

Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação

Getting started with prosody crucial theoretical and methodological aspects of the sound component that shapes our utterances

Plínio A. Barbosa
Universidade Estadual de Campinas - Unicamp

Resumo

O trabalho tem a finalidade de apresentar aspectos teóricos e metodológicos da área de prosódia da fala para servir de ponto de partida para o pesquisador iniciante. A partir da significação platônica do termo “prosódia” desenvolve as acepções ligadas ao termo na pesquisa científica contemporânea, apresentando noções como proeminência, fronteira prosódica, acento frasal, grupo acentual, foco, ênfase, entoação e ritmo. Aspectos metodológicos da área como a montagem de *corpora* de fala de laboratório com relação com a fala espontânea, a normalização da duração silábica e a diferenciação da prosódia na produção da prosódia na percepção da fala são discutidos com o fim de apresentar algumas questões de relevo para a área de pesquisa.

Palavras-chave

Prosódia, Experimentação, Entoação, Ritmo.

Abstract

This work aims at presenting some theoretical and methodological aspects of prosody research to newcomers. The work gives the Platonic origin of the word “prosody” from which the contemporary meanings developed, to include notions such as prominence, prosodic boundary, phrase stress, stress group, focus, emphasis, intonation and rhythm. Selected methodological aspects are shown, such as the construction of corpus, the relationship between lab and spontaneous speech, syllabic duration normalization, and the difference between the notions attached to prosody production and perception. From these aspects some crucial issues in speech prosody are discussed.

Keywords

Prosody, Experimentation, Intonation, Rhythm.

1. Introdução: o que é a prosódia?

O termo *prosódia* se origina da palavra bimorfêmica grega *pros-wdiā*, utilizada pela primeira vez na *República* de Platão na expressão *phthongous te kai prosōdias* (LIDDEL; SCOTT; JONES, 1996, p. 399a) para opor o conteúdo segmental às variações melódicas presentes em formas de narrar por imitação. Essas variações melódicas são comparadas no texto platônico a narrações imitativas semelhantes ao canto ou em “sintonia com o canto” (*pros* = “em direção a, junto com”, *wdiā* = “canto”). Na *República*, Sócrates condena essas narrativas pela possibilidade de corromperem a imagem que os jovens tinham de personagens célebres, como heróis de guerra. Mais tarde identificada na poesia às características quantitativas e melódicas da versificação, acaba recebendo um caráter normativo, associado a ou incluso em regras da boa pronúncia ou ortoépia. Na Linguística, o termo *prosódia* receberá um sentido especializado (DUCROT; TODOROV, 1972, p. 228).

Definida pelos estruturalistas e funcionalistas, num primeiro momento, de forma negativa, como “todos os fatos de fala que não entram no quadro fonemático, isto é, aqueles que não concernem, de uma forma ou outra, a segunda articulação” (MARTINET, 1991, p. 83), a *prosódia* está, no cenário de pesquisa atual, associada a fatores linguísticos como acento, fronteira de constituinte, ênfase, entoação e ritmo, a fatores paralinguísticos como marcadores discursivos (e.g., “né”, “entendo”, “an-han”) e atitudes proposicionais (e.g., “confiante” e “duvidoso”) e sociais (e.g., “hostil” e “solidário”), além de tratar de fatores extralinguísticos como as emoções. Todos esses fatores se combinam com aspectos sociais e biológicos indiciais como gênero, faixa etária, classe social, nível de escolaridade, entre outros.

Aos estudos de *prosódia* cabe a análise de unidades fônicas e de suas relações a partir da sílaba até trechos de discurso cuja extensão máxima tem se ampliado de acordo com os interesses dos especialistas. É a *prosódia* que molda nossa

enunciação imprimindo a “o que se fala” um “modo de falar” que é dirigido intencionalmente ou não ao ouvinte.

Do ponto de vista do especialista, a análise prosódica é feita segundo os eixos linguísticos tradicionais, o eixo sintagmático e o eixo paradigmático, tanto do ponto de vista fonológico quanto fonético. Essa análise se dá em uma amostra ou instantâneo de fala que se denomina *corpus*. O *corpus* é montado com trechos da fala de um ou mais informantes e o seu conteúdo decorre da escolha prévia de variáveis selecionadas pelo pesquisador, para testar uma teoria científica. Como em toda área de pesquisa com os sons da fala, o *corpus* pode ser inteiramente concebido e gravado em laboratório ou provir de gravações de fala espontânea. Convém esclarecer acerca das vantagens e desvantagens desses dois tipos de *corpus*.

