different domains (university lectures and map-task dialogues). In this study, we also perform a multiclass automatic classification task based on prosodic features to verify in both corpora which words are discourse markers, which are disfluencies, and which are *sentence like-units* (SUs). Results show that the selection of discourse markers varies across domain and between speakers. As for the classification task, results show that the discourse markers are better classified in the lectures corpus (87%) than in the dialogue corpus (84%). However, cross-domain experiments evidenced that data trained with the dialogue corpus predicts better the events in the lecture corpus, since this domain displays more speakers and therefore complex patterns. In both corpora, markers are more easily classified as SUs than as disfluencies.

Keywords: *discourse markers, prosody, speech processing, multiclass automatic classification.*

Palavras-chave: marcadores discursivos, prosódia, processamento de fala, classificação automática multiclasse.

1. Introdução

O presente estudo visa contribuir para a descrição prosódica de marcadores discursivos em dois *corpora* de fala espontânea em Português Europeu (PE), em domínios diferentes, nomeadamente aulas universitárias e diálogos em formato *map-task*. A principal motivação deste trabalho é a realização de uma tarefa de classificação automática multiclasse para determinar, com base nas propriedades prosódicas das palavras, quais as que são marcadores discursivos, quais as que são disfluências e quais as que são constituintes similares a frases (do inglês *sentence like-units*, SUs). No domínio do processamento automático de fala, as marcas de pontuação, que delimitam SUs, as disfluências e os marcadores discursivos fazem parte de um conjunto de eventos designados no inglês *structural metadata events* (Liu *et al.*, 2006; Ostendorf *et al.*, 2008). Nesta linha de análise, pretende-se recuperar automaticamente a pontuação e as maiúsculas em fronteiras de frase, bem como a anotação e filtragem de disfluências e de marcadores.

A tarefa de classificação multiclasse permitir-nos-á, num futuro próximo, incluir os marcadores discursivos nas transcrições automáticas do reconhecer de fala (Neto *et al.*, 2008). A identificação de marcas de pontuação (Batista, 2011; Batista *et al.*, 2012; Moniz, 2013) e de disfluências (*e.g.*, pausas preenchidas lexicalizadas, como “aam” e/ou “mm”, apagamentos, substituições, entre outros¹) nas transcrições já permitiu uma melhoria significativa do *output* do sistema, o que resultou numa diminuição da taxa de erro de reconhecimento (Moniz *et al.*, 2014b). Com a recente disponibilização de uma grande quantidade de *corpora* de fala espontânea, foi

¹ Para uma descrição detalhada sobre as tipologias de disfluências, ver Moniz (2006, 2013).

possível analisar os eventos em falta, nomeadamente, os marcadores discursivos (Liu *et al.*, 2006; Ostendorf *et al.*, 2008), tópico do presente estudo.

Em PE, tal como acontece noutras línguas, os marcadores discursivos desempenham diferentes funções no discurso, sendo utilizados principalmente como estratégias paralinguísticas para manter o nexos discursivo entre enunciados (Schiffirin, 1987; Fraser, 1990; Schourup, 1999; Aijmer, 2013). No âmbito do processamento automático de fala, Goldwater *et al.* (2010) mostraram que estas estruturas, a par de outras palavras que tendem a ocorrer em posição inicial, são difíceis de classificar e apresentam elevados erros de reconhecimento quando processadas automaticamente. Assim, de forma a enriquecer as transcrições automáticas, é necessária uma caracterização linguística destas estruturas, bem como a identificação das suas propriedades prosódicas mais proeminentes.

Os marcadores discursivos também suscitam algumas questões problemáticas na área da tradução, quer humana quer automática. Estas estruturas são caracterizadas pela sua natureza idiomática, o que coloca dificuldades na procura de equivalentes exatos em diferentes línguas. Em Lopes *et al.* (2015), os autores comparam os marcadores discursivos de dois *corpora* espontâneos em PE (CORAL e LECTRA, utilizados no presente trabalho) com o seu equivalente em inglês, extraído do *corpus* Europarl, o *corpus* do Parlamento Europeu². Os resultados mostram que, dos cerca de setenta marcadores discursivos em PE, apenas dezoito têm tradução em inglês no Europarl. Verifica-se, assim, que a construção de um inventário destas palavras e expressões, a par da descrição das suas funções no discurso, é crucial para diversas tarefas.

Em PE, os marcadores discursivos correspondem a classes de palavras distintas, nomeadamente advérbios, conjunções, interjeições. O facto de um mesmo marcador poder ser associado a diferentes funções pragmáticas, sendo, como tal, multifuncional ou polissémico (Schiffirin, 1987; Fischer, 1998, 2000; Soares da Silva, 2006), representa mais um desafio na sua classificação. Adicionalmente, a maioria dos estudos desenvolvidos em PE baseiam-se em textos escritos e não contemplam ainda todos os marcadores discursivos presentes na fala espontânea. Note-se que aqui se entende por textos escritos textos de carácter mais formal e cuidado, não se contemplando, por exemplo, *tweets*. Estes, tal como defendido por Brogueira *et al.* (2014), apresentam características de comunicação oral, pelo uso de uma linguagem coloquial, com contrações de palavras e até reduções. No mesmo *corpus* utilizado por Brogueira *et al.* (2014), uma coleção de *tweets* geolocalizados com um total de 307 mil *tweets*, produzidos em Portugal por cerca de 11 mil utilizadores diferentes num período de oito dias, foram encontrados cerca de 45 mil marcadores discursivos em início de enunciado, sendo que a maioria é produzida como expressão fixa, utilizada como se fosse interjeição ou desbloqueador de conversa, sem conteúdo proposicional (e.g. “ó pá”, “ya”, “fogo”, “bom”, “pronto”, etc.).

Assim, este estudo visa contribuir para a descrição linguística de marcadores discursivos utilizados maioritariamente em diálogos espontâneos. Para tal, adopta-se uma abordagem baseada nos dados para identificação dos marcadores discursivos presentes nos dois *corpora* de fala analisados, seguindo-se uma classificação automática dos mesmos, com base nas suas propriedades acústico-prosódicas. Nesta tarefa, pretende-se diferenciar os marcadores discursivos das disfluências e das SUs (Shriberg, 1994; Liu *et al.*, 2006; Moniz, 2013).

2. Enquadramento teórico

Os marcadores discursivos são utilizados como estratégias de gestão dos diálogos e como mecanismo para estruturação de enunciados, com vista à manutenção das relações de coesão no discurso. Estes podem igualmente ser usados como estratégias de tomada de palavra e/ou como forma de manifestar cortesia, sendo associados a diferentes valores pragmáticos (e.g. Schiffirin, 1987; Fraser, 1990; Beeching, 2002). Vários autores demonstraram a importância destas estruturas em diferentes línguas, tendo em conta a sua alta frequência e multifuncionalidade, quer na fala como em textos escritos. Tem sido feito também um crescente esforço para definir o conjunto de palavras que podem pertencer à classe dos marcadores discursivos (Fraser, 1990 vs. Schiffirin, 1987), existindo uma variedade de definições para abranger esta classe heterogénea, nomeadamente marcadores discursivos (Schiffirin, 1987; Schourup, 1999), marcadores conjuntivos (Fraser, 1988; 1990; 1999), partículas pragmáticas (Beeching, 2002; 2014), marcadores fáticos (Fraser, 2009), marcadores pragmáticos

² No estudo de Lopes *et al.* (2015), foi utilizada a versão 7 do *corpus* Europarl, disponível em <http://www.statmt.org/europarl/>

(Redeker, 1990; Aijmer *et al.*, 2006; Denke, 2009), marcadores conversacionais (Urbano, 2003; Borreguero e López, 2010), entre outros.

De acordo com Fraser (1999), os marcadores discursivos são estruturas linguísticas que derivam das classes morfossintáticas conjunções, advérbios e frases preposicionais. Neste trabalho, interjeições não são consideradas como um marcador discursivo, na medida em que sinalizam uma mensagem adicional e não apenas uma relação entre segmentos. O autor afirma que os marcadores são uma classe pragmática, uma vez que contribuem para a interpretação dos enunciados, mas não do seu conteúdo proposicional. Baseando-se nesta distinção, Fraser (1999) define dois tipos de marcadores, os que se relacionam com as mensagens e os que se relacionam com os tópicos. O autor defende ainda que os marcadores discursivos sinalizam sempre uma relação entre a interpretação do segmento que introduzem, S2, e o segmento anterior, S1, que não precisa necessariamente de ser adjacente.

