

Universidade Federal de Juiz de Fora
Pós-Graduação em Linguística
Mestrado em Linguística

Patrícia Ribeiro do Valle Coutinho

COGNIÇÃO E PROSÓDIA:
O EMBATE PELA PALAVRA NA COLETIVA DE IMPRENSA

Juiz de Fora

2010

Patrícia Ribeiro do Valle Coutinho

Cognição e prosódia:
O embate pela palavra na coletiva de imprensa

Dissertação apresentada ao
Programa de Pós-Graduação em
Linguística da Faculdade de Letras
da Universidade Federal de Juiz de
Fora, como requisito parcial para
obtenção do título de Mestre em
Linguística.

Orientador: Prof. Dr. Luiz Fernando Matos Rocha

Juiz de Fora

2010

Coutinho, Patrícia Ribeiro do Valle.

Cognição e prosódia: o embate pela palavra na coletiva de imprensa / Patrícia Ribeiro do Valle Coutinho. – 2010.
165 f. : il.

Dissertação (Mestrado em Linguística)–Universidade Federal de Juiz de Fora, Juiz de Fora, 2010.

1. Linguística. 2. Cognição. 3. Prosódia. 4. Metáfora. I. Título.

CDU 801

Patrícia Ribeiro do Valle Coutinho

Cognição e prosódia:
O embate pela palavra na coletiva de imprensa

Dissertação apresentada ao Programa de Pós-Graduação em Linguística da Faculdade de Letras da Universidade Federal de Juiz de Fora, como requisito parcial para obtenção do título de Mestre em Linguística.

Aprovada em: 15/03/2010.

BANCA EXAMINADORA

Prof. Dr. Luiz Fernando Matos Rocha (Orientador)
Universidade Federal de Juiz de Fora

Profª. Dra. Heliana Ribeiro de Mello
Universidade Federal de Minas Gerais

Profª. Dra. Luciana Teixeira
Universidade Federal de Juiz de Fora

Aos meus pais e irmãos, maior motivação não há.

AGRADECIMENTOS

Ao cumprir a tarefa de escrever essa dissertação, encerrando um *turno* e dando início a outro, reflito sobre o que estive *emparelhado* nessa *construção*. Certamente, *em conjunto* com meu esforço e dedicação, foram decisivas as participações de algumas pessoas. E são a elas que destino os meus agradecimentos.

Minha gratidão total e incondicional ao professor e orientador Luiz Fernando Matos Rocha, pela acolhida e confiança, pela amizade e generosidade e por me mostrar que uma orientação pode significar crescimento.

Aos professores do Programa de Pós-Graduação em Linguística da UFJF, por contribuírem para que o mestrado tenha feito toda a diferença na minha vida acadêmica. Às professoras Heliana Mello e Luciana Teixeira, por aceitarem compor a banca da minha defesa. A Luciana Ávila, sempre presente nos momentos importantes. À secretária Rosângela, por ser tão prestativa e participante nos meus passos durante essa etapa.

Aos colegas do mestrado, em especial, Gabriela, pelo companheirismo, pela torcida, e por me ajudar a compreender a união de um *grupo*. Agradeço também a Simone e Ana Eliza, por dividirem comigo as angústias de um período delicado.

Minha gratidão também a Luciana Lucente, por ser tão receptiva às minhas dúvidas e anseios com relação a se fazer a análise acústica do meu trabalho.

Obrigada às minhas sempre carinhosas tias Marilda, Isa e Celina e aos amigos que, de alguma forma, me fortalecem: Fabiana, Paula, Lívia, Ronaldo, Luciene, Raphaela, Mariana, Laís, Monique e Renata.

E finalizo com o princípio de tudo: Deus, obrigada por aumentar cada vez mais minha fé.

As sílabas

Luiz Tatit

Cantiga diga lá
A dica de cantar
O dom que o canto tem
Que tem que ter se quer encantar
Só que as sílabas se embalam
Como sons que se rebelam
Que se embolam numa fila
E se acumulam numa bola
Tem sílabas contínuas:
Ia indo ao Piauí
Tem sílabas que pulam:
Vox populi
Tem sílaba que escapa
Que despenca
Rola a escada
E no caminho
Só se ouve
Aquele boi-bumbá
Tem sílaba de ar
Que sopra sai o sopro
E o som não sai
Tem sílaba com esse
Não sobe não desce
Tem sílaba legal
Consoante com vogal
Tem sílaba que leve oscila
E cai como uma luva na canção