2. Da fala espontânea e de laboratório

O contraste entre fala de laboratório (*lab speech*), fala semi-espontânea e fala espontânea, que faz a literatura fonética, mascara dois aspectos essenciais: (1) o elemento primordial do ponto de vista de uma distinção metodológica, qual seja, a dimensão de controle do experimentador sobre a eliciação do *corpus*, (2) bem como o papel dos gêneros discursivos no contraste acima explicitado. O gênero do material coligido para análise é uma dimensão ortogonal à dimensão do nível de controle do experimentador sobre a eliciação, isto é, podemos variar graus ou tipos de uma dimensão independentemente da outra. É preciso desvincular definitivamente o termo de fala espontânea da conversa livre, pois um ditado, uma leitura, uma narrativa para uma criança e uma entrevista são igualmente espontâneos se ocorrerem em situação natural de comunicação. São, assim, instâncias legítimas da comunicação humana em que o pesquisador não intervém, apenas seleciona o material para análise a partir de critérios teóricos. Da mesma forma, o termo “fala de laboratório” não é um sinônimo do resultado da leitura de frases isoladas preparadas pelo experimentador. Uma fala eliciada a partir de figuras ou a partir de um tema fornecido pelo experimentador ou uma narração solicitada após a leitura de um texto ou ainda uma entrevista guiada pelo experimentador são todos exemplos de fala de laboratório.

A dimensão essencial para a caracterização da espontaneidade é o grau de intervenção do experimentador em relação à eliciação do material. Essa intervenção pode ser extrema, como no caso em que o experimentador solicita

a um informante a imitação fiel da fala de um outro (nesse caso o que se fala e como se fala são rigidamente apresentados como modelo) ou a intervenção do experimentador pode ser bastante livre, como na situação em que solicita que, a partir de um mapa, um informante explique a outro como se faz para ir de um lugar para outro (*map task*). São exemplos como o último, em que não há texto prévio, apenas itens lexicais definidos pelo mapa (escolhido pelo experimentador!) e pelo mecanismo de explicação de uma rota (que de certa forma impõe uma sintaxe própria: o experimentador escolheu a tarefa!), que se fala de fala semi-espontânea. A fala semi-espontânea é exemplo de fala de laboratório, eliciada para o estudo prosódico. Como se deduz do exposto, o que diferencia os vários tipos de fala de laboratório é o grau de intervenção do experimentador, que, sendo cada vez menor, aproxima a fala de laboratório da fala espontânea para o gênero discursivo ou enunciativo específico. Fala espontânea é a situação em que não há intervenção do experimentador. Assim, o termo “espontânea”, usado restritivamente para uma conversa, perde um pouco o sentido, pois a fala espontânea pode se dar em diversos gêneros enunciativos ou discursivos. Por isso, na descrição de todo *corpus* é preciso informar o nível de controle do experimentador sobre a eliciação.

Mas não basta informar esse grau para caracterizar um *corpus*, é preciso informar o gênero enunciativo. Isso é válido mesmo se não há controle algum do experimentador sobre a eliciação do *corpus*, se, por exemplo, o experimentador fez uma seleção de material numa situação habitual da comunicação. O gênero determina um *modus operandi* do sujeito em relação a sua enunciação. Assim, a mudança no comportamento de um informante durante uma entrevista, por exemplo, não ocorre somente por conta de possível controle do experimentador, mas decorre também da situação. Essa dimensão é ortogonal ao grau de controle do experimentador e é própria do gênero entrevista. Assim, durante a entrevista, a situação de gravação e o tema podem produzir segmentos discursivos que soam mais ou menos como uma conversa ou depoimento, ou opinião. Assim, o gênero enunciativo precisa ser claramente descrito. Exemplos de gêneros que já receberam atenção em estudos prosódicos são: entrevista, narrativa, ditado, leitura em sala de aula habitual, imitação de uma pessoa, interpretação de um texto em cena.

A FIG. 1 apresenta de forma ilustrativa a proposta de classificação dos *corpora* segundo as duas dimensões mencionadas. Observe-se que as narrativas com tema ou texto a ser recontado proposto pelo experimentador são exemplos

para os quais o eixo de controle está muito próximo do extremo correspondente ao eixo de espontaneidade, para o qual não há controle do experimentador. Outro exemplo é o caso em que o experimentador seleciona o texto a ser lido em voz alta, guiado pela teoria que vai avaliar.