Numa perspectiva diferente, Schiffrin (1987) define os marcadores discursivos como dispositivos multifuncionais que delimitam unidades de fala, que são sintaticamente destacáveis e que apresentam variados contornos prosódicos. Estas estruturas correspondem a expressões linguísticas, como conjunções, interjeições, advérbios e frases lexicalizadas, bem como dispositivos não-verbais, como gestos ou características paralinguísticas. Contrariamente ao defendido por Fraser, Schiffrin (1987) mostra que os marcadores discursivos exibem relações não só locais mas também globais, ou seja, relações entre enunciados adjacentes e ao longo de todo o discurso, respetivamente. Em última análise, a autora evidencia o papel dos marcadores na estruturação de um discurso coerente, englobando tanto dimensões linguísticas como não-linguísticas da situação comunicativa, além do exclusivo uso de tais dispositivos como modo de indexar relações proposicionais num texto.

Aijmer *et al.* (2006) distinguem entre marcador discursivo e marcador pragmático, preferindo a última designação. A principal diferença baseia-se no facto de um marcador pragmático corresponder a uma palavra ou expressão que não contribui para o conteúdo proposicional de um enunciado, ao contrário do marcador discursivo, usado para descrever palavras que sinalizam relações de coerência. Embora os autores distingam estas estruturas, a fronteira entre elas não é muito clara, devido ao facto de as palavras poderem operar tanto a nível proposicional como não-proposicional.

O facto de os marcadores discursivos serem multifuncionais (Schiffrin, 1987, 2001) ou polissémicos (Fischer, 1998, 2000), na medida em que o mesmo item pode ter diferentes funções ou interpretações semânticas no discurso, faz com que seja mais difícil a sua desambiguação e, por conseguinte, a sua classificação. De acordo com Beeching (2014), os marcadores pragmáticos têm funções distintas, dependendo da natureza da interação. Na fala espontânea, permitem hesitações, retoma de enunciados, reparos e repetições, podendo ocorrer também como estratégias para tomar a palavra. Além disso, numa perspetiva social, os marcadores pragmáticos podem ser sociolinguisticamente marcados, são frequentemente associados a expressões de naturalidade, simpatia e cordialidade e também podem ser um sinal de educação.

Como brevemente descrito, não há consenso sobre a terminologia utilizada, quer para identificação ou classificação de marcadores discursivos. O principal critério utilizado no presente trabalho é empregar a designação mais comum, marcador discursivo, na sequência dos trabalhos de Schiffrin (1987, 2001). Uma vez que o objetivo principal deste estudo é identificar estas estruturas em diferentes *corpora*, enfatizando o seu papel na fala e não em textos escritos, esta análise segue uma orientação prosódica, com ênfase nos marcadores que ocorrem em posição inicial no enunciado³. Aqui, tal como demonstrado no estudo de Goldwater *et al.* (2010), são verificadas as maiores taxas de erros de reconhecimento automático de palavras, sendo este também o local de maior ocorrência de marcadores discursivos, bem como de disfluências (Moniz, 2013).

Sabe-se que os marcadores discursivos e as disfluências partilham as mesmas propriedades, como a alta frequência na fala espontânea (Jucker e Smith, 1998), a tendência para ocorrer em posição inicial no enunciado (Heeman e Allen, 1999; Goldwater *et al.*, 2010), sendo seguidos, principalmente, por uma pausa prosódica ou silenciosa ou por uma ruptura melódica (Schiffrin, 1987; Moniz, 2013), bem como, no caso das pausas preenchidas, pela ausência de conteúdo proposicional (Aijmer, 2013; Moniz, 2013).

De acordo com Popescu-Bellis e Zufferey (2011), existem duas perspetivas principais no processamento automático de marcadores e de disfluências: a remoção destas estruturas das transcrições automáticas e a sua inclusão para enriquecimento das transcrições com informação paralinguística. Lease e Johnson (2006)

³ Neste estudo, entende-se por enunciado um constituinte maior ou menor, ou seja, com índices de fronteira 3 ou 4, níveis de ruptura temporal e/ou melódica entre palavras, geralmente silêncio e/ou diferenciação de contornos de f_0 . Para uma descrição detalhada sobre tipos de fraseamento em PE, ver Viana *et al.* (2007).

comparam os marcadores discursivos a disfluências, nomeadamente a pausas preenchidas lexicalizadas, e procedem à sua extração das transcrições automáticas. Numa abordagem diferente, por exemplo, em Hirschberg e Litman (1993) e Samuel (1999), os autores afirmam que os marcadores discursivos podem ser usados para inferir a estrutura do discurso e, como tal, devem ser considerados no processamento automático.

Quanto à frequência destes itens, Jucker e Smith (1998), num *corpus* de diálogos para estabelecer o *common ground* entre falantes, encontraram um total de 31,8% de marcadores discursivos (marcadores de receção e marcadores de apresentação) por unidade de tópico (5 minutos). Os autores mostram que a frequência de ambos os tipos de marcadores varia de acordo com o grau de familiaridade entre os falantes (amigos ou estranhos), embora o número total de marcadores discursivos seja semelhante em ambos os pares (32,8% vs. 30,7%). Quanto às disfluências, a sua frequência varia em função de o interlocutor ser um sistema automático ou um humano. De acordo com resultados obtidos por Shriberg (2001), as disfluências apresentam um intervalo de 5% a 10% em diálogos humano-humano. Heeman e Allen (1999) defendem que as tarefas de segmentação de enunciados e a reparação da fluência juntamente com a identificação de marcadores discursivos são cruciais para estudar as intenções do falante em diálogos.

Estudos preliminares em PE (Batista, 2011; Moniz, 2013; Moniz et al. 2014a) mostraram que com base num conjunto de pistas puramente prosódicas é possível prever e sistematizar padrões regulares na identificação das distintas marcas de pontuação e de reposição de fluência. Para o efeito, as pistas que mais contribuem para a predição da reposição de fluência a seguir a uma sequência disfluente contemplam: i) duas palavras contíguas idênticas; ii) subida dos níveis de f_0 e de energia na palavra que inicia uma reposição de fluência e um contorno estacionário de f_0 na palavra anterior; e iii) grau de confiança da palavra que inicia a reposição superior ao da disfluência propriamente dita.

No que diz respeito às pistas associadas à predição de pontos finais, estas integram: i) contorno descendente na palavra antes de um ponto final; ii) nível estacionário de energia na mesma palavra; iii) duração relativa entre essa palavra e a seguinte; e iv) grau superior de confiança em relação à palavra seguinte. Este conjunto de pistas é ilustrativo do comportamento de uma declarativa neutra no PE. Os pontos de interrogação, por sua vez, são caracterizados por dois padrões diferenciados: i) contorno de f_0 ascendente na palavra antes de um ponto de interrogação e declive de energia ascendente nessa e na palavra seguinte; ii) contorno de f_0 estacionário na palavra antes de um ponto de interrogação e declive de energia descendente nessa mesma palavra. As vírgulas são o evento que menos depende de uma caracterização prosódica, sendo sobretudo morfossintática.

Para se prever e processar os marcadores discursivos de forma automática para uma dada língua, é necessário estabelecer, em primeiro lugar, quais as palavras que podem ser classificadas como marcadores e quais as suas características acústico-prosódicas. Em PE, os estudos sobre marcadores discursivos são escassos, não sendo ainda conhecidas a sua distribuição e frequência relativa em dados de fala, bem como o conjunto das suas características linguísticas mais informativas. A maioria dos estudos foca-se apenas no comportamento de um marcador discursivo específico e exclusivamente em textos escritos de carácter mais formal.