RESUMO

Prosódia e cognição são os temas centrais do presente trabalho. Reivindica-se o lugar da fonologia na representação formal de construções gramaticais, uma vez que emparelhadas com informações semânticas, sintáticas e pragmáticas estão as informações fonológicas. A partir de contribuições teóricas da Linguística Cognitiva, o conceito de fonema é revisto, sendo tratado em termos de categoria de base corpórea e unidade construcional. Para essa nova visão, é fundamental, ao tratar do fenômeno, uma concepção que envolva a aquisição, o uso e a frequência de construções da língua. Ao se direcionar a atenção para o nível suprasegmental, enfatiza-se a função que a prosódia desempenha na interpretação e na negociação de sentidos. Defende-se que há uma ação conjunta e coordenada na produção da melodia da fala. A fim de analisar como a entoação pode apontar para a construção do significado, foi realizada, através de um programa computacional (PRAAT), a leitura acústica de dados de fala semiespontânea. Os resultados da submissão dos dados foram descritos no sentido de se observar as curvas da Frequência Fundamental (F0) nos trechos que encerram as respostas de entrevistados em coletivas de imprensa. A hipótese é a de que os contornos melódicos nessa transição de turno são descendentes e uma das razões para isso é o desgaste da pressão subglotal que ocorre no fluxo da fala. Outro ponto discutido é que a ação das pregas vocais pode ser considerada um caso particular entre as gestualidades produzidas pelo nosso aparato físico e perceptual. Estabelecendo uma projeção metafórica entre o domínio da coletiva de imprensa e o domínio da guerra, aposta-se na analogia entre os gestos produzidos pelas pregas vocais e os gestos dos braços em um combate. Enfim, quanto mais gestos “sonoros”, de alta F0, ocorrem, maior é o poder de combate interacional do falante, na luta pelo turno. Se a F0 decai, dá-se a entrega do turno. São movimentos de caráter físico que estão estreitamente relacionados com o simbolismo conceptual da linguagem.

Palavras-chave: Linguística Cognitiva. Prosódia. Categorização. Unidades construcionais. Ação coordenada. Coletiva de imprensa. Troca de turnos. Frequência fundamental. Metáfora.

ABSTRACT

Prosody and cognition are the central subjects of the present work. The place of phonology in the formal representation of constructions is demanded, since phonological information is paired with semantic, syntactic and pragmatic information. From theoretical contributions of Cognitive Linguistic, the concept of phoneme is reviewed and reformulated with the notions of phoneme as an embodied category and phoneme as a constructional unit. For this new vision, it is fundamental a conception that involves the acquisition, the use and the frequency of language. To direct the attention for the suprasegmental level, it is emphasized the function that the prosody performs in the interpretation and management of meaning. One defends that there is a joint and coordinated action in the production of speech's melody. In order to analyze how the intonation can point out to the construction of meaning, an acoustics reading of data was accomplished through a computation program (PRAAT). The results of the submission of the data were described in order to observe the curves of fundamental frequency (F0) in the stretches that close the interviewed answers in the press conference. The work hypothesis is that the turn transition is configured by falling contours and one of the reasons for this is the deterioration of the subglotal pressure in the flow of the speech. Another point is that the action of vocal folds can be considered a particular case between the gestures produced by our physical and perceptual apparatus. From a metaphorical extension of the press conference's domain and the war's domain, it can be speculated the analogy between the gestures produced by the focal folds and the gestures of the arms in a combat.

Keywords: Cognitive Linguistics. Prosody. Categorization. Constructional units. Coordinated action. Press conference. Turn Exchange. Fundamental frequency. Metaphor.

LISTA DE ILUSTRAÇÕES

Imagem 1	Mulher pronunciando “både” (ambos), em dinamarquês.....	58
Imagem 2	O brigadeiro encerrando sua resposta à primeira pergunta dos jornalistas.....	74
Imagem 3	O jogador Daniel sorrindo quando brinca com os jornalistas sobre treinar com o pé direito, já que ele e os demais jogadores do quadrado mágico são canhotos.....	95
Imagem 4	O jogador Álvaro encerrando sua resposta à primeira pergunta dos jornalistas.....	96
Imagem 5	Ivete respondendo a pergunta dos jornalistas sobre a crise na economia brasileira.....	106
Foto 1	Golpes corporais análogos aos “golpes” sonoros.....	92

LISTA DE GRÁFICOS

Gráfico 0 Modelo.....	68
Gráfico 1 Fragmento A.1: “para que nós possamos identificar se há algum indício de sobrevivente”	76
Gráfico 2 Trecho: “se há algum indício de sobrevivente.....	76
Gráfico 3 Fragmento A.1: “para que nós possamos identificar se há algum indício de sobrevivente” (linha traceja).....	77
Gráfico 4 Trecho: “declarar o luto [mas] na minha na minha opinião”.....	80
Gráfico 5 Fragmento A.2: eu acho que é é é da sua competência declarar o luto [“mas” (sobreposição dos jornalistas)] na minha na minha opinião.....	81
Gráfico 6 Fragmento A.3: então nós levamos três horas para localizar os destroços em função do que foi acompanhado pelos radares.....	83
Gráfico 7 Trecho: “do que foi acompanhado pelos radares”	84
Gráfico 8 Fragmento A.4: “para que possamos trazer os corpos para o DF” (linha tracejada).....	86
Gráfico 9 Trecho: corpos para o DF	86
Gráfico 10 Fragmento A.5: “vai ser em função do da quantidade de corpos que foram que forem encontrados no local”	89
Gráfico 11 Trecho: “vai ser em função do da quantidade de corpos que foram que forem encontrados no local”	90
Gráfico 12 Fragmento B.1: “aí vai ser outra história outra novela e a gente vai esperar pra deixar pra dezembro tu resolver”	99
Gráfico 13 Trecho: “deixar pra dezembro tu resolver”	99
Gráfico 14 Fragmento B.1: “agora eu vou ter a oportunidade de de acompanhar de jogar e de viver o campeonato brasileiro que que esse era o meu desejo”...100	
Gráfico 15 Trecho: “esse era o meu desejo”	101
Gráfico 16 Fragmento B.3: “me encontro em em em condições de fazer um trabalho forte essas semanas e e e pra entrar em condições de jogo o mais rápido possível”.....	102
Gráfico 17 Trecho: “mais rápido possível”	103
Gráfico 18 Trecho: “isso é meio incitado”	108
Gráfico 19 Fragmento C.1: “o respeito mútuo que temos uns com os outros a	