A proposta não estabelece que um gênero não possa contar traços de um outro, muito embora o eixo das ordenadas possa dar a entender isso. No caso de narrativas por recontagem de história, por exemplo, a extensão, o tema e / ou o vocabulário do texto lido previamente fazem com que a narração apresente muitos fenômenos da conversa espontânea, como hesitações, reformulações, falsos inícios, repetições, alongamentos, bloqueios, entre outros. A própria conversa espontânea contém instâncias narrativas diversas.

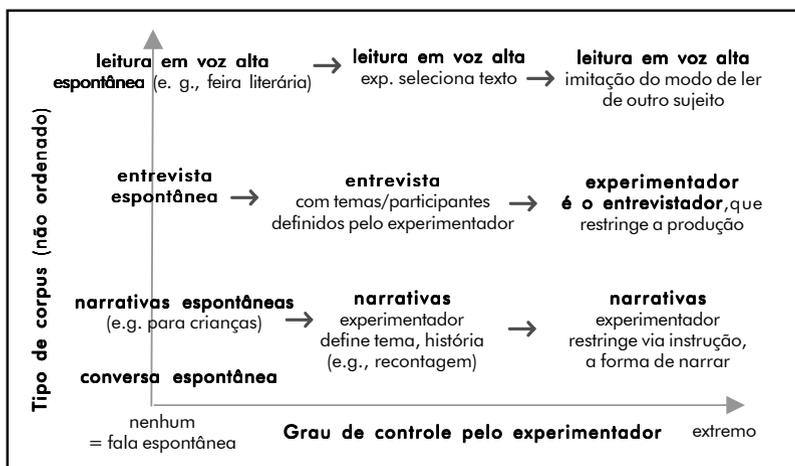


FIGURA 1 – Proposta de classificação de *corpora* segundo eixo “grau de controle do experimentador” e “gênero ou tipo de *corpus*”

A montagem teoricamente fundamentada e experimentalmente apropriada de *corpora* de fala é essencial para a pesquisa sobre as funções comunicativas da prosódia da fala.

3. Funções prosódicas na comunicação

É possível identificar funções prosódicas tanto no plano linguístico quanto expressivo. Quando se fala de função comunicativa, evidentemente refere-se a

um aspecto audível da prosódia, passível, assim, de ser interpretado pelo nosso sistema perceptivo. Do ponto de vista da produção da fala, pode haver saliências acústicas implementadas pelo falante que não ultrapassam o limiar de audibilidade em função do contexto de sua realização ou do ouvinte. Voltaremos a esse ponto na quinta seção.

Nos planos linguístico e paralinguístico, distinguem-se: (1) as funções discursivas dialógicas como os marcadores de turno e os *backchannels* (e.g., *um-hum*, *entendo*) e não dialógicas como a modalidade de um enunciado de um trecho lido ou de um monólogo, (2) as funções demarcativas que assinalam limites de constituintes prosódicos como sílabas, palavras fonológicas e grupos acentuais, (3) bem como as funções de marcação de proeminência que assinalam a saliência de um constituinte prosódico em relação a outro.

Como exemplo de função discursiva, tem-se a oposição entre o enunciado assertivo “Pedro fez bem seu trabalho.” e o enunciado interrogativo “Pedro fez bem seu trabalho?” Linearmente, a distinção entre as formas de organização prosódica realiza sozinha o contraste entre os dois enunciados. No contraste entre essas modalidades no português brasileiro, destaca-se a forte subida da frequência fundamental (o correlato acústico da frequência de vibração das pregas vocais), sobretudo na última sílaba tônica (no exemplo dado, em “-ba-”) do enunciado interrogativo, em relação à subida lenta do mesmo parâmetro no enunciado assertivo.

Como exemplo de função demarcativa ou de segmentação (*phrasing*), pode-se apontar o papel da pausa (silenciosa ou preenchida), para assinalar uma fronteira prosódica: compare o enunciados “Maria | dança bem. ||” com o enunciado “Maria dança | divinamente bem. ||”, em que o número de barras verticais assinala pausas de maior duração. A razão da diferença na extensão dos constituintes e do deslocamento da fronteira após “Maria”, no primeiro enunciado, para a fronteira após “dança”, no segundo enunciado, é a necessidade de dividir ambos os enunciados em constituintes prosódicos de tamanho semelhante quanto ao número de sílabas. Por conta disso, a produção da prosódia é autônoma em relação à sintaxe, pois é regida por outros princípios. Observe que o primeiro constituinte prosódico do segundo enunciado não corresponde a nenhum constituinte válido em teorias sintáticas fundamentadas nessa noção.