Coutinho (2009), adaptando as classificações de Fraser (1999) e Adam (2008), descreve os marcadores discursivos de acordo com o contexto e a especificidade do texto, a fim de distinguir os seus diferentes usos discursivos. Autores como Lopes (1997), Freitas e Ramilo (2003) e Soares da Silva (2006) apresentaram estudos referentes a apenas um marcador discursivo, nomeadamente “então”, “portanto” e “pronto”, respetivamente, e as suas diferentes funções de acordo com o contexto. O estudo de Pimentel (2012) analisa um conjunto de marcadores discursivos numa perspetiva de aquisição de segunda língua e o trabalho de Lopes (2014) estuda o comportamento diacrónico do marcador “aliás”.

Apesar do crescente esforço para descrever e classificar marcadores discursivos em PE, estas estruturas são ainda pouco estudadas na nossa língua, especialmente no que diz respeito às suas propriedades idiossincráticas na fala espontânea. Pelo que nos é dado a conhecer, este trabalho é o primeiro a analisar marcadores discursivos em diferentes *corpora* de fala espontânea. O principal objetivo deste estudo é identificar e classificar essas estruturas, tendo em conta as suas funções e contribuições para o conteúdo paralinguístico do discurso. Esta análise levará em conta não só os padrões de distribuição dos marcadores nos dados, mas também as suas características prosódicas. Com estas últimas, será possível caracterizar estas estruturas, prevê-las automaticamente e estabelecer uma comparação entre estes marcadores e outro tipo de *structural metadata events*, nomeadamente as disfluências.

3. Metodologia

3.1. *Corpora*

Este estudo utiliza dois *corpora* de fala espontânea em diferentes domínios, nomeadamente aulas universitárias (*corpus* LECTRA) e diálogos em formato *map-task* (*corpus* CORAL). Ambos estão disponíveis através da plataforma ELRA. O *corpus* de aulas universitárias foi recolhido no âmbito do projeto nacional LECTRA (Trancoso *et al.*, 2008), visando a produção de conteúdos multimédia para aplicações de *e-learning* e para alunos com deficiência auditiva. O *corpus* LECTRA (ISLRN - 298-379-572-530-5) tem um total de sete unidades curriculares, seis gravadas na presença de estudantes e um gravado apenas com o professor para uma audiência *online*. Todos os professores (seis homens e uma mulher) são falantes nativos de português. O LECTRA tem um total de 75 horas de fala, das quais 32 horas foram ortograficamente transcritas, num total de 155 mil palavras. O processo de transcrição foi realizado por 3 anotadores, com o mesmo percurso académico na área da linguística. A anotação das disfluências baseou-se em Moniz (2006) e em Trancoso *et al.* (2008) e a das marcas de pontuação, em Duarte (2000). A concordância entre anotadores foi avaliada (Moniz, 2013), tendo sido obtida uma concordância quase perfeita entre um anotador (A1) e os restantes (A2 e A3, com uma exatidão de 0.82 e 0.79, respetivamente) e uma concordância substancial entre A2 e A3 (com uma exatidão de 0,69).

O *corpus* CORAL (ISLRN 499-311-025-331-2) foi recolhido no âmbito do projeto nacional com o mesmo nome (Trancoso *et al.*, 1998), compreendendo 64 diálogos em formato *map-task* entre 32 falantes. Os diálogos ocorrem entre dois falantes com papéis diferentes, um é o dador de informação e o outro é o seguidor. O dador tem um mapa com uma rota traçada e alguns marcos e a sua tarefa é fornecer informações e orientações para o seguidor reconstruir o mesmo caminho no seu mapa incompleto. Há também diversas inconsistências entre os nomes e os locais dos pontos de referência para suscitar o diálogo. O CORAL é equilibrado em termos de género e de papel desempenhado pelos falantes. O *corpus* tem 7 horas de fala ortograficamente transcritas, num total de 61 mil palavras.

3.2. Identificação dos marcadores discursivos

Este estudo segue uma abordagem orientada pelos dados, tendo sido tomados vários passos para identificar os marcadores discursivos. Este foi um processo trabalhoso, que se baseou em primeiro lugar na audição de ambos os *corpora*. Durante a prospeção dos dados, foram identificadas cerca de sessenta itens que partilham características de marcadores discursivos (ver Quadro 1), uma vez que (i) são muito frequentes em fala espontânea (Heeman e Allen, 1999) e raras em textos escritos; (ii) ocorrem, maioritariamente, em início de enunciado (Goldwater *et al.*, 2010); e (iii) podem ser extraídas do discurso sem comprometer o seu conteúdo proposicional (*e.g.*, “pronto” e “bem”) (Aijmer *et al.*, 2006). Para esta tarefa, foi também levado em consideração o facto de estas estruturas já estarem marcadas na transcrição ortográfica disponível para ambos os *corpora*, sendo constantemente seguidas por vírgula ou ponto final. Neste processo de seleção, incluíram-se também casos de coocorrência do mesmo marcador, como “então” e “mas então”, para se dar conta do maior número possível de marcadores. Esta abordagem orientada pelos dados permitiu estudar marcadores discursivos que, apesar de serem muito frequentes na fala espontânea, ainda não estão descritos na gramática do Português.

Marcadores discursivos		
<i>(E) já agora</i>	<i>Ora vejamos</i>	<i>A seguir</i>
<i>Atenção</i>	<i>Pá</i>	<i>Agora</i>
<i>Bem</i>	<i>Pois é</i>	<i>(Mas) eu agora</i>
<i>Bom</i>	<i>Pois bem</i>	<i>E agora</i>
<i>Desculpem lá</i>	<i>Pois então</i>	<i>Mas agora</i>
<i>Eh pá; epá</i>	<i>Pois muito bem</i>	<i>Até agora</i>
<i>Então é assim</i>	<i>Portantos</i>	<i>Enfim</i>
<i>É assim</i>	<i>Pronto</i>	<i>Mas enfim</i>
<i>Mas então</i>	<i>Quer-se dizer; quer dizer</i>	<i>Então</i>
<i>Então (isto) agora</i>	<i>Tomem lá atenção</i>	<i>Entretanto</i>
<i>Então a seguir</i>	<i>Tudo bem</i>	<i>Ok</i>
<i>Muito bem</i>	<i>Vamos lá</i>	<i>Pois</i>

<i>Olha; olhe</i>	<i>Vamos lá (a) começar</i>	<i>Portanto</i>
<i>Olha lá</i>	<i>Vamos lá estar então</i>	<i>E portanto</i>
<i>Ora</i>	<i>E pronto</i>	<i>E depois</i>
<i>Ora bem</i>	<i>Mas pronto</i>	<i>De facto</i>
<i>Ora bom</i>	<i>Prontos</i>	<i>Por acaso</i>
<i>Ora então agora</i>	<i>Com atenção</i>	<i>No entanto</i>
<i>Ouve; ouça</i>	<i>Vamos lá ver</i>	<i>Se calhar</i>
<i>Ora isto</i>	<i>Ó; oh</i>	<i>Ou seja</i>

Quadro 1: Marcadores discursivos no *corpus* de aulas universitárias (LECTRA) e no *corpus* de diálogos (CORAL).

Nesta lista de marcadores discursivos, optou-se por incluir os marcadores “e” e “mas” apenas quando em coocorrência com outros marcadores, como unidades lexicais multipalavra. Na prospeção dos dados, verificou-se que quer “e” ou “mas” são usados de duas formas distintas: (i) como conectores com conteúdo proposicional e (ii) como marcadores discursivos com informações interpessoais, no sentido de Beeching⁴ (2014). Na maioria dos casos, especialmente no *corpus* de aulas universitárias, estas conjunções veiculam conteúdo proposicional. Uma vez que o que se pretende aqui é uma caracterização geral dos marcadores discursivos e não a exploração das funções de cada marcador individualmente, optou-se pela não inclusão destas estruturas desde que não coocorram com outros marcadores discursivos. Uma análise mais detalhada sobre ambos os itens e contextos adjacentes é necessária, mas não será efetuada no âmbito do presente estudo. Por outro lado, optou-se pela inclusão de dois conectores frequentes, “portanto” e “agora”, uma vez que a audição dos *corpora* indicou que ambos, quando ocorrem em posição inicial no enunciado, são principalmente marcadores discursivos.

4. Distribuição dos marcadores discursivos

Os marcadores discursivos, nas aulas universitárias e nos diálogos, correspondem a 3% do número total de palavras em cada *corpus* (o *corpus* LECTRA tem um total de 5103 vs. 1719 no CORAL). Num *corpus* de diálogos em formato *map-task* entre humanos em Inglês, com 58298 palavras, Heeman e Allen (1999) mostram que 14% das palavras são marcadores discursivos, evidenciando que estas estruturas são muito frequentes na fala espontânea.