prova disso é que como ela falou com a Bahia ninguém pode”	109
Gráfico 20 Trecho: “com a Bahia ninguém pode”	110
Gráfico 21 Fragmento C.2: “se você fizer a conta da passagem do tecido com as lantijola e as pedra as costura as agulha sai tudo em conta meu bem”	111
Gráfico 22 Trecho: “sai tudo em conta meu bem”	112
Gráfico 23 Fragmento C.3: “a relação do público de fora é muito intensa com o Brasil é maior do que a gente imagina do que a consciência da gente muito maior”	113
Gráfico 24 Trecho: “do que a consciência da gente muito maior”	114
Gráfico 25 Fragmento C.4: “sempre estarei pronta pra desfrutar da companhia deles porque é um prazer será um prazer”	115
Gráfico 26 Trecho: “porque é um prazer será um prazer”	116

SUMÁRIO

1	INTRODUÇÃO	12
2.	PRESSUPOSTOS TEÓRICOS	15
2.1	Fonologia Cognitiva: estado da arte.....	15
2.2	Bases corpóreas do fonema	24
2.3	Fonema em e como construção.....	28
2.4	Prosódia: percurso e relevância	40
2.5	Sistemas dinâmicos e cognição prosódica.....	48
3	PROCEDIMENTOS METODOLÓGICOS	60
3.1	A cena interacional e a troca de turnos: “Quando um burro fala, o outro abaixa a orelha”.....	61
3.2	Contexto de pesquisa: O gênero coletiva de imprensa.....	64
3.3	Materiais e procedimentos.....	66
3.4	Sistema de notação.....	69
4	ANÁLISE DOS DADOS	71
5	CONSIDERAÇÕES FINAIS	118
	REFERÊNCIAS	120

1. INTRODUÇÃO

A missão que este trabalho se destina a cumprir não é a das mais fáceis, mas as paisagens são belas. Nesse percurso, nós nos questionamos: como fazer um trabalho em Linguística Cognitiva que contemple os estudos sobre prosódia? Quais serão os autores cognitivistas a partir dos quais nos guiaremos para estudar entoação? Nada como expor as próprias angústias para nos fazermos claros quanto aos primeiros passos dados nessa empreitada. E foi assim que começamos a descobrir as belezas de se fazer o imprevisível. O argumento mais forte ao se justificar este trabalho é o de agregar idéias, conceitos, teorias e objetos para se tentar criar o novo.

Nossa intenção é a de avançar nos estudos em Linguística Cognitiva, usufruindo o que essa abordagem teórica postula, e trazer à tona um objeto de estudo ainda pouco focalizado por esse recente campo de investigação linguística: a melodia da fala. Acreditamos que a prosódia contribui, de forma decisiva, para a produção e interpretação dos sentidos, integrada construcionalmente com a semântica, sintaxe e pragmática. Dessa maneira, preocupar-se com semiose da linguagem implica preocupar-se com a prosódia da fala.

O objetivo geral desta dissertação é reivindicar a inserção de unidades construcionais prosódicas no emparelhamento forma/sentido. Nosso recorte analítico se dirige a construções que encerram a resposta de entrevistados em coletivas de imprensa, ou seja, em construções que se configuram como lugar relevante de transição de turno, o momento a partir do qual os jornalistas lançam novas perguntas. Uma questão importante é: o que faz os entrevistadores entenderem que a fala do entrevistado se encerrou, ou está se encerrando? Além disso, refletimos sobre o papel do nosso corpo como responsável e também limitador no processo de produção e percepção da fala. Pensaremos, por exemplo, na ação das pregas vocais como um caso particular entre as gestualidades produzidas pelo nosso aparato físico e perceptual; e discutiremos como, a partir daí, se criam unidades simbólicas sonoras.

Sabemos que o tempo e o espaço reservados para este trabalho nos restringem muito quanto às considerações geradas. Não temos a intenção de tecer comentários conclusivos sobre o assunto e, sobretudo, sabemos que esta dissertação se trata, ainda, de um estudo prospectivo.

