

Estrutura discursiva como limitadora de padrões entoacionais.

Luciana Lucente

Grupo de Estudos de Prosódia da Fala
Universidade Federal de Alagoas - Arapiraca
lucente.luciana@gmail.com

Resumo – Este trabalho busca mostrar que as proeminências presentes na prosódia do português brasileiro, como o acento frasal e os contornos entoacionais, estão alinhadas a uma segmentação discursiva ancora a divisão de seus segmentos discursivos no que chamamos neste trabalho de “palavras fortes”. A verificação de tal alinhamento corrobora a hipótese apresentada em Lucente (2012) de que existe um oscilador glotal presente no modelo de produção da fala proposto por Barbosa (2006).

Palavras-chave – estrutura discursiva; entoação; alinhamento; acento frasal.

Abstract – This paper aims to show that Brazilian Portuguese prosodic prominence, as the phrase accent and the intonational contours, are aligned to the discourse segmentation which anchors the division of its discursive segments in terms of we call “strong words”. The investigation of these alignments supports the hypothesis presented in Lucente (2012) that there is a glottal oscillator present in the speech production model proposed by Barbosa (2006).

Keywords – *discourse structure; intonation; alignment; phrasal stress.*

I. INTRODUÇÃO

A relação entre prosódia e discurso é bastante discutida, porém são poucos os trabalhos experimentais que mostram essa relação, sobre tudo no português brasileiro (PB). Em Lucente (2012) essa relação é explorada ao investigar o alinhamento entre fronteiras entoacionais e acentos frasais, segundo o sistema DaTo de notação entoacional (Lucente, 2008, 2012) e unidades discursivas, segmentadas segundo o modelo de computacional de Grosz e Sidner (1986).

O modelo de Grosz e Sidner (doravante modelo G&S) foi desenvolvido a princípio como uma teoria computacional que, ao descrever a estrutura do discurso, oferece as bases necessárias para a descrição e seu 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. Paralelamente, existem ainda dois fatores importantes na constituição desse modelo, que são os papéis dos participantes em uma conversa e o significado contido no discurso.

Assumindo que o discurso seja uma peça do comportamento linguístico que envolve múltiplos enunciados e participantes, ele pode ser produzido por um ou mais participantes desse evento, pois em uma conversa, por exemplo, mais de um dos envolvidos pode falar e compor diferentes enunciados do discurso. Portanto falar de falante e ouvinte, ou locutor e interlocutor, é inviável, pois em uma situação de conversação não é claro o papel que cada participante desempenha.

Sendo assim, podemos dizer que o discurso possui três componentes responsáveis pela 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 determinado enunciado se encaixa com as outras partes do discurso, possibilitando que os participantes entendam por que algo foi dito e o que isso significa sem que se toque em aspectos sobre o significado do discurso como um todo (Grosz & Sidner, 1986, p. 177).

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* (doravante DS – *discourse segments*).

O desafio desse modelo é a forma de delimitar os segmentos de discurso, pois embora diferentes pesquisas em diferentes teorias (Cohen & Levesque, 1980; Polanyi & Scha, 1986) tenham analisado uma variedade de gêneros de discurso não houve concordância suficiente sobre como delimitar as fronteiras dos segmentos de discurso. Existem várias pistas linguísticas que apontam para tais fronteiras, como pistas prosódicas tais como pausas, alongamentos (Chafe, 1979,

1980), taxa de elocução (Butterworth, 1975) e pistas textuais, como o uso de subordinações, coordenações, pronomes, referentes e marcadores discursivos (Hirshberg & Litman, 1993). Essas pistas textuais podem indicar inclusive as mudanças na estrutura das intenções ou o estado atencional do discurso. O exemplo de segmentação dos enunciados no Quadro 1 a seguir, extraído do Corpus VoCE, é ilustrativo de como tais expressões linguísticas (em itálico) funcionam como marcadores de fronteiras discursivas.