Como exemplo de função de proeminência (*prominence*), examinemos o adjetivo “verde” no enunciado “Não, eu vi uma moto VERDE.”, pressupondo um contexto em que algo sobre uma moto de outra cor fora comentado. Nesse caso, há

uma relação sintagmática no enunciado, pois o trecho “VERDE” tem mais energia, maiores valores de frequência fundamental e maior duração do que os elementos da vizinhança, bem como uma oposição paradigmática em que o foco de contraste no adjetivo chama a atenção para a cor verde da moto em substituição a outra cor evocada em pelo menos um enunciado anterior. A FIG. 2 ilustra a curva ou contorno de frequência fundamental do enunciado referido acima para um informante paulista masculino de cerca de 40 anos na época da gravação. Sua finalidade é mostrar que pontos singulares dessa curva são correlatos da proeminência percebida pelos ouvintes nessa situação de enunciação. A duas interrupções da curva, após “Não” e entre “moto” e “VERDE” correspondem respectivamente a uma pausa silenciosa e ao intervalo de oclusão do [t] de “moto”. O que é usado pelo informante para assinalar para o ouvinte a importância da palavra “verde” é a realização de uma forte descida da curva de frequência fundamental na sílaba tônica da palavra “VERDE”. Para realizar essa descida, o informante eleva o valor da frequência fundamental com um movimento de subida iniciado durante a palavra precedente “moto”. Apesar de o pico da curva se encontrar ao final da sílaba pós-tônica “to”, isso não torna essa palavra proeminente para o ouvinte, pois são relevantes para um falante do português brasileiro apenas os movimentos singulares que se passam durante as sílabas tônicas das palavras fonológicas.

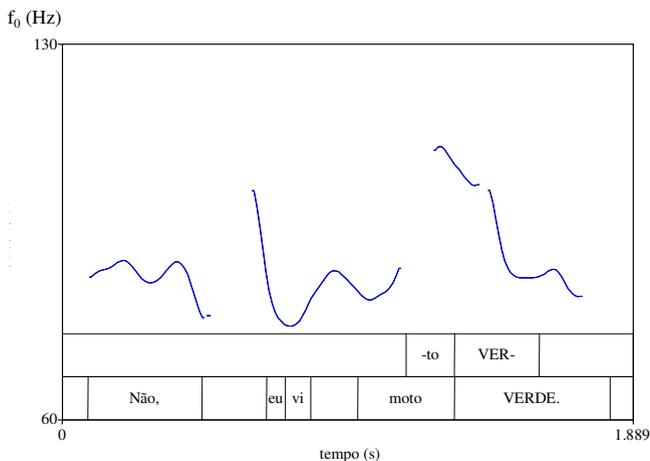


FIGURA 2 – Curva de frequência fundamental do enunciado “Não, eu vi uma moto VERDE.”

Já no plano expressivo, distinguem-se as funções prosódicas atitudinais (atitude, postura interpessoal, estilo de elocução), afetivas (emoções como tristeza, alegria e raiva e afetos como o humor) e indiciais (marcas de gênero e sexo, origem social e dialetal, entre outras). Exemplos dessas funções são encontrados em todo enunciado, tendo em vista que uma atitude, uma emoção e marcas indiciais são traços dificilmente disfarçáveis na enunciação. Afinal, não reconhecemos, mesmo na fala de desconhecidos, se são homens ou mulheres, qual sua faixa etária e seu estado afetivo, entre outros aspectos?

Para uma introdução aos aspectos expressivos da prosódia, ver os trabalhos de Fónagy (1986), Bolinger (1986) e Scherer (1984). A chamada Fonoestilística estuda as funções prosódicas atitudinais e a variabilidade prosódica intra- e inter-sujeito, ligada a escolhas voluntárias ou não (cf. LÉON, 1971).