Ainda que os marcadores ocorram com maior frequência em início de enunciado, também se verifica que o mesmo marcador pode ocorrer tanto em posição inicial como em qualquer outra posição no enunciado, mesmo em posição final devido à sobreposição de voz e estratégias de tomada de palavra. Deste modo, os padrões de distribuição aqui descritos têm em consideração todos os marcadores discursivos encontrados, independentemente da sua posição no enunciado.

Os marcadores mais frequentes (Quadro 2) nas aulas universitárias são “portanto” (34% das ocorrências); “agora” e “ok” (ambos com 10%), e “então” (8%). Os restantes marcadores são bastante residuais no *corpus*. Quanto aos diálogos, os marcadores discursivos mais frequentes são “então” (21%), “portanto” (19%), “pronto” e “ok” (ambos com 15%). Nas aulas universitárias, os professores tendem a usar estas estruturas para gerir o tempo, enquanto planificam as unidades subsequentes e o seu nexos de coesão. Nos diálogos, os falantes têm interações mais curtas e num ritmo mais rápido, sendo estas estruturadas com base em perguntas e respostas para fins de esclarecimento e procura de informação.

Marcadores discursivos	LECTRA	CORAL
<i>(E) Portanto</i>	34,0%	19,0%
<i>(E) Agora</i>	10,0%	5,3%
<i>Então</i>	8,0%	21,0%
<i>(E) Pronto</i>	3,0%	15,0%
<i>Ok</i>	10,0%	15,0%
<i>E depois</i>	4,0%	7,0%

⁴ A autora afirma que os marcadores discursivos podem ser socialmente marcados, havendo marcadores mais usados em determinadas regiões, por falantes de uma certa faixa etária e grupo social.

<i>Ou seja</i>	3,5%	0,4%
<i>Pois</i>	0,4%	3,3%
<i>(Mas) É assim</i>	1,5%	2,0%
<i>Bom</i>	1,2%	0,1%
<i>Olha</i>	1,0%	0,3%
<i>Ora bem</i>	1,2%	0,4%
<i>Prontos</i>	0,0%	1,5%

Quadro 2: Marcadores discursivos mais frequentes em ambos os *corpora* (LECTRA e CORAL).

A distribuição global dos marcadores discursivos por falante revela que estes tendem a usar marcadores semelhantes no mesmo domínio, mas com diferentes combinações, como “ora”, “ora bem” e “ora bom” ou até mesmo “portanto”, “e portanto”, “portanto é assim”. Os resultados também mostram que a sua distribuição varia substancialmente entre falantes, por exemplo, o marcador discursivo mais comum no *corpus* das aulas universitárias não é produzido por todos os falantes ou nas mesmas proporções: “portanto” é produzido 40% das vezes apenas por um falante (S7) e 29% por outro (S3). Também no *corpus* dos diálogos, há variação entre falantes, ainda que menos expressiva. “Então”, o marcador mais comum, é produzido quase 10% das vezes por dois falantes (S7 e S9), uma percentagem que cai para menos de 8% para os restantes falantes. Esta diferença repete-se em quase todos os outros marcadores. No entanto, o facto de o marcador “ok” pronunciado como [O'kap6] ser produzido apenas por um falante (S17) pode sugerir, de forma semelhante ao “portanto” na produção do falante S7 no *corpus* LECTRA, que a seleção de marcadores discursivos é influenciada pelas escolhas pessoais de cada falante.

Também se verifica que existem marcadores que só ocorrem num *corpus* mas não no outro, como “prontos”. Neste caso particular, o uso do jargão de “pronto” mostra que há um ambiente menos formal entre os falantes nos diálogos, uma vez que é utilizado um discurso mais coloquial, que se opõe ao discurso mais formal associado às aulas universitárias.

Em suma, os marcadores analisados apresentam variação entre os dois *corpora*, bem como dependem do falante, não apenas em domínios diferentes mas também no mesmo *corpus*. Nos dados analisados, não foram encontradas evidências de relação entre o género do falante e a seleção de determinados marcadores discursivos.

Tendo em conta o trabalho de Goldwater *et al.* (2010), em que os autores mostram que as palavras em início de enunciado ou como marcadores discursivos apresentam elevadas taxas de erro quando reconhecidas automaticamente, optou-se por descrever com mais pormenor os marcadores que ocorrem nessa posição, bem como os que ocorrem como enunciados isolados.

Os marcadores discursivos nestas posições (inicial e isolado) totalizam, nas aulas universitárias, 36% dos marcadores encontrados e, nos diálogos, cerca de 67%. Estes resultados já eram expectáveis considerando a natureza dos dados, ou seja, sendo o CORAL um *corpus* caracterizado por declarações curtas e interações sequenciadas entre os falantes. Todos estes marcadores foram classificados como discursivos, devido ao facto de poderem ser eliminados ou substituídos, por exemplo, por uma pausa preenchida, sem adicionar conteúdo proposicional.

5. Classificação automática multiclasse

5.1. Configuração das experiências

Após o processo de identificação de marcadores discursivos, foram analisadas as características acústico-prosódicas dos marcadores que ocorrem em início de enunciado ou em posição isolada, de forma a compreender quais as propriedades que estes partilham com as disfluências. Nesta tarefa, foram agrupados os marcadores nestas posições, uma vez que estes são os locais de maior ocorrência de marcadores e de disfluências, bem como onde ocorrem também os maiores erros de reconhecimento (ver secções 2 e 4). Adicionalmente, tratando-se esta de um tarefa automática, importa dar conta do maior número de casos possíveis, para um melhor desempenho do classificador. O principal objetivo é realizar uma tarefa de classificação automática multiclasse com base em características acústico-prosódicas, para verificar quais as palavras em ambos os *corpora* que são classificadas como marcadores discursivos, as que são disfluências, e as que não são nem marcadores nem disfluências (SUs).

Este processo permitirá, num futuro próximo, incluir marcadores discursivos nas transcrições automáticas do reconhecedor de fala.

Para a tarefa de classificação automática, foram extraídas 6300 características acústico-prosódicas dos marcadores discursivos em ambos os *corpora*, através da ferramenta openSMILE (Eyben et al., 2010). O mesmo conjunto de características já tinha sido usado anteriormente em estudos sobre disfluências e marcas de pontuação em PE (e.g., Moniz, 2013, Batista, 2011). Nestes, a tarefa de classificação de disfluências e de marcas de pontuação, para delimitação de constituintes similares a frases (SUs), teve como base o seguinte conjunto de pistas prosódicas: presença e duração de pausas nas fronteiras dos eventos, declinação global de f_0 , reinicializações de f_0 e de energia, alongamento final dos segmentos, ocorrência de pausas e ocorrência de fragmentos (Moniz et al., 2014b).

As experiências foram realizadas com a ferramenta WEKA (Hall et al., 2009), tendo sido utilizados diversos métodos de aprendizagem estatística, como Árvores de Decisão, Naïve Bayes e Máquinas de Suporte Vetorial (*Linear Regression Support Vector Machines* - SVM), sendo este último o que apresentou os melhores resultados.

A fim de estabelecer uma base de comparação para outras experiências, também foi aplicado o método ZeroR do WEKA, que reflete exclusivamente a seleção da classe mais frequente. Os resultados foram avaliados usando métricas de desempenho padrão (Makhoul et al., 1999), nomeadamente precisão (fração de instâncias recuperadas que estão corretas), cobertura (fração de instâncias relevantes que são recuperadas), medida-F (média ponderada entre precisão e cobertura) e exatidão (grau de aproximação da classificação ao seu valor correto). Algumas experiências também foram medidas em termos de estatística Kappa, uma medida de concordância entre as classificações obtidas e as verdadeiras classes. Um valor próximo de zero indica que os resultados podem ser alcançados por acaso, enquanto um valor próximo de 1,0 significa que o modelo é adequado ao problema em causa.