Quadro 1: Exemplo de expressões linguísticas como marcadores de fronteiras discursivas

0 - 1 - (1) *mas hoje em dia* por exemplo eu to lançando o u::
um meu disco novo no formato dum livro também
(2) *então* vem o livro e o CD juntos
(3) *que é* o Sambazz :::
2 - (4) *e o Samba/* e o livro explica a produção do disco
(5) *então* além de ter o disco ainda tem o livro :
3 - (6) *que* eu escrevi todo ::: falando da produção do
disco desd'a:: das composições até a finalização
(7) *também* pra despertar um interesse na pessoa
[assim] não simplesmente em baixar as o disco
na in/ na internet de graça e :: ouvir as músicas ::
(8) *mas* também pro cara : ter :: ter a vontade de ler
o livro ::: e:: e saber da história do disco né?

Quadro 2: segmentação hierárquica do discurso em diferentes DS.

DS1 - (1) porque : tem uma con/ existia uma convicção de
que o mundo ia acabar ::: né?
(2) no dia que concid/ coincidisse isso :::
DS2 - (3) só que assim eh ::: a superstição o que que a
superstição é?
(4) ela nada mais é do que aquela velh' insegurança né?
(5) então eu acredito em tudo :: eu todo mundo o que
todo mundo me fala eu eu : coloco em :::
» **DS3** - (6) porque a a superstição ela vem :: de uma pessoa
que num pensa ::: né?
(7) se você for parar e pensar como é que um
» número pode ter poder e força em cima de você?
DS4 - (8) agora o que qu' ele tem poder e força?
(9) porque na bíblia fica/ na bíblia sagrada católica
apostólica romana fala que o número meia meia
meia : é o número da besta
» » **DS5** - (10) só que o número da best' ele num tem zero
meia zero meia de dois mil e seis :
(11) é o número meia meia meia ::
(12) é diferente :: né? :::
» **DS6** - (13) então é por causa disso que acredito que há ::
essas superstições

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 um dado discurso existe mais do que um único objetivo que os leva a participar de uma 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*), como aparece na segmentação do *corpus* VoCE (Lucente, 2012) transcrita no Quadro 2.

Aplicando esse modelo à análise da estrutura discursiva da fala espontânea pode-se observar que para cada segmento discursivo estão alinhadas fronteiras de *grupos acentuais* (Barbosa, 2006) e *contornos entoacionais* específicos, que são componentes prosódicos.

Os grupos acentuais são unidades delimitadas por dois acentos frasais consecutivos. Os acentos frasais não se referem a aspectos sintáticos, mas sim a aspectos fonéticos, pois são definidos como proeminências no domínio da produção da fala, assinaladas por picos locais de duração ao longo dos enunciados (Barbosa, 2006, p.10).

Os contornos entoacionais são definidos de acordo com o sistema DaTo de notação entoacional (Lucente, 2008; 2012), na Tabela 1 abaixo, e se referem a *uma unidade tonal que contém elementos comunicativos expressos em uma trajetória ideal da curva entoacional, especificada por um alvo a ser atingido e associada a uma unidade segmental linguística*, representado pela Figura 1, a seguir. Esse alinhamento pode ser explicado ao se adotar o modelo de produção a fala presente em Lucente (2012) a partir de um acréscimo de um oscilador glotal ao modelo proposto originalmente por Barbosa (2006).

Tabela 1: Contornos dinâmicos do Sistema DaTo.

Contornos Dinâmicos	
LH	<i>rising</i>
>LH	<i>late rising</i>
vLH	<i>compressed rising</i>
HLH	<i>falling-rising</i>
LHL	<i>rising-falling</i>
HL	<i>falling</i>
>HL	<i>late falling</i>
vHL	<i>compressed falling</i>
Níveis de Fronteira	
L	<i>low</i>
H	<i>high</i>

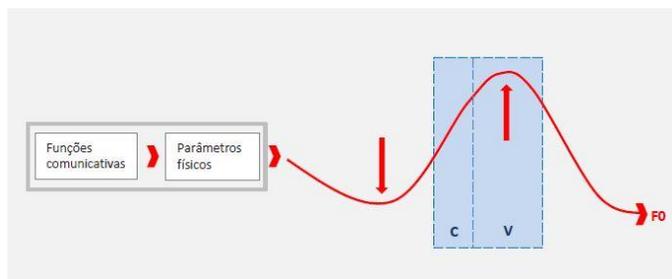


Figura 1: representação do contorno entoacional dinâmico.