No segundo capítulo, detemo-nos a expor os pressupostos teóricos a partir dos quais construímos o trabalho. Na seção 2.1, fazemos uma breve revisão da agenda de

estudos cognitivistas com o intuito de situar a Fonologia Cognitiva no panorama da teoria linguística emergente. Localizamos quem são os autores que já tiveram ou têm o nível fonológico como objeto de estudo. Além disso, chamamos a atenção para como a nova perspectiva inova no tratamento da fonologia.

A seção 2.2 inicia a discussão sobre a nossa experiência corporal na produção dos sons segmentais. Outro ponto questionado é a noção de prototipicidade na categorização dos fonemas da língua. Assumiremos que o corpo tem preferências naturais e que isso interfere nas ações coordenadas e conjuntas que resultam na nossa fala. Ademais, bem como uma licença física, postulamos que há uma motivação cultural na categorização dos sons.

Uma vez que falamos em ações conjuntas, anunciamos, na seção 2.3, nossa proposta de encararmos o fonema através de uma perspectiva construcional. As noções de uso, frequência e aquisição da linguagem se fazem presentes e fundamentais nesse momento. Além disso, usamos a teoria da Gramática das Construções para postular a existência do pareamento entre informações semânticas, sintáticas, pragmáticas e fonológicas na construção de unidades simbólicas.

Na seção 2.4, traçamos um panorama em torno da prosódia, principalmente no que diz respeito à sua função. Reconhecemos alguns dos principais autores da área e descrevemos os parâmetros prosódicos (e seus correlatos fisiológicos e perceptuais) mais reconhecidos e estudados. Encerrando essa seção, apresentamos resultados interessantes de estudos envolvendo o simbolismo sonoro, resultados esses que nos possibilitaram avançar em algumas de nossas considerações finais.

Por reconhecermos, de certa maneira, um diálogo entre algumas ideias da Linguística Cognitiva e o que a Teoria dos Sistemas Dinâmicos afirma sobre o funcionamento da linguagem, a seção 2.5 aborda temas comuns a ambas as vertentes teóricas. Tais temas integram a linguagem com aspectos sociais, psicológicos, biológicos e culturais. Ao comentarmos sobre bases neurológicas que licenciam o ser humano a simular ações, discutimos, de modo sucinto, sobre o papel das metáforas na nossa experiência e como elas nos ajudam a explicar nossos dados.

O capítulo metodológico conta com quatro seções. Na 3.1, um pouco é falado sobre a troca de turnos em uma cena interacional. Esse breve passeio por questões como cenários conversacionais é relevante, pois eles têm ligação direta com o que vamos analisar.

A seção 3.2 explica o gênero discursivo escolhido para análise. Descrevemos o contexto de cada uma das quatro entrevistas utilizadas, em virtude de que tais informações constituem um dos nossos focos na análise.

Em 3.3, relatamos o percurso metodológico de confecção do trabalho, desde a proposta inicial do projeto da dissertação até as etapas percorridas no manuseio do programa computacional utilizado para análise acústica.

Fechando nossas bases metodológicas, na seção 3.4, apresentamos o sistema de notação entoacional escolhido para transcrever as curvas resultantes da submissão dos dados.

O quarto capítulo, a análise dos dados, apresenta e comenta os gráficos gerados pelo programa de leitura acústica. As seções A, B e C são dispostas conforme o contexto das entrevistas, a saber, “Entrevistado *versus* entrevistador”, “Entrevistado *e* entrevistador” e “Entrevistado ‘*entrevista*’ entrevistador”, respectivamente. Os questionamentos que guiam a análise nos cobram respostas que atendam a aspectos semântico-interacionais, prosódicos e cognitivos envolvidos na cena analisada, especificamente no momento percebido pelos jornalistas como oportuno para se tomar o turno.

Enfim, após introduzirmos o que consistem as belas paisagens que encontramos no nosso, a princípio, distante caminho, partiremos para o primeiro capítulo do trabalho.

2. PRESSUPOSTOS TEÓRICOS

2.1 Fonologia Cognitiva: estado da arte

Quando se reivindica a legitimidade de um objeto exaustivamente estudado, porém ainda muito subfocalizado no domínio de uma teoria emergente, torna-se imprescindível expor, logo de início, o limite dos reais avanços obtidos até o momento por essa teoria em relação a tal objeto. Com isso, já se reconhece de antemão que há muito por fazer, porquanto se vislumbram possibilidades inéditas de contribuição para os estudos na área a qual se anuncia. Tais reflexões generalizam o que os recentes trabalhos em Linguística Cognitiva (doravante LC) têm feito para a fonologia ao criarem o que vem se denominando de Fonologia Cognitiva (doravante FC), abordagem teórica que concebe os sons da fala como constructos de base conceptual e corpórea. Então, qual justificativa a Gramática Cognitiva apresenta para direcionar seu olhar aos estudos fonológicos?