Para a configuração das experiências, quanto à seleção do parâmetro C (complexidade), utilizou-se o valor de 0,01, uma vez que, em trabalhos anteriores sobre disfluências, este provou ser o melhor valor para lidar com a quantidade de características extraídas a partir do openSMILE. Procedeu-se também a uma validação cruzada das amostras, feita com base numa amostragem não supervisionada considerando 5 subconjuntos. Quanto ao tipo de dados, realizaram-se experiências com amostras equilibradas e não-equilibradas. Em relação ao número total de instâncias, portanto, às amostras não-equilibradas, no CORAL, existem 6381 SUs, 1834 disfluências e 723 marcadores discursivos. Quanto ao LECTRA, há um total de 15068 SUs, 6066 disfluências e 1286 marcadores discursivos. Como a classe marcadores tem muito menos instâncias, o número de SUs e disfluências foi selecionado aleatoriamente para um total de 2000 no CORAL e 4000 no LECTRA, para que as experiências pudessem ser feitas também com uma amostra mais equilibrada.

5.2. Parâmetros prosódicos

Para a caracterização prosódica dos marcadores discursivos foram utilizadas as pistas acústico-prosódicas extraídas com o openSMILE. Estas compreendem 2 conjuntos: i) pistas de f_0 , como energia e parâmetros de voz (*jitter* e *shimmer*) com valores absolutos, sendo designadas por Eyben et al. (2010) de Descritores de Baixo Nível (Quadro 3), e (ii) Funcionais extraídos a partir dos valores absolutos, ou seja, cálculos aplicados sobre esses valores (estatísticas, regressão polinomial, transformadas – Quadro 4).

Feature Group	Description
Waveform	Zero-Crossings, Extremes, DC
Signal energy	Root Mean-Square & logarithmic
Loudness	Intensity & approx. loudness
FFT spectrum	Phase, magnitude (lin, dB, dBA)
ACF, Cepstrum	Autocorrelation and Cepstrum
Mel/Bark spectr.	Bands 0- N_{mel}
Semitone spectr.	FFT based and filter based
Cepstral	Cepstral features, e.g. MFCC, PLP-CC
Pitch	F_0 via ACF and SHS methods Probability of Voicing
Voice Quality	HNR, Jitter, Shimmer
LPC	LPC coeff., reflect. coeff., residual Line spectral pairs (LSP)
Auditory	Auditory spectra and PLP coeff.
Formants	Centre frequencies and bandwidths
Spectral	Energy in N user-defined bands, multiple roll-off points, centroid, entropy, flux, and rel. pos. of max./min.
Tonal	CHROMA, CENS, CHROMA- based features

Quadro 3: Descritores de Baixo Nível do openSMILE (Eyben et al., 2010: 1461).

Category	Description
Extremes	Extreme values, positions, and ranges
Means	Arithmetic, quadratic, geometric
Moments	Std. dev., variance, kurtosis, skewness
Percentiles	Percentiles and percentile ranges
Regression	Linear and quad. approximation coefficients, regression err., and centroid
Peaks	Number of peaks, mean peak distance, mean peak amplitude
Segments	Number of segments based on delta thresholding, mean segment length
Sample values	Values of the contour at configurable relative positions
Times/durations	Up- and down-level times, rise/fall times, duration
Onsets	Number of onsets, relative position of first/last on-/offset
DCT	Coefficients of the Discrete Cosine Transformation (DCT)
Zero-Crossings	Zero-crossing rate, Mean-crossing rate

Quadro 4: Funcionais disponíveis no openSMILE (Eyben et al., 2010: 1461).

A importância das pistas prosódicas dos marcadores discursivos, disfluências e SUs foi aferida através de Regressão Logística, de modo a hierarquizar quais as mais proeminentes em cada classe (ver secção 6.2).

6. Resultados

6.1. Predição automática de classes

A tarefa de classificação automática provou ser mais robusta no *corpus* de aulas universitárias do que nos diálogos, tanto com amostras equilibradas como não-equilibradas (Quadro 5). Estes resultados podem dever-se ao facto de o *corpus* LECTRA ter um menor número de falantes do que o CORAL (7 vs. 20 falantes), o que corresponde também a uma menor variação.

Os resultados mostram que a inclusão das pistas acústico-prosódicas na classificação automática produzem melhorias significativas relativamente à base de comparação: 20% nas aulas universitárias e 13% nos diálogos, na amostra não-equilibrada, e 16% e 8%, respectivamente, na amostra equilibrada. Observa-se também que a classificação tem uma maior exatidão na amostra não-equilibrada em ambos os *corpora*, ainda que o valor de kappa seja ligeiramente inferior nas aulas universitárias e consideravelmente mais baixo nos diálogos.

Como se pode observar na matriz de confusão (Quadro 6), nos diálogos, a classe SUs é a mais bem predita, seguida das disfluências e, por fim, dos marcadores. Estes apresentam os valores mais baixos de cobertura (62%), mas uma precisão (77%) mais elevada do que as disfluências (77%). Estes resultados mostram que, apesar de ainda haver muitas palavras que não são identificadas como marcadores discursivos, as que são têm uma classificação correta. Nas aulas universitárias (Quadro 7), o número de marcadores discursivos corretamente classificados é maior, sendo as disfluências a classe que apresenta maior ambiguidade com as SUs. Os marcadores apresentam uma precisão de 75% e uma cobertura de 63%, o que mostra que estes são difíceis de identificar e de classificar.

Corpora	Amostra não-equilibrada			Amostra equilibrada	
	Kappa	Exatidão		Kappa	Exatidão
		ZeroR	SVM		
Diálogos	0,599	71,4%	83,9%	0,659	78,9%
Aulas	0,717	67,2%	87,1%	0,725	83,3%

Quadro 5: Classificação automática com SVM para amostras equilibradas e não-equilibradas.

Amostra não-equilibrada – Diálogos			
Classificada como	SUs	Disf	MarcDisc
SUs	6032	270	79
Disf	761	1020	53
MarcDisc	240	37	446

Quadro 6: Matriz de confusão para o *corpus* de diálogos (CORAL)

Amostra não-equilibrada – Aulas			
Classificada como	SUs	Disf	MarcDisc
SUs	14216	704	148
Disf	1422	4521	123
MarcDisc	272	210	804

Quadro 7: Matriz de confusão para o *corpus* de aulas universitárias (LECTRA)

Para testar a robustez do sistema de classificação entre domínios, realizou-se uma outra experiência com dados de um *corpus* para treino e do outro para teste e vice-versa. Assim, para testar o *corpus* CORAL, foi utilizado o conjunto de treino não-equilibrado do LECTRA (para dar conta de um maior número de casos) e para o *corpus* LECTRA, os modelos foram treinados com o CORAL, também não-equilibrado.

Os resultados desta experiência foram semelhantes aos obtidos quando o conjunto de treino e de teste foram selecionados aleatoriamente a partir do mesmo *corpus*. Apesar de a cobertura ser um pouco mais baixa (ver Quadro 8), os melhores resultados são ainda encontrados para o *corpus* LECTRA não-equilibrado (80%).

Tendo em conta a complexidade das estruturas incluídas na classe marcadores discursivos, esta é, como expectável, a mais difícil de prever. Como se verifica nas matrizes de confusão (Quadros 9 e 10), a classe dos marcadores discursivos é a que apresenta mais problemas de classificação em ambos os *corpora*. A medida-F dos marcadores é igualmente baixa no CORAL (27%) e no LECTRA (18%). No entanto, também em ambos os *corpora*, ainda que se verifique um menor número de estruturas identificadas como marcadores discursivos (a cobertura no CORAL é de 18% e no LECTRA de 11%), aqueles que são identificados mostram uma classificação correta (a precisão no CORAL é de 53% e no LECTRA de 72% - Quadro 8).

Corpora	Kappa	Exatidão		MarcDisc		
		ZeroR	SVM	Precisão	Cobertura	Medida-F
Aulas para Diálogos	0,41	71,39%	74,97%	53%	18%	27%
Diálogos para Aulas	0,55	67,21%	79,97%	72%	11%	18%

Quadro 8: Classificação automática com SVM para amostras não-equilibradas com diferentes conjuntos de treino.