II. EXPERIMENTO

Para verificação da hipótese, um experimento foi realizado em duas etapas. Na primeira etapa do experimento foram selecionados quatro arquivos de áudio do corpus VoCE com exemplos de fala espontânea, sendo uma em taxa de elocução lenta, uma normal e uma rápida. O arquivo que apresenta fala em taxa de elocução considerada normal se trata de uma fala *nonsense*, porém com estrutura sintática preservada.

Foi proposto primeiramente a um grupo de 12 estudantes de letras (todos nascidos em Belo Horizonte com faixa etária entre 20 e 25 anos) a indicação de proeminências na transcrição das falas em formato impresso ao mesmo tempo em que lhes eram apresentadas as respectivas gravações dos textos. Essa marcação de locais de ocorrência de proeminências foi comparada à marcação de grupos acentuais feita automaticamente pelo programa SG Detector (Barbosa, 2006).

Na segunda etapa do experimento foi proposto para seis estudantes (um subgrupo do grupo anterior), em um intervalo de 10 dias após o primeiro experimento, a segmentação dos mesmos textos em unidades discursivas, seguindo o modelo de segmentação de Grosz e Sidner. Antes da realização no segundo experimento os alunos foram submetidos a um treinamento de 6 horas para aprendizagem da segmentação segundo o modelo referido.

Para a primeira parte foi feita uma contagem da porcentagem de ocorrências por item marcado como proeminente, sendo consideradas relevantes as marcações que continham mais de 50% de ocorrências (igual ou mais que seis marcações). Para a segunda parte do experimento também foi analisada a concordância entre os sujeitos quanto à segmentação em unidades discursivas, sendo que as marcações que obtiveram valor acima de 50% de ocorrências como fronteiras na segmentação discursiva foram consideradas “marcações fortes”, ou ocorrendo em “palavras fortes”.

A partir das indicações de proeminência e segmentação foi feito um cruzamento desses dados a fim de verificar o quanto a estrutura discursiva e a prosódica estão inter-relacionadas, como nos exemplos:

a) segmentação em unidades discursivas com palavras fortes em negrito:

*olha eu queria **chorar***
*porque eu não tenho condição de prever o meu **futuro***
*além do ponto final se tornando uma outra língua **brasileira***
*porque eu acho que o futuro a Deus **pertence***
*e **Aton***
*então o futuro **pertence** ao ponto **Aton***
*porque eu acho que o nosso futuro é muito **reluzente***
*que no futuro a gente **pira***
*não vai ter mais a que se **reportar***
*então as réplicas elas são **constantes** (...)*

b) marcação de proeminências :

*olha(8) eu queria **chorar**(6) porque eu não tenho assim(3)*
*ahh(4) **condição**(8) de prever o meu **futuro**(10) além do*
*ponto(3) final se tornando uma(1) outra **língua brasileira**(6)*
*porque eu acho que o **futuro**(10) a deus **pertence**(8) e **aton**(9)*
*então o futuro **pertence**(1) ao **ponto**(4) **aton**(6) porque eu*
*acho que o nosso futuro é muito **reluzente**(6) que no futuro a*
*gente **pira**(8) não(2) **vai**(7) ter(3) mais a que(3) se **reportar**(4)*
*então as réplicas(4) elas são **constantes**(10) (...)*

III. PRIMEIROS RESULTADOS OBTIDOS

Os resultados preliminares mostraram: i) probabilidade alta de ocorrer o alinhamento entre as marcações de proeminências feitas pelos sujeitos e pelo programa SG Detector; ii) probabilidade alta de alinhamento entre as palavras marcadas como mais proeminente (tanto manualmente quanto automaticamente) e fronteiras de segmentos discursivos quando ocorrem nas chamadas fronteiras fortes, como mostra a Figura 2. Quanto à taxa de elocução, foi observada menor concordância no alinhamento entre as marcações de proeminências no áudio que apresentava taxa de elocução mais rápida e as fronteiras de seus segmentos discursivos marcados pelas palavras fortes.

Em relação à marcação dos contornos entoacionais, como previsto, estes acompanham a marcação de proeminências, e são predominantemente contornos de fronteira LHL e HL, o que indica a estrutura discursiva influenciando a prosódica, e vice versa. Os outros tipos de contornos, como LH, >LH e HLH ficam restritos à marcação de foco, e raramente ocorrem em posição de fronteira. Tais contornos indicam os itens marcados como proeminentes que não se alinham discursivamente com fronteiras, como se pode ver nas palavras “chorar” e “futuro” assinaladas na Figura 2.