4. Demarcação da prosódia em relação ao ritmo e à entoação

Há uma ambiguidade entre os termos prosódia e aqueles de ritmo e entoação, sobretudo entre o primeiro e o último que tentarei esclarecer aqui. Autores como Hirst e Di Cristo (1998, p. 1-44) tomam a prosódia como o termo mais geral, compreendendo tanto o domínio propriamente lexical, ao qual apontam os estudos de tom em línguas tonais, do acento lexical, quanto o domínio não-lexical ou pós-lexical, ao qual referem a variação no enunciado da frequência fundamental, da duração e a marcação de fronteiras prosódicas. A entoação é o estudo das relações abstratas que envolvem o domínio não-lexical, independentemente do correlato físico ou perceptivo que a veicula. Sendo assim, para esses autores (entre outros), entoação não concerne apenas a percepção do *pitch* (altura) ao longo do tempo. A entoação também abarcaria o estudo da duração e seu correlato físico e perceptivo ao longo do enunciado, o ritmo.

Outro recorte possível é pela via dos efeitos perceptivos ao longo do enunciado das sensações de duração, *pitch* (altura) e volume (*loudness*). Nesse recorte, “prosódia” também é o termo mais geral, compreendendo ainda o domínio lexical e o pós-lexical, mas a “entoação” restringe-se à análise, ao longo do enunciado, das variações de altura, ou seja, das sensações de grave e agudo. A oposição entre a sensação de tom grave *vs.* tom agudo em trechos de um enunciado é uma oposição que concerne o *pitch*, termo usado na literatura

científica em língua portuguesa para evitar a ambiguidade da palavra “altura”. Como o parâmetro físico que controla direta e primariamente a sensação de *pitch* é a frequência fundamental, isto é, o correlato acústico da frequência de vibração das pregas vocais, os estudos fonéticos da entoação privilegiam a análise da curva de frequência fundamental ao longo dos enunciados.

Nesse segundo recorte o ritmo faz parte da prosódia, mas é independente da entoação, uma vez que compreende as variações de duração percebida de unidades do tamanho da sílaba ao longo do enunciado. Como a duração objetiva de unidades do tamanho da sílaba (medida em milissegundos), é o parâmetro de controle por excelência da duração percebida, a análise do ritmo privilegia o estudo da duração.

5. Aspectos metodológicos ligados ao estudo da entoação e do ritmo

Geralmente quem estuda a entoação da fala analisa a estruturação de movimentos singulares ou salientes na curva de frequência fundamental ao longo dos enunciados representativos de uma instância comunicativa “congelada” num *corpus* de fala. Esses pontos singulares são de dois tipos: (1) proeminências expressas por picos e vales da frequência fundamental, chamados de acentos de *pitch* e (2) movimentos da mesma curva para níveis altos e baixos que marcam fronteiras entoacionais. Exemplo desses pontos singulares podem ser vistos na FIG. 2: proeminências assinaladas por picos de frequência fundamental nas palavras “eu” e “VERDE” e marcação de fronteiras entoacionais com movimento para um tom baixo após a palavra “não” e após a palavra “VERDE”.

Por outro lado, quem estuda o ritmo da fala analisa a estruturação de contornos de duração de unidades do tamanho da sílaba também nos enunciados de um *corpus* representativo de instância comunicativa. A FIG. 3 ilustra esse contorno duracional das sílabas fonéticas do enunciado “Manuel tinha entrado para o mosteiro há quase um ano, mas ainda não se adaptara àquela maneira de viver.”, realizado por falante paulista do sexo feminino de cerca de 30 anos na época da gravação.

da duração. Exemplos de efeitos microprosódicos na curva de frequência fundamental podem ser vistos na FIG. 2 no vale durante a palavra “não”, por conta da nasalização e nas interrupções na curva por conta do não vozeamento de consoantes, como no [t] de “moto”. Essas alterações na curva não são audíveis, pelo menos para um adulto. Essa forma de ouvir é aprendida durante a aquisição da linguagem.