Aulas para Diálogos			
Classificada como	SUs	Disf	MarcDisc
SUs	5470	844	67
Disf	685	1100	49
MarcDisc	461	131	131

Quadro 9: Matriz de confusão para o *corpus* CORAL treinado com o *corpus* não-equilibrado LECTRA

Diálogos para Aulas			
Classificada como	SUs	Disf	MarcDisc
SUs	13642	1396	30
Disf	1891	4151	24
MarcDisc	651	497	138

Quadro 10: Matriz de confusão para o *corpus* LECTRA treinado com o *corpus* não-equilibrado CORAL

Os resultados obtidos são melhores para as disfluências do que para os marcadores discursivos, uma vez que, tal como descrito em Moniz *et al.* (2014a), existe uma marcação de contraste prosódico que se manifesta no

aumento de energia entre a região a reparar e a reparação de fluência em ambos os *corpora*. Esta estratégia de produção de disfluências não se observa nos marcadores discursivos de forma consistente, exibindo estes uma variedade de padrões prosódicos.

Ainda assim, estes resultados são muito promissores, uma vez que permitem prever que esta classificação será possível em diferentes conjuntos de teste de diferentes domínios.

6.2. Propriedades acústico-prosódicas mais relevantes

As experiências efetuadas confirmam a dificuldade em realizar uma predição automática dos marcadores discursivos e das disfluências. O elevado número de casos de SUs em ambos os *corpora* pode explicar por que razão as disfluências e os marcadores são mais facilmente confundidos com SUs do que entre si. No entanto, também era expectável uma menor precisão na discriminação entre marcadores discursivos e disfluências, uma vez que estes partilham vários comportamentos, nomeadamente: (i) ambos são considerados como parte da classe marcadores discursivos; (ii) ocorrem geralmente na mesma posição (Heeman e Allen, 1999; Goldwater *et al.*, 2010); (iii) contribuem para o conteúdo paralinguístico da fala (Aijmer, 2013; Moniz, 2013); e (iv) tendem a ser seguidos por uma pausa prosódica ou silenciosa ou uma ruptura melódica (Schiffirin, 1987; Moniz, 2013). Assim, uma vez que estas estruturas partilham todas estas propriedades, resta saber em que características diferem, de modo a ser possível identificar a função pragmática de cada uma, bem como prevêê-las automaticamente.

Tanto nas aulas como nos diálogos, as características prosódicas mais proeminentes extraídas do openSMILE (Quadros 3 e 4) entre as três classes de palavras, ou seja, marcadores discursivos, disfluências e SUs, são as seguintes: PCM, MFCC SMA (Banda de frequência Mel), vozeamento final, log HNR (*Harmonic Noise Ratio* – medida de rouquidão) e audSpec (variação espectral de áudio). Os primeiros, PCM e MFCC, correspondem às características mais informativas em muitas tarefas de processamento de fala, tais como identificação de falantes e de língua, informação paralinguística, entre outros.

A análise das características mais marcantes entre os marcadores discursivos e as SUs revela que estas são semelhantes em ambos os *corpora*. Nos diálogos, audSpec e vozeamento final são as características mais informativas e, nas aulas universitárias, estas são vozeamento final, f_0 final e audSpec. No entanto, as características *jitter* e *shimmer*, que também se mostraram relevantes nos estudos de deteção de disfluência por Moniz *et al.* (2014a), são encontrados somente no *corpus* dos diálogos. Estes resultados apontam para a importância das propriedades ao nível da fronteira dos marcadores discursivos, nomeadamente no que diz respeito à maior variação e declives de f_0 , bem como evidências de *creaky*, ou seja, perturbações nas sílabas finais.

Analisando os resultados já disponíveis para disfluências, observa-se que as características diferem das encontradas entre as classes SU e marcadores discursivos. De acordo com Moniz *et al.* (2014a, 2014b), sobre a caracterização prosódica de disfluências, as características mais informativas extraídas do openSMILE são MFCC e comprimento de audSpec. O que é específico das disfluências são as diferenças espectrais de áudio, em linha com Shriberg (2001), que mostram que as disfluências também apresentam perturbações fonético-fonológicas ao nível da fronteira de final de palavra. Os autores mostram evidências de uma marcação prosódica contrastiva entre a disfluência e a reparação de fluência para todos os falantes. Este comportamento consistente pode ser a razão pela qual as disfluências são mais bem preditas do que os marcadores discursivos.

De modo a comparar o comportamento prosódico dos marcadores discursivos com as disfluências, foram selecionados 30 exemplos dos marcadores discursivos mais frequentes de cada *corpus*. Devido à pequena quantidade de exemplos analisados, não é possível comparar o seu comportamento entre os diferentes *corpora*. O que se pretende aqui é destacar a variabilidade dos padrões prosódicos encontrados.

O marcador discursivo “agora” (Figura 1) tende a ser acentuado, com uma gama de variação alta dentro da sílaba acentuada e com valores de f_0 semelhantes ou mais elevados do que os constituintes prosódicos adjacentes. Estes dados mostram que há um esforço para marcar esta palavra e distingui-la dos contextos adjacentes. Por outro lado, os marcadores “portanto” (Figura 2) e “ok” são maioritariamente desacentuados, exibem contornos *plateau* e têm valores de f_0 menores do que a unidade prosódica seguinte, sendo, por conseguinte, proferidos num espaço tonal intermédio, quase como um suporte vocálico das unidades prosódicas adjacentes. Quer “então” como “pronto” e “portanto” apresentam declives de f_0 reduzidos, embora este último tenda a ser desacentuado e o primeiro seja, maioritariamente, acentuado.

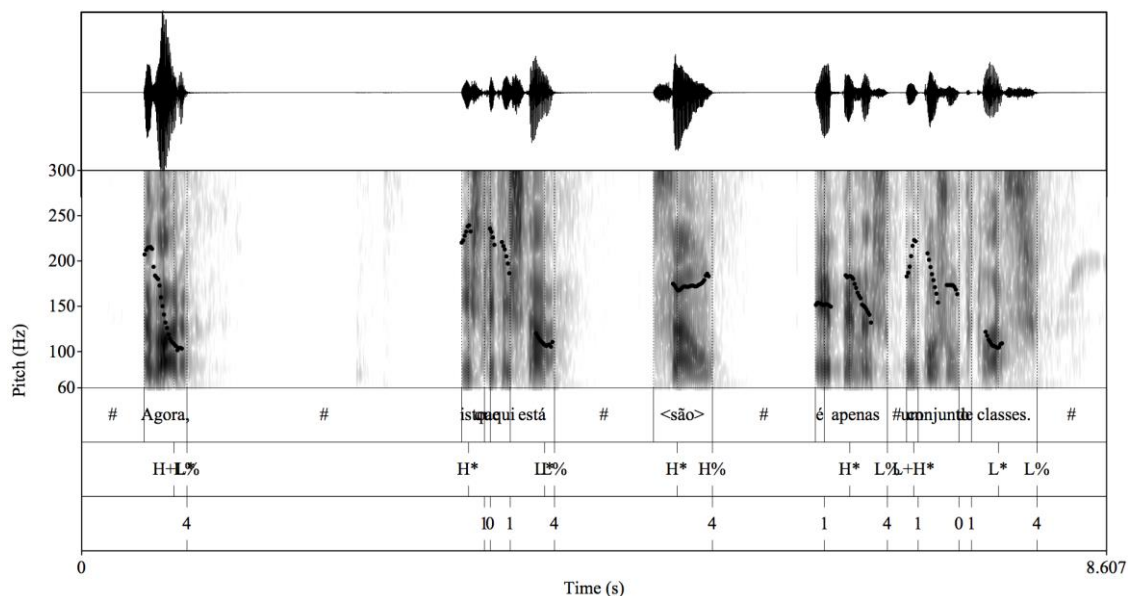


Figura 1: Exemplo do marcador discursivo “agora” no excerto “Agora, isto que aqui está <são> é apenas um conjunto de classes”. *Corpus* de aulas universitárias.