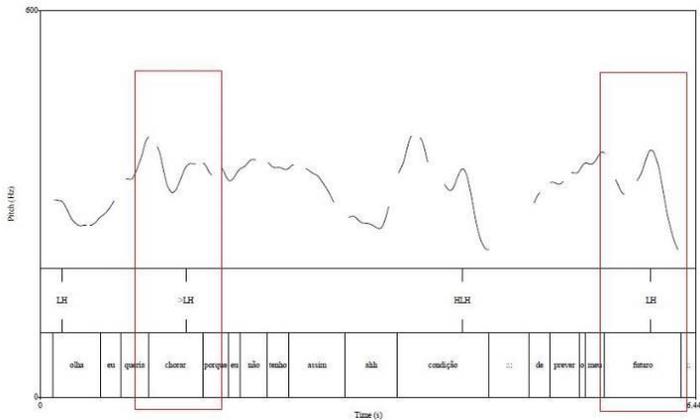


Figura 2: contornos >LH e LH marcando fronteiras discursivas, respectivamente, em palavras fortes como “chorar” e “futuro”.

IV. DISCUSSÃO

A discussão a que o resultado do experimento preliminar conduz é sobre o fato de as estruturas discursiva e prosódica estarem ligadas de forma bastante intrínseca e como partes de um sistema dinâmico de produção da fala, podendo ter como atrator os grupos acentuais. Outro ponto a ser discutido é em que nível se dá essa interação, pois de acordo com o modelo de Barbosa (2006) e Lucente (2012), o processamento do discurso está em um nível superior na produção da fala, juntamente com a sintaxe e a semântica, como um componente cognitivo, e a prosódia em um nível executivo, ou cognitivo-executivo.

V. REFERÊNCIAS

Barbosa, P. A. (2006). *Incursões em torno do ritmo da fala*. Campinas: Pontes.
 Chafe, W. (1976) “Givenness contrastiveness de_niteness subjects topics and point of view”. In Li, C., (ed.) *Subject and Topic*, Academic Press New York, p. 25-55.

Chafe, Wallace L. (1979) The Flow of Thought and the Flow of Language. In Givon, T., Ed., *Syntax and Semantics, Vol. 12, Discourse and Syntax*. Academic Press, New York, New York: 159- 182.
 Chafe, W.L. (1980) The Deployment of Consciousness in the Production of a Narrative. In Chafe, W.L., Ed., *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production. Vol. 3. Advances in Discourse Processes*. Ablex Publishing Corp, Norwood, New Jersey: 9-50.
 Grice, M., Ladd, D. R., Arvaniti, A. (2000) On the place of phrase accents in intonational phonology. *Phonology* 17: p. 143-185.
 Grosz, B.J., and Sidner, C.L., (1986) "Attention, Intentions, and the Structure of Discourse", *Computational Linguistics*, p. 12:3.
 Lucente, L. (2012) *DaTo: Um sistema de notação entoacional do português brasileiro baseado em princípios dinâmicos. Ênfase no foco e na fala espontânea*. Dissertação de Mestrado. Unicamp.
 Lucente, L., Barbosa, P. A. (2010) The role of alignment and height in the perception of LH contours. *Proceedings of Fifth Conference on Speech Prosody*. Chicago.
 _____ (2008) Narrow focus in Brazilian Portuguese: spatial and temporal constraints. *Proceedings of Fourth Conference on Speech Prosody*. Campinas.
 _____ (2007) Notação Entoacional do Português Brasileiro em Corpora de fala Semi-Espontânea e Espontânea. *Revista Intercâmbio* 16.
 Prince, E. F., (1981) “Toward a taxonomy of given-new information”, In Peter Cole (ed.), *Radical Pragmatics*, The Academic Press, New York, p. 223-255
 Saussure, F. et al. (2002) *Curso de linguística geral*. 24. ed. São Paulo, SP: Cultrix.
 Sonntag, G, P. & Portele, T. (1998) Comparative evaluation of synthetic prosody with the PURR method. *International Conference on Spoken Language Processing*.