Os efeitos microprosódicos do contorno duracional da FIG. 3 foram eliminados por meio de uma técnica de normalização da duração em duas etapas. Essa técnica permite preservar, com acuidade de cerca de 80%, aquilo que é percebido como proeminente pelos ouvintes, seja pela marcação de uma fronteira, seja pelo assinalamento de uma proeminência. Na primeira etapa da normalização, calcula-se um valor normalizado da duração da sílaba fonética que é padrão na Estatística, o *z-score*. Por definição, esse valor expressa o grau de afastamento de um valor em relação à média em unidades de desvio-padrão. Dizer, por exemplo, que a duração de uma sílaba tem valor de *z-score* igual a 1,3 quer dizer que a duração se encontra a 1,3 desvio-padrão à direita da média da duração. Os descritores de média e desvio-padrão podem ser obtidos de um *corpus* de fala que não precisa ser com a fala do mesmo informante (cf. BARBOSA, 2006, p. 489, para uma tabela de referência). Após essa etapa, a curva com valores de *z-score* de duração de sílaba fonética é suavizada por um procedimento de média ponderada conhecida como média móvel de 5 pontos (BARBOSA, 2010). O resultado da aplicação dessa técnica foi mostrado na FIG. 3. A curva duracional com a duração bruta, antes da normalização, é mostrada na FIG. 4. Há doze picos locais nessa curva, mas os sujeitos só ouvem até seis palavras como proeminentes ou pré-fronteiriças nesse enunciado. Cinco dessas palavras são reveladas pela curva normalizada da FIG. 3. A sexta palavra, “Manuel”, é assinalada para o ouvinte por um movimento de frequência fundamental pela informante. A normalização da duração é um procedimento padrão e básico da pesquisa sobre o ritmo da fala para revelar os trechos de um enunciado relevantes do ponto de vista prosódico. Observe ainda, pela proeminência estritamente entoacional da palavra “Manuel”, que os parâmetros fonético-prosódicos se conjugam para assinalar a proeminência para um ouvinte.

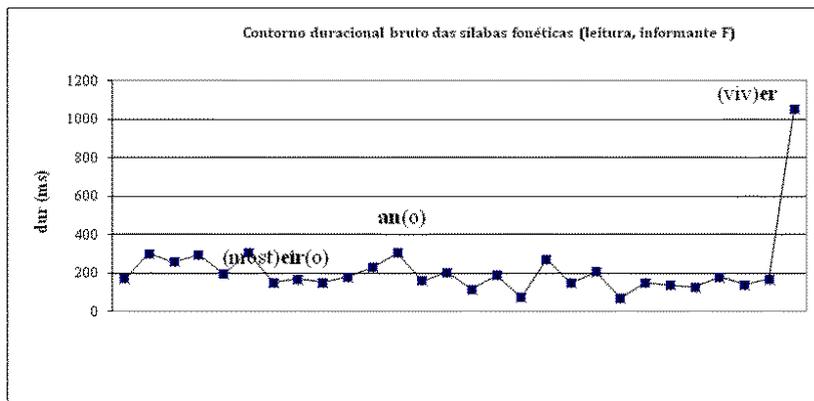


FIGURA 4 – Contorno duracional bruto das sílabas fonéticas do enunciado da FIG. 3

Auxiliada pela análise estatística inferencial, a análise fonético-prosódica busca regularidades nos padrões duracionais e melódicos que revelam o emprego do conhecimento prosódico de suas línguas por seus falantes. Esses estudos revelam que, em português brasileiro, as proeminências e as fronteiras prosódicas são marcadas por picos de duração e de frequência fundamental na vizinhança das sílabas tônicas salientes no enunciado.

Esses procedimentos metodológicos possibilitam definir, para além da sílaba, os grupos acentuais como unidades prosódicas fundamentais. Esses grupos são delimitados por sílabas salientes consecutivas no enunciado. Essa saliência pode se dar por proeminência ou marcação de fronteira prosódica. No enunciado ilustrado acima, “Manuel tinha entrado para o mosteiro há quase um ano, mas ainda não se adaptara àquela maneira de viver.”, as palavras “Manuel”, “entrado”, “mosteiro”, “ano”, “adaptara” e “viver” são salientes para ou ouvintes. Definem assim uma anacruse (do início do enunciado à tônica de “Manuel”) que é seguida de cinco grupos acentuais que se terminam na tônicas das palavras salientes seguintes. Graus distintos de saliência nessas sílabas assinalam níveis prosódicos distintos, passíveis de formarem constituintes prosódicos de domínios distintos. Basta observar na FIG. 3 os dois principais constituintes que terminam em “ano” e “viver”, além de constituintes de nível inferior delimitados pelas palavras “entrado”, “mosteiro” e “adaptara”. A análise estatística revela que o número desses domínios não ultrapassa quatro (BARBOSA, 2006). Os pontos salientes que delimitam os grupos acentuais são também denominados de

“acentos frasais” (*phrase stress*), embora a literatura os associe mais frequentemente à função estrita de proeminência. O acento frasal é, em português brasileiro, a indicação da proeminência de um elemento do enunciado, realizado em torno da sílaba tônica ou lexicalmente acentuada. Nem toda sílaba tônica é marcada como proeminente ao nível do enunciado por um ponto singular de duração ou frequência fundamental. Isso não significa que não haja marca fonético-acústica alguma do acento lexical de uma palavra não proeminente. A queda da intensidade total ou a mudança de qualidade da vogal da sílaba tônica para a pós-tônica frequentemente o assinalam. Para uma revisão da terminologia na área de prosódia consultar Botinis *et al.* (2001).