Essas reflexões são pertinentes levando-se em conta o carro-chefe que guiou os estudos em LC desde seu início: a preocupação com a produção e a interpretação do significado. A inquietação com a questão semântica se pautava na revisão da ideia de o sentido ser uma propriedade intrínseca da linguagem e ser equivalente às condições de verificação em um modelo no mundo real. Rejeitou-se a separação entre conhecimento linguístico e conhecimento extra-linguístico em prol de um entendimento de linguagem numa perspectiva de continuidade essencial entre os níveis microlinguísticos, a saber: semântica, pragmática, sintaxe, morfologia e fonologia. Silva (2007) elenca aspectos da construção do significado sob a ótica da LC, segundo a qual o significado é:

- (i) *perspectivista*, uma vez que não reflete objetiva e diretamente o mundo, mas modela-o, constrói-o através de perspectivas diferentes;
- (ii) *enciclopédico*, pois se associa ao conhecimento de mundo, não sendo autônomo e separado de outras capacidades cognitivas;
- (iii) *flexível*, visto que é dinâmico e adaptável às mudanças do mundo e das circunstâncias; e
- (iv) *baseado no uso*, isto é, na nossa experiência individual corpórea ou biológica e na nossa experiência social e cultural.

Como sustenta Fauconnier (1997), as expressões linguísticas passaram a ser entendidas como instruções parciais e indeterminadas para a construção de domínios interligados. Enfim, a primazia da semântica decorre da própria perspectiva cognitivista: se a função primária da linguagem é a de categorização, então a significação será o fenômeno linguístico primário (SILVA, 1997). Os princípios que tradicionalmente guiam o paradigma nomeado como Linguística Cognitiva são:

- (i) Escassez do significante;
- (ii) Semiologização do contexto;
- (iii) Drama das representações (SALOMÃO, 1999).

O princípio (i) é postulado em virtude de ser o significante insuficiente para abarcar os significados, funcionando a pista léxico-semântica como guia para a construção de sentidos. A significação é atingida quando o falante percebe as outras semioses que enriquecem a forma. Conforme o princípio (ii), no evento comunicativo, há uma continuidade entre linguagem, conhecimento e realidade. Por fim, o que o princípio (iii) sugere é que, assim como representar é interpretar um papel, na cena comunicativa exercemos um papel comunicativo, assumindo determinadas perspectivas.

Com o avanço dos estudos na área, notou-se que era preciso buscar respostas para a contraparte formal da linguagem, não mais em função de um significado isolado, mas integrado ao contexto de enunciação. Evidentemente, não se abandona a noção de estrutura, porém em vez de se primar por uma estrutura fixa, entende-se o sistema linguístico como algo plástico e flexível. Como não se produz significado linguístico sem forma, seria mesmo necessário propor um aporte teórico condizente com as premissas de que:

- (i) A linguagem não é uma faculdade cognitiva autônoma;
- (ii) A gramática é conceptualização;
- (iii) O conhecimento da linguagem emerge do uso da linguagem (CROFT; CRUSE, 2004).

Assim, seguindo o percurso traçado pelos linguistas cognitivistas de segunda geração (sendo a primeira composta pelos seguidores do gerativismo), é importante destacar, nesse breve panorama, a resposta sintática da LC, que atenta para unidades linguísticas complexas, cuja relação forma/sentido é imprevisível a partir de simples

cálculos combinatórios, bem como para o fenômeno da variação das estruturas argumentais de uma predicação (SALOMÃO, 2006). Visões como essas possibilitaram o despontar da Gramática das Construções (doravante GC) (FILLMORE; KAY. O'CONNOR, 1988; GOLDBERG, 1995; CROFT, 2001), reação ao modelo estritamente composicional que tentava, de maneira não plenamente satisfatória, dar conta das expressões linguísticas.

Assumimos então que construções, do ponto de vista da LC, são emparelhamentos dos pólos forma e sentido, tradicional e respectivamente estabelecidos entre sintaxe e semântica. No decorrer deste trabalho, advogamos em favor de uma inserção mais vigorosa da fonologia no pólo formal da construção. As representações e notações feitas pelos estudiosos em GC ainda não incluem informação fonológica. Defendemos a ideia de que é relevante em uma análise mono-estratal, isto é, que considera o estrato das construções e não a derivação, a inserção de informações fonológicas e, especialmente, prosódicas. Estas últimas estabelecem a ponte entre fonética e fonologia, nos termos de Barbosa (2006). Mais adiante falaremos, especificadamente, do paradigma construcionista, uma vez que este está diretamente atrelado à nossa proposta de trabalho.

Com o tratamento da fonologia em LC, a dialética não foi diferente: assim como nos outros níveis da língua (sintaxe e semântica, primordialmente), fazer fonologia no panorama da Gramática Cognitiva implicava de início uma revisão radical da Gramática Gerativa. Para entender como isso funcionou, recuperamos um texto de Lakoff (1993), cujas problematizações provavelmente abriram o caminho para o surgimento de uma nova reflexão acerca de fonologia. No trabalho intitulado *Cognitive Phonology*, o autor sugere uma considerável revisão dos postulados gerativistas. Para ele, não são convincentes os estágios derivacionais os quais, segundo a teoria gerativa, os falantes percorrem passo a passo ao pronunciarem palavras e sentenças. É questionado, pois, se tais estágios intermediários são cognitivamente reais quando se observa o desempenho do falante e não sua competência. Lakoff sugere a oposição entre um tempo real e um tempo abstrato, sendo que as derivações não seriam colocadas em correspondência com o primeiro.

