

## Proposta de delimitação de elementos de segmentação discursiva baseada na prosódia

A análise da conversação explora diversas vertentes sob o mesmo título. Diversos trabalhos da área exploram teorias Bakhtinianas e levam em consideração elementos ideológicos e da anunciação, outros tendem para a linguística textual segundo uma perspectiva textual-interativa, na qual os trabalhos de Jubran (2006), Koch (2006) e Marcuschi (2006) se inserem. Tal perspectiva se preocupa com o funcionamento da língua em contextos de uso e, portanto, com a utilização da atividade discursiva em textos.

Uma perspectiva computacional de análise da conversação que considera aspectos formais e funcionais foi proposta por Grosz e Sidner (1986), e adotada por autores como Hirschberg e Litman (1993) e Lucente (2012). Tal proposta, conhecida como modelo de Grosz & Sidner, ou modelo G&S, visa a segmentação de textos orais em unidades discursivas menores, seguindo uma hierarquia que se apoia em três princípios: intenção, atenção e estrutura linguística.

O Modelo G&S foi desenvolvido como um modelo computacional que, ao descrever a estrutura do discurso, oferece as bases necessárias para sua descrição e significado. De acordo com as proponentes desse modelo, a descrição da estrutura do discurso desempenha um papel central no processamento da linguagem à medida que estipula restrições nas porções do discurso (GROSZ & SIDNER, 1986). Essa descrição é intimamente relacionada a duas questões: *o que distingue o discurso*, e *o que o faz coerente*? A tentativa de resposta a essas questões leva a dois aspectos não linguísticos fundamentais no desenvolvimento desse modelo, que são a *atenção* e a *intenção* dos participantes de uma conversa. A atenção é um fator essencial no processamento de enunciados em um dado discurso, enquanto a intenção desempenha um papel importante na explicação de como se estrutura o discurso, proporcionando coerência a este e ao próprio termo “discurso” (GROSZ & SIDNER, 1986, p. 175).

A hipótese defendida nesse modelo é de que qualquer discurso é composto por três elementos essenciais e distintos, mas que interagem entre si a todo momento, que são: i) a estrutura sequencial dos enunciados do discurso em um dado momento; ii) a estrutura das intenções envolvidas no discurso; e iii) o estado de atenção dos participantes envolvidos no discurso.

Sendo assim, podemos dizer que o discurso possui três componentes responsáveis por sua estruturação e interação: i) estrutura da sequência dos enunciados, ou estrutura linguística (*linguistic structure*); ii) a estrutura dos propósitos, ou estrutura das intenções (*intentional structure*); e iii) o estado do foco de atenção, ou estado de atenção (*attentional state*). Juntos, esses três constituintes da estrutura do discurso suprem a informação necessária para que os participantes da conversa possam determinar como um enunciado se encaixa com outras partes do discurso, possibilitando que os participantes entendam por que algo foi dito e o que isso significa (GROSZ & SIDNER, 1986, p. 177), sem que se toque em aspectos sobre o significado do discurso como um todo.

A estrutura linguística, entendida no modelo G&S como a estrutura dos enunciados que compõem o discurso, é responsável por agregar tais enunciados em segmentos de discurso. A estrutura das intenções compreende os propósitos que subjazem ao discurso e seus componentes e suporta a distinção entre os propósitos fundamentais ao discurso. Entre os participantes de uma conversa existe mais do que um único objetivo que os leva a participar de tal conversa. A distinção desses objetivos, ou intenções, é fundamental para o entendimento do discurso. Cada

intenção que subjaz a um discurso em particular é chamada de propósito do discurso (doravante DP – *discourse purpose*).

E por fim, o estado de atenção é definido como uma propriedade intrínseca do discurso e não dos participantes do discurso. Esse estado é (i) inerentemente dinâmico, fazendo um registro de objetos, propriedades e relações salientes em cada ponto do discurso (GROSZ & SIDNER, 1986) e (ii) modelado por um conjunto de espaços focais (doravante FS – *focus space*), que são mudanças que ocorrem no estado de atenção motivadas por um conjunto de regras de transição que especificam as condições para se acrescentar ou excluir esses espaços.

O modelo G&S se aplica diretamente à análise da estrutura do discurso por meio da segmentação discursiva e da construção de uma relação hierárquica entre esses segmentos, que se conectam por meio de elementos textuais que compreendem, além de conectivos como conjunções e preposições, elementos de construção textual. No entanto a detecção de quais elementos textuais se utilizar para a segmentação se apresenta bastante arbitrário, pois alguns segmentadores podem realizar uma segmentação mais fina, considerando quaisquer unidades sintáticas, enquanto outros podem optar por uma segmentação mais ampla, considerando apenas conjunções, por exemplo.

Uma forma de facilitar a segmentação está na análise prosódica dos enunciados, uma vez que os participantes de um diálogo utilizam desta informação, associada à estrutura linguística, para detectar os propósitos de cada segmento discursivo.

Lucente (2012) apresenta a transcrição e segmentação de um corpus de fala espontânea – o corpus VoCE (LUCENTE, 2012) – feito manualmente seguindo os pressupostos de segmentação estabelecidos no modelo G&S. Este estudo mostrou que existe uma coincidência entre as fronteiras destas unidades discursivas e acentos frasais detectados automaticamente. Uma amostra de aproximadamente 30 minutos de fala mostrou um agrupamento de dados em que para cada grupo acentual – intervalo entre dois acentos frasais – estão inseridos três segmentos discursivos.

Baseando-se nestes dados um experimento foi feito com uma nova amostra de fala espontânea, extraída do *Corpus Moda*, composto pela fala de profissionais da moda. Neste experimento as gravações foram segmentadas primeiramente em grupos acentuais, e posteriormente foi feita a verificação no texto para detectar que elementos prosódicos ou sintáticos correspondem no alinhamento aos trechos em que segmentos discursivos são delimitados.

Os resultados mostraram que o texto oral se segmenta em hesitações, parênteses, marcadores discursivos, como “tipo”, “bom” e “sei lá”, artigos, conjunções e elementos de negação.

A detecção de tais elementos pode auxiliar no estabelecimento de critérios de segmentação sintático-prosódicos para dados do PB mais claros para os usuários do modelo G & S do que os estabelecidos pelo próprio modelo.

Referências:

- GROSZ, B.J., AND SIDNER, C.L., (1986) "Attention, Intentions, and the Structure of Discourse", *Computational Linguistics*, p. 12:3.
- HIRSCHBERG, J., AND LITMAN, D. (1993) Empirical Studies on the Disambiguation of Cue Phrases, *Computation Linguistics*, 19-3, p. 501-530.

- JUBRAN, C. C. A. S. (2006) A perspectiva textual-interativa. In Gramática do português culto falado no Brasil. Volume 1: Construção do texto falado. JUBRAN, C. C. A. S. & KOCH, I. G. V. (Org.). Campinas: Editora da Unicamp.
- KOCH, I.G.V. (2006) Especificidade do texto falado. In Gramática do português culto falado no Brasil. Volume 1: Construção do texto falado. JUBRAN, C. C. A. S. & KOCH, I. G. V. (Org.). Campinas: Editora da Unicamp.
- LUCENTE, L. Aspectos dinâmicos da entoação e da fala no português brasileiro. Tese (Doutorado em Linguística) - Unicamp, Campinas, SP, 2012.