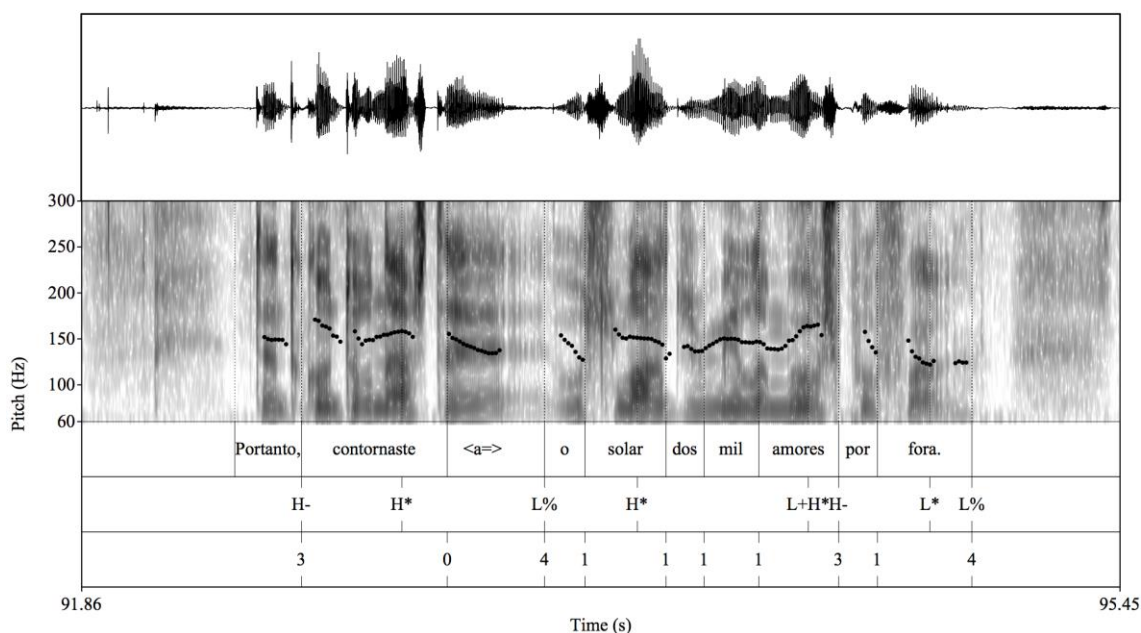


Figura 2: Exemplo do marcador discursivo “portanto” no excerto “Portanto, contornaste <a=> o solar dos mil amores por fora”. *Corpus* de diálogos.

Tendo em conta estes resultados, é possível salientar que, dependendo do marcador discursivo, existem comportamentos prosódicos distintos, embora todos eles compartilhem a mesma posição – início de enunciado. O facto de os marcadores discursivos serem associados a diferentes funções pragmáticas pode ser uma explicação para a variação encontrada. Pode-se supor que os marcadores que têm uma função semelhante às disfluências, por exemplo, ganhar tempo enquanto se pensa no que dizer em seguida, podem partilhar com eles algumas propriedades prosódicas, nomeadamente os contornos *plateau* que contrastam com a subida de f_0 nos constituintes prosódicos seguintes. Outros marcadores discursivos, como “agora”, introduzem um novo tópico e

são prosodicamente proeminentes. Para confirmar esta análise preliminar, será necessário realizar um estudo tonal e acústico mais detalhado dos marcadores discursivos e contextos prosódicos adjacentes.

7. Conclusão

Este trabalho representa uma primeira tentativa de identificar marcadores discursivos em dois *corpora* diferentes em PE, nomeadamente, aulas universitárias e diálogos em formato *map-task*, seguida de uma classificação automática dos mesmos, complementada por uma discriminação entre estes, disfluências e SUs.

Quanto ao tipo de marcadores discursivos utilizados em ambos os *corpora*, os resultados mostram que a seleção de marcadores discursivos depende do domínio e do falante. No mesmo *corpus*, existem falantes que tendem a utilizar o mesmo marcador e existem aqueles que variam entre diversas estruturas. Também se observa que os marcadores mais frequentes são semelhantes em ambos os *corpora*.

Considerando o facto de que quer os marcadores como as disfluências são descritos na literatura como estruturas paralinguísticas que partilham algumas propriedades (Goldwater *et al.*, 2010; Liu *et al.*, 2006), pretendeu-se aferir também se estas estruturas se distinguem prosodicamente. Para tal, efetuou-se uma tarefa de classificação automática multiclasse, para classificar as palavras que correspondem a marcadores discursivos, a disfluências ou a nenhuma destas (SUs).

Os resultados da tarefa de classificação mostram que a inclusão de pistas acústico-prosódicas melhora a predição destes eventos em 20%. Esta tarefa também evidencia que os marcadores discursivos são os mais difíceis de prever, devido principalmente à variabilidade do seu comportamento prosódico. As suas propriedades prosódicas são mais semelhantes à classe SU do que à das disfluências, o que coloca uma questão importante para uma análise mais aprofundada.

Para testar a robustez do sistema de classificação entre domínios, realizou-se uma outra experiência com dados de um *corpus* para treino e do outro para teste e vice-versa. Tal como expectável, nesta tarefa, os valores obtidos são mais baixos do que utilizando o mesmo *corpus* para treino e teste. No entanto as melhorias obtidas face à base de comparação em ambos os *corpora* mostram que esta classificação pode ser utilizada em dados fora do domínio.

No geral, apesar da complexidade desta tarefa, os resultados são muito encorajadores para a classificação automática multiclasse. Num trabalho futuro, pretende-se analisar com mais detalhe as características prosódicas e acústicas mais proeminentes, para compreender a sua relevância no processo de classificação. Pretende-se também incluir marcadores de discurso nos modelos de língua para PE já treinados com outros eventos, o que resultará em transcrições automáticas enriquecidas, bem como integrar os classificadores em sistemas de diálogo.

Agradecimentos

Este trabalho foi financiado com fundos nacionais através da FCT – Fundação para a Ciência e a Tecnologia, com as referências UID/LIN/00214/2013 e UID/CEC/50021/2013, pela bolsa de doutoramento SFRH/BD/96492/2013 e pós-doutoramento SFRH/PBD/95849/2013 e pelo projeto europeu SPEDIAL (FP7 611396).

Referências

- Adam, Jean-Michel (2008) *A lingüística textual. Introdução à análise textual dos discursos*. São Paulo: Cortez Editora.
- Aijmer, Karin (2013) *Understanding Pragmatic Markers: a Variational Pragmatic Approach*. Edinburgh: Edinburgh University Press.
- Aijmer, Karin, Ad Foolen & Anne-Marie Simon-Vandenberg (2006) Pragmatic markers in translation: a methodological proposal. In Kerstin Fischer (ed), *Approaches to Discourse Particles*. 101-114 Oxford/Amsterdam: Elsevier.
- Batista, Fernando (2011) *Recovering Capitalization and Punctuation Marks on Speech Transcriptions*. Dissertação de Doutoramento, Instituto Superior Técnico.
- Batista, Fernando, Helena Moniz, Isabel Trancoso, Nuno Mamede & Ana Isabel Mata (2012) Extending automatic transcripts in a unified data representation towards a prosodic-based metadata annotation and evaluation. *Journal of Speech Sciences*, 2, 115-138