Lakoff propõe um modo alternativo a fim de simplificar a fonologia, mais especificamente a Fonologia Gerativa. Para o autor, todo o ordenamento rigoroso de regras gerativistas (em que uma regra interage com outra, uma regra antecede outra etc.) pode ser deixado de lado, se adotada uma nova proposta de níveis de representação. Tal

proposta baseia-se na correspondência entre morfemas – como armazenados na mente – (nível morfêmico) e sequências fonéticas (nível fonético). E, ainda, o modelo se atenta para o nível da palavra. O autor sugere que há restrições para condições de boa formação no nível fonético e nossa interpretação é a de que tais restrições são ditadas pelo sistema de coordenação corpórea (o que será mais desenvolvido adiante).

O objetivo geral do texto de Lakoff é o de mostrar como derivações, guiadas por regras, ciclos e princípios – tal como era feito na Gramática Gerativa – podem ser eliminadas, pelo menos em alguns casos. Para tanto, uma contribuição importante é a concepção de gramática caracterizando correlações simultâneas entre dimensões fonológicas e aspectos da sintaxe, semântica e pragmática. Percebemos, nesse texto inicial, uma abertura para a mono-estratagem da gramática, que é um ponto central e de consenso entre os modelos da GC.

Podemos, então, inferir que naquele momento já se negavam as regras gramaticais e se apostava em uma perspectiva construcionista. Assim como a sintaxe era explorada na GC, a causa da fonologia seria abraçada também por um modelo cognitivo de gramática. No entanto, assumia-se que as possibilidades da FC somente começavam a ser investigadas.

Ao mapearmos a Gramática Funcional (em sentido amplo, incluindo a LC), constatamos que dois ramos de estudos relacionados à fonologia vêm sendo desenvolvidos: o primeiro vê a fonologia como representação do nosso conhecimento da experiência corpórea (NATHAN, 1986), e o segundo se expande nos modelos baseados no uso (BYBEE, 1994). E é sobre eles que seguiremos dissertando.

Um pesquisador que vem se dedicando ao estudo da FC é Geoffrey S. Nathan (1986, 2006, 2007, 2008). No seu livro intitulado *Phonology: a cognitive Grammar introduction*, Nathan (2008) explica que a fonologia é o estudo da organização e estrutura dos sons da língua e, como a maioria das áreas da gramática, lida com princípios universais e específicos. Trata do que a língua faz com esses sons – como certos sons são selecionados, como esses sons são ajustados a seus ambientes e como são construídos dentro de unidades maiores como as sílabas, os pés, as palavras etc. A fonética, por outro lado, lida com todos os sons possíveis da fala humana – um inventário de possibilidades, definidas pelo aparato vocal, ou o sistema perceptual humano.

Para tornar mais claras as definições e diferenciações, o autor faz uma analogia com a arquitetura e os materiais de construção. Para ele, a fonética é como o estudo da

natureza dos materiais de construção; enquanto a fonologia é o estudo do modo pelo qual esses materiais são realmente usados para construir prédios. Afirma-se que há mais na arquitetura do que simplesmente uma lista de materiais. Um arquiteto tem que entender das propriedades estruturais dos materiais de construção (quais podem suportar peso, quais resistem ao tempo) a fim de determinar quais deles fariam um bom teto ou uma boa parede. Fonética e fonologia interagem sob muitos aspectos tratando a última dos princípios que a língua usa para tomar os materiais brutos e construir palavras (NATHAN, 2008).

A preocupação central de Nathan é construir uma teoria, no que diz respeito à fonologia, capaz de explicar como os falantes processam a linguagem, tanto consciente quanto inconscientemente. Ele destaca que, ainda hoje, muitos dos constructos teóricos surgidos desde o início dos estudos em fonologia continuam a vigorar mesmo com tantas propostas de revisões. Para ilustrar como ainda é forte a influência do legado estruturalista ao longo dos tempos, o autor faz um breve apanhado da história do fonema, desde as primeiras formulações do seu conceito.

Uma atenção especial é dada a Baudouin de Courtenay (1972), considerado o primeiro fonologista sincrônico, cuja visão era de que o fonema é visto como um som ideal; a fusão, na mente, de impressões que esta recebe das diferentes pronúncias de um mesmo som (na realidade, tais concepções são datadas aproximadamente de 1895). Trata-se de um ponto de vista psicológico, sendo o fonema o equivalente psíquico dos sons que falantes sistematicamente realizam no processo da fala. Baudouin recebeu críticas às suas postulações, pois se alegava que o interesse do autor era demasiadamente psíquico e pouco linguístico. Tal ataque vinha dos linguistas da Escola de Praga como Trubetzkoy e Jakobson. A partir de 1930, nos trabalhos do Círculo Linguístico, o conceito de fonema é formulado enquanto unidade mínima distintiva do sistema de sons. Mais tarde (década de 60) o conceito é reformulado por Jakobson com a proposta de fonema como feixe de traços distintivos.