Relacionado ao termo de “acento frasal” encontram-se frequentemente na literatura prosódica os termos “foco” e “ênfase”. Há uma grande divergência a respeito de como definir esses termos, embora não se discuta que desempenham uma função de proeminência. O foco é normalmente entendido do ponto de vista fonológico ou, num maior grau de abstração, como uma marca num determinado nó sintático que assinala que o constituinte que a tem é “forte” em relação a outro, presente ou ausente. Pode também ser visto como uma marca semântica. É de qualquer forma uma marca discreta, opondo um constituinte focado a um não focado. A ênfase pode ser entendida como a realização desse foco na cadeia da fala como uma proeminência manifesta de uma unidade linguística com função de insistência ou para chamar a atenção para uma informação crucial, entre outras funções. O exemplo do enunciado sobre a moto verde, acima, ilustra um foco contrastivo em sua face discreta e realizada a partir de movimento da curva de frequência fundamental. Por essa última razão pode-se falar de ênfase na palavra “verde”. Por vezes esse termo só é empregado se o falante pronuncia o adjetivo, passando um significado adicional de insistência ou importância, mas a delimitação do limiar especificado pelo adjetivo “adicional”, além do qual há ênfase, não é determinável seguramente em cada caso.

6. A não congruência entre prosódia e sintaxe

Um dos aspectos mais importantes da relação entre prosódia e níveis linguísticos superiores à palavra fonológica diz respeito ao não isomorfismo ou não congruência entre sintaxe e prosódia. Isso significa que os constituintes prosódicos não coincidem com os constituintes sintáticos, como exemplificado pelos enunciados “Maria dança | divinamente bem. ||” e “O projeto | permite

avanços porque visa || à interação entre povos | de tribos diferentes. |||”, em que o primeiro constituinte prosódico de cada enunciado não corresponde a nenhum constituinte sintático.

Dizer que não há congruência não significa dizer que a formação de constituintes prosódicos prescinde do componente sintático ou mesmo do semântico. Significa dizer que sua construção depende de outros fatores além dos linguísticos. Tais fatores são restrições dos mecanismos de produção e percepção da fala que se relacionam com memória de trabalho no que tange a seus limites temporais e / ou espaciais. É de conhecimento psicolinguístico que planejamos nossos enunciados com quatro a seis palavras de antecedência (LEVELT, 1989). Esses limites correspondem, em média, a limites temporais da ordem de um a dois segundos, que, não coincidentemente, correspondem à extensão temporal média dos grupos acentuais. Uma das consequências das restrições na produção e percepção da fala é a tendência de produzir constituintes prosódicos com número de sílabas não muito díspares. Sob esse ângulo, a prosódia pode ser vista como um compromisso entre um sistema de produção e percepção sujeito a restrições de ordem cognitiva e um sistema linguístico sujeito a restrições linguísticas.

7. Da relação entre prosódia e segmentos acústicos menores do que a sílaba

A teoria do *Frame / Content*, de Macneilage (1998), supõe uma separação original da enunciação de unidades do tamanho da sílaba, que agem como máscaras (*frame*), da produção de unidades do tamanho do fonema (*content*). Vogais e consoantes seriam assim um subproduto de uma oscilação eminentemente silábica. Procurando explicar a própria evolução da linguagem falada a partir dessa separação fundamental, o autor pressupõe o primado da prosódia sobre o segmental. A prosódia seria mais primitiva, possivelmente derivada de emoções articuladas durante a oscilação mandibular, ainda por encontrar palavras para especificar seu conteúdo. Independentemente da corrente teórica, é um fato que a segmentação do enunciado, a proeminência e as marcas discursivas afetam a realização das unidades fônicas. Os elementos que constituem a sílaba tônica de uma palavra proeminente, por exemplo, são mais longos, têm valor de frequência fundamental mais elevado em sua vizinhança e têm valores de intensidade mais elevados que os dos segmentos vizinhos. A esses aspectos também se acrescentam valores mais periféricos ou extremos de

frequências de formantes, correlatos acústicos da hiperarticulação do trato vocal. Para exemplificar, um simples [s] extremamente longo e intenso pode constituir um enunciado com a função de fazer silêncio numa classe.