- Beeching, Kate (2002) *Gender, Politeness and Pragmatic Particles in French*. Amsterdam: John Benjamins.
- Beeching, Kate (2014) Just a suggestion - Pragmatic marker just/e in French and English. Oral communication in the *International Workshop - Pragmatic Markers, Discourse Markers and Modal Particles: What do we know and where do we go from here?*, Università dell'Insubria, Como (Italy), 16-17 de Outubro, 2014.
- Borreguero, Margarita & Araceli López (2010) Los marcadores del discurso y la variación lengua hablada vs. lengua escrita. In Lamas, O.L. & Villa E. A. (coord.). *Los estudios sobre marcadores del discurso en español, hoy*. Madrid: Arco/Libros, 415-496.
- Brogueira, Gaspar, Fernando Batista, João Paulo Carvalho, Helena Moniz (2014) Expanding a Database of Portuguese Tweets. In *SLATE'14 3rd Symposium on Languages, Applications and Technologies*, Schloss Dagstuhl, vol. 4569, series OpenAccess Series in Informatics (OASICs), 275-282, Bragança, Portugal, Junho de 2014.
- Coutinho, Maria Antónia (2009) Marcadores discursivos e tipos de discurso. *Estudos Linguísticos / Linguistic Studies* 2, 193-210.
- Denke, Annika (2009). Nativelike Performance. Pragmatic Markers, Repair and Repetition in Native and Non-native English Speech. Saarbrücken: VDM Verlag Dr. Müller.
- Duarte, Inês (2000). Língua Portuguesa. Instrumentos de Análise. Lisboa: Universidade Aberta.
- Eyben, Florian, Martin Wöllmer, Bjorn Schuller (2010) openSMILE - The Munich Versatile and Fast Open-Source Audio Feature Extractor. *ACM Multimedia (MM), ACM, Firenze, Itália*, 25-29.10.2010.
- Fischer, Kerstin (1998) *Discourse particles, turn-taking, and the semantics-pragmatics interface*. *Revue de Sémiotique et Pragmatique* 8: 111-137.
- Fischer, Kerstin (2000) *From Cognitive Semantics to Lexical Pragmatics: The Functional Polysemy of Discourse Particles*. Mouton de Gruyter, Berlin/New York.
- Fraser, Bruce (1988) Types of English discourse markers. *Acta Linguistica Hungarica* 38, 19-33.
- Fraser, Bruce (1990) An approach to discourse markers. *Journal of Pragmatics* 14, 383-395.
- Fraser, Bruce (1999) What are discourse markers? *Journal of Pragmatics* 31, 931-952.
- Fraser, Bruce (2009) Topic orientation markers. *Journal of Pragmatics* 41, 892-898
- Freitas, Tiago, Maria Celeste Ramilo (2003) O actual estatuto da palavra portanto. In *Actas do XVIII Encontro da Associação Portuguesa de Linguística*, Lisboa.
- Goldwater Sharon, Dan Jurafsky, Christopher D. Manning (2010) Which words are hard to recognize? Prosodic, lexical, and disfluency factors that increase speech recognition error rates. *Speech Communication* 52:181-200.
- Hall, Mark, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann & Ian H. Witten (2009) The WEKA data mining software: an update. *SIGKDD Explorations*, 11(1):10-18.
- Heeman, Peter & James Allen (1999) Speech repairs, intonational phrases and discourse markers: Modeling speakers' utterances in spoken dialogue. *Computational Linguistics*, 25, 527-571.
- Hirschberg, Julia, Diane J. Litman (1993) Empirical studies on the disambiguation of cue phrases. *Computational Linguistics* 19 (3), 501-530.
- Jucker, Andreas H. & Sara W. Smith (1998) And people just you know like "wow": Discourse markers as negotiating strategies. In Jucker, Andreas H. and Yael Ziv (eds.). *Discourse Markers: Description and Theory*. Amsterdam: John Benjamins.
- Lease, Matthew, Mark Johnson (2006) Early deletion of fillers in processing conversational speech. In *Proceedings of HLT-NAACL 2006 (Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics)*, Companion Volume: Short Papers, 73-76. New York, NY.
- Liu, Yang, Elizabeth Shriberg, Andreas Stolcke, Dustin Hillard, Mari Ostendorf, & Mary Harper (2006) Enriching Speech Recognition with Automatic Detection of Sentence Boundaries and Disfluencies. In *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, n. 5, pp. 1526-1540.
- Lopes, Ana Cristina Macário (1997) Então: elementos para uma análise semântica e pragmática. *Actas do XII Encontro Nacional da APL*, vol. 1, 177-189. Lisbon: Colibri.
- Lopes, Ana Cristina Macário (2014) Aliás: A contribution to the study of a Portuguese discourse marker. In Piera Molinelli & Chiara Ghezzi, (Eds.), *Discourse and Pragmatic Markers from Latin to the Romance Languages*. Oxford University Press.
- Lopes, António, David Martins de Matos, Vera Cabarrão, Ricardo Ribeiro, Helena Moniz, Isabel Trancoso, Ana Isabel Mata (2015) Towards Using Machine Translation Techniques to Induce Multilingual Lexica of Discourse Markers, <http://arxiv.org/abs/1503.0914>.

- Makhoul, John, Francis Kubala, Richard Schwartz & Ralph Weischedel (1999) Performance measures for information extraction. In *Proceedings of the DARPA Broadcast News Workshop*, 249-252. Herndon, VA.
- Moniz, Helena (2006) *Contributo para a Caracterização dos Mecanismos de (Dis)Fluência no Português Europeu*. Dissertação de Mestrado, Universidade de Lisboa.
- Moniz, Helena (2013) *Processing disfluencies in European Portuguese*. Dissertação de Doutoramento, Universidade de Lisboa.
- Moniz, Helena, Fernando Batista, Ana Isabel Mata, Isabel Trancoso (2014a) Speaking style effects in the production of disfluencies. *Speech Communication*, vol. 65, 20-35, doi: 10.1016/j.specom.2014.05.004, Agosto de 2014.
- Moniz, Helena, Fernando Batista, Ana Isabel Mata da Silva, Isabel Trancoso (2014b) Discriminação automática de sinais de pontuação e de disfluências, In *Textos Seleccionados do XXIX Encontro Nacional da Associação Portuguesa de Linguística, ENAPL*, Outubro de 2014.
- Neto, João, Hugo Meinedo, Márcio Viveiros, Renato Cassaca, Ciro Martins & Diamantino Caseiro (2008) Broadcast news subtitling system in Portuguese. *Proceedings of ICASSP'08*, 1561–1564.
- Ostendorf, Mari, Benoît Favre, Ralph Grishman, Dilek Hakkani-Tür, Mary Harper, Dustin Hillard, Julia Hirschberg, Heng Ji, Jeremy G. Kahn, Yang Liu, Sameer Maskey, Evgeny Matusov, Hermann Ney, Andrew Rosenberg, Elizabeth Shriberg, Wen Wang, and Chuck Wooters (2008) Speech Segmentation and Spoken Document Processing. In *IEEE Signal Processing Magazine*, 59-69.
- Pimentel, Ana (2012) *Os marcadores conversacionais no ensino de Português Língua Estrangeira: um estudo de caso*. Dissertação de Mestrado, Faculdade de Letras, Universidade do Porto.
- Popescu-Bellis, Andrei, Sandrine Zufferey (2011) Automatic identification of discourse markers in dialogues: An in-depth study of like and well. *Computer Speech & Language* 25 (3), 499-518.
- Redeker, Gisela (1990) Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics* 14: 367-381.
- Samuel, Kenneth Brian (1999) *Discourse learning: an investigation of dialogue act tagging using transformation-based learning*. Dissertação de Doutoramento, Universidade de Delaware.
- Schiffirin, Deborah (1987) *Discourse Markers*. Cambridge University Press, Cambridge.
- Schiffirin, Deborah (2001) Discourse markers: Language meaning and context. In Deborah Schiffirin, Deborah Tannen and Heidi Hamilton (Eds.) *Handbook of Discourse analysis*. Oxford: Basil Blackwell
- Schourup, Lawrence (1999) Discourse markers. *Lingua* 107: 227-65.
- Shriberg, Elizabeth (1994) *Preliminaries to a Theory of Speech Disfluencies*. Dissertação de Doutoramento, Universidade da Califórnia.
- Shriberg, Elizabeth (2001) To "errrr" is human: Ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, 31:153–169.
- Soares da Silva, Augusto (2006) The polysemy of discourse markers: the case of pronto in Portuguese. *Journal of Pragmatics*, 38, 2188-2205.
- Trancoso, Isabel, Rui Martins, Helena Moniz, Ana Isabel Mata, & Maria do Céu Viana (2008). The LECTRA Corpus - Classroom Lecture Transcriptions in European Portuguese. In *Proceedings LREC'08*, Marrakech, Morocco.
- Trancoso, Isabel, Maria do Céu Viana, Inês Duarte & Gabriela Matos (1998) Corpus de Diálogo CORAL. In *Proceedings of PROPOR'98*, Porto Alegre, Brasil.
- Urbano, Hudinilson (2003). Marcadores conversacionais. In Preti, D. (org.). *Análise de textos orais*. 6a Edição, 93-116. São Paulo: Humanitas.
- Viana, Maria do Céu, Sónia Frota, Isabel Falé, Flaviane Fernandes, Isabel Mascarenhas, Ana Isabel Mata, Helena Moniz & Marina Vigário (2007). Towards a P_ToBI. In <http://www.ling.ohio-state.edu/~tobi/>.