Para o estruturalismo, a língua é um sistema auto-suficiente. O resultado dessa visão autônoma era a concepção de fonema distante do conhecimento do falante ou o ouvinte (NATHAN, 2006). Em outras palavras, para os estruturalistas pouca relevância tinham os aspectos extra-linguísticos quando o assunto era fonologia.

Outro autor citado no rol de reflexões acerca do fonema, e que se alia às ideias de Baudouin, é Sapir (1972), para quem falantes não usam a metodologia de pares mínimos para aprender a língua. Trata-se de uma posição destoante da Fonologia

Gerativa que permanece, essencialmente, com a definição contrastiva dos estruturalistas. Taylor (2006), da mesma maneira, problematiza a definição de fonema a partir da ideia de pares mínimos, pois alega que sua existência dependerá do tamanho do léxico de uma pessoa. Por isso, ele não acredita que crianças aprendam com base em pares mínimos. O autor ainda enfatiza que muitas coisas que as crianças dizem são criações únicas, nunca antes encontradas, mas não são, por essa razão, agramaticais.

A Fonologia Natural – uma vertente não-gerativa de bastante alcance nos anos 70 e 80, mas que desde então pouco tem sido explorada – defende que fonemas são imagens mentais de sons; e adaptando-se aos achados de Sapir e Baudouin, enfatiza que a fala real diverge desses sons armazenados. Assim, fonemas são imagens mentais de sons que falantes modificam na fala e que, ao perceber a fala do outro, os mesmos falantes percebem o *output* “deformado” como o que eles teriam pretendido dizer. Isto é, as “deformações” são adaptações mentais causadas pela natureza e pelas limitações do trato vocal e do sistema perceptual, que são universais, mas aprendidas no mesmo sentido de que crianças aprendem a controlar suas mãos e seus pés (NATHAN, 2007). Tratando-se da percepção, ouvintes participam de um processo chamado *sympathetic reconstruction*, o qual poderia ser parafraseado como algo assim: “Ele ou ela disse X, que soou como se eu tivesse pretendido dizer Y” (NATHAN, 2006, p. 4). Ou seja, para ele, é central a noção de intenção por trás de todo o processamento fonológico, noção essa inaugurada por Baudouin.

Segundo Nathan (2006), há uma afinidade entre a Gramática Cognitiva e a Fonologia Natural, cujo principal líder é Stampe (1969). O que une ambas as abordagens é a defesa de uma relação entre estruturas linguísticas e aspectos não-linguísticos da cognição humana. Conforme a Fonologia Natural, os sistemas de sons das línguas são regidos por forças implícitas na vocalização e na percepção humanas:

Na versão moderna da teoria, as forças fonéticas implícitas manifestam-se através de processos, na interpretação de Sapir – substituições mentais que sistematicamente, embora subconscientemente, adaptam as nossas intenções fonológicas às nossas capacidades fonéticas e que, inversamente, nos permitem a compreensão nos outros falantes, das intenções subjacentes a estas adaptações fonéticas superficiais. (SOUZA e CAGLIARI, 2001, p. 3-4)

Essa vertente da Fonologia concebe a linguagem como um reflexo natural das necessidades, das capacidades e do universo dos falantes.

Situando, portanto, o fonema no modelo da Gramática Cognitiva, percebe-se a insistência em uma base psicológica para todas as unidades linguísticas, isto é, não há princípios estritamente linguísticos como um todo. Uma das possibilidades de se olhar para o fonema é encarando-o como uma representação de imagens mentais acústicas e articulatórias.

Os fonemas, na nova roupagem recebida dentro da Gramática Cognitiva, são categorias de nível básico. Enfatiza-se a habilidade de humanos categorizarem o mundo e questiona-se qual a natureza dessa categorização. A razão de falantes terem a necessidade de armazenar a fala em termos de um pequeno número de unidades de percepção/produção do som pode ser encontrada na noção original de categorias de nível básico, como proposta por Rosh:

(...) two contrasting principles influence the taxonomic level which people prefer to categorize. The first is the efficiency principle, according to which people prefer to minimize the number of categories they must consider in making a categorization (...) The second principle is the informativeness principle, according to which people tend to maximize the informativeness of their categorizations. (ZBIKOWSKI, 2002 *apud* NATHAN, 2006, p. 14)

(...) dois princípios contrastantes influenciam o nível taxonômico pelo qual as pessoas preferem categorizar. O primeiro é o princípio da eficiência, segundo o qual as pessoas preferem minimizar o número de categorias a serem consideradas na categorização (...) O segundo princípio é o princípio da informatividade, de acordo com o qual as pessoas tendem a maximizar a informatividade de suas categorizações. (ZBIKOWSKI, 2002 *apud* NATHAN, 2006, p. 14, tradução nossa).