8. A prosódia no campo da percepção

À prosódia auditiva cabe o estudo das sensações de altura (*pitch*), duração (*duration*) e volume (*loudness*) através de testes de percepção empregados para confirmar ou não constituintes e proeminências prosódicos dos enunciados. A partir dessas técnicas, em que o ouvinte é submetido a uma série de testes de reconhecimento de contrastes prosódicos *in praesentia*, todo o campo da investigação da percepção da prosódia ainda está por se delinear e encontrar técnicas de experimentação apropriadas. Para tanto, o conhecimento metodológico da psicolinguística experimental será crucial. Vale lembrar que não há uma relação de 1:1 entre os parâmetros físicos e perceptivos, por duas razões. A primeira, porque a relação entre parâmetro físico e correlato perceptivo não é linear (e.g., a sensação de *pitch* vai se tornando menos discriminante à medida que a frequência fundamental sobe); a segunda, porque contribuem para uma dada sensação todos os parâmetros acústicos. Para o *pitch*, por exemplo, tanto a frequência fundamental quanto a intensidade e a duração objetiva são integradas para compor a sensação de altura. Essa relação complexa entre o físico e o sensorial se dá pelas características fisiológicas do sistema periférico auditivo, aliadas às características integrativas das vias auditivas superiores no cérebro.

Componentes marginais da prosódia, os estudos da variação do volume e do timbre não têm recebido atenção por parte dos pesquisadores, com raras exceções. O volume é o correlato perceptivo, no eixo forte / fraco, das variações de intensidade, seu parâmetro físico mais importante. O timbre é de mais difícil delimitação por conta de proximidade com a versificação e as acepções não científicas. Pode ser caracterizado por relações entre frequências dos formantes das vogais e consoantes ao longo do enunciado, com certa relação com as interpretações sonoras dos tratados de versificação (cf. GRAMMONT, 1965). No entanto, não há uma métrica aceita unanimemente para seu estudo. A qualidade de voz, por outro lado, ganhou seu espaço nos estudos prosódicos. Associada aos estados de fonação das pregas vocais, a qualidade de voz é correlata ou sofre as consequências do esforço vocal, da soproidade, do tônus vocal, entre outros aspectos do estudo prosódico da expressividade.

Referências

- BARBOSA, P. A. *Incurções em torno do ritmo da Fala*. Campinas: FAPESP/Pontes Editores, 2006.
- BARBOSA, P. A. Automatic duration-related salience detection in Brazilian Portuguese read and spontaneous speech In: *Proc. of the Speech Prosody 2010 Conference*, 2010, Chicago, 2010. 100067:1-4.
- BOLINGER, D. *Intonation and Its Parts: Melody in Spoken English*. Stanford: Stanford University Press, 1986.
- BOTINIS, A.; GRANSTRÖM, B.; MÖBIUS, B. (2001) Developments and paradigms in intonation research. *Speech Communication*, v. 33, p. 263-296, 2001.
- DUCROT, O.; TODOROV, T. *Dictionnaire encyclopédique des sciences du langage*. Paris: Éditions du Seuil, 1972.
- FÓNAGY, I. Les langages de l'émotion, *Quaterni di semantica*, 7/2, M. Alinei ed., Bolonha, p. 305-318, 1986.
- GRAMMONT, M. *Petit traité de versification française*. Paris: Armand Collin, 1965.
- HIRST, D.; DI CRISTO, A. *Intonation systems: a survey of twenty languages*. Cambridge: Cambridge University Press, 1998.
- LÉON, P. R. *Essais de phonostylistique*. Ottawa: Didier, 1971.
- LEVELT, W. J. M. *Speaking: from Intention to Articulation*. Cambridge, MA: M.I.T. Press, 1989.
- LIDDEL, H. G.; SCOTT, R.; JONES, H. S. *A Greek-English lexicon*. Oxford: Clarendon, 1996.
- MACNEILAGE, P. F. The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*. 21, 499-511, 1998.
- MARTINET, A. *Éléments de linguistique générale*. 3. ed. Paris: Armand Collin, 1991.
- SCHERER, K. R. On the nature and function of emotion: a component process approach. In SCHERER, K. R. EKMAN, P. (Ed.). *Approaches to emotion*. Hillsdale, NJ: Lawrence Erlbaum. 1984. p. 293-318.