Será, então, a partir das instâncias armazenadas que os fonemas poderão ser categorizados, ou seja, a grande contribuição da Gramática Cognitiva para a fonologia é a de que a identificação de fonemas é uma questão de categorização.

Nos termos da Gramática Cognitiva, fonemas são unidades de nível básico e estão facilmente disponíveis à consciência do falante, enquanto os traços são classificações mais abstratas não diretamente acessíveis ao ouvido ou à boca destreinados. É importante salientar, contudo, que nessa visão não se negam os traços fonéticos. Embora falantes possam não estar conscientes dos traços, eles são linguisticamente significativos. Traços representam gestos físicos reais ou pelo menos equivalentes perceptuais deles.

Para entendermos como se dá a categorização dos sons, é importante deixar claro que fonemas individuais não podem ser considerados unidades semânticas isoladas, no mesmo sentido de que, não há semântica associada com sons tais como o trovão ou o vento (NATHAN, 2007). No entanto, isso não os exclui de pertencerem a uma categoria. É possível que fragmentos de palavras possam adquirir algumas associações semânticas, porém o fonema já deve estar em interação com o morfema.

Os autores em LC costumam lançar mão de analogias com imagens para tentar dar conta de suas propostas teóricas. A imagem específica que Lakoff (1987) propõe quando discute categorização é a de uma roda com raios, similar à roda de uma bicicleta, com formato de círculo que sugere simetrias e regularidades. Trata-se do que ele chama protótipo de categoria radial. Nathan (2008) defende que a única desvantagem dessa analogia é a de que o formato de uma categoria está longe de ser circular. Ele prefere acionar a imagem de uma ameba cujo formato é assimétrico e irregular, capaz de se movimentar com seus pseudo-pés para abarcar novos membros de acordo com alterações de nível social e interacional.

Dedicando-se a essa analogia, Nathan (1986) apresenta uma análise dos aspectos problemáticos do fonema do inglês americano /t/. Tais aspectos dizem respeito ao fato de que as instanciações do fonema /t/ são amplamente diversas e, tomadas como um conjunto, não compartilham um local ou uma maneira de articulação. O autor demonstra que há, pelo menos, cinco diferentes variantes do fonema /t/ no inglês americano, dispostas abaixo:

[t^h] *tall*,

[t] *stall*

[□] *Button*

[tʰ] *what!*

[□] *water*

Como apontado no trabalho de Nathan (1986), não há um traço único do qual todos esses sons compartilhem, nem mesmo o [- sonoro] (o flap é sonoro); no entanto, não há dúvidas de que falantes nativos categorizam todos eles como tipos de /t/. Na verdade, leva-se alguns meses para estudantes de fonética começarem a se conscientizar da existência de todas essas variantes. Dessa maneira, Nathan argumenta que essa categoria de som pode ser entendida como um exemplo clássico de conjunto radial, com um membro central prototípico e outros membros do fonema irradiando para fora de

acordo com princípios fonéticos bem definidos, análogos às extensões descritas por Lakoff envolvendo princípios como metáfora, metonímia etc. Evidências sugerem que o /t/ surdo e não-aspirado é o membro central.

A teoria fonológica inovou quando criou a noção de que as variações dos sons pretendidos poderiam ser agrupadas em conjuntos, com o fonema funcionando como o nome do conjunto, com cada variante individual sendo chamada de alofone. Sapir, como Baudoiu, acreditava que esses conjuntos eram modos de percepção, sendo que falantes ouvem todos os membros do conjunto como se fossem o mesmo.

Nos termos da Gramática Cognitiva, alofones são transformações de esquemas imagéticos de sons prototípicos que criam ambientes particulares. Para tornar a questão mais clara, Nathan (2006) faz uma analogia com a discussão de Lakoff (1987) sobre a palavra *over*:

We prototypically think of this Word as referring to a trajector located ‘above’ the landmark (i. e., oriented vertically with respect to the force of gravity), but we can certainly put some wallpaper “over” a hole in the ceiling. Just as we see the hole in the ceiling “upside down” without noticing that we have made any kind of change or extension, so we produce an aspirated /t/ without being aware that we have made it sound any different than the original target, which, as I argued above, is unaspirated. The idea that we come pre-equipped with natural responses to motor difficulties that we need to unlearn is not mysterious, nor is it all foreign to the basic principles of Cognitive Grammar (NATHAN, 2007, p. 617-618).

Nós prototipicamente pensamos nessa palavra como se referindo a um trajector localizado ‘acima’ do alvo (isto é, orientado verticalmente com relação à força da gravidade), mas certamente podemos colocar um papel de parede “over” um buraco no teto. Vemos o buraco no teto invertido sem notar que fizemos um tipo de mudança ou extensão, então nós produzimos um /t/ aspirado sem estarmos conscientes de que ele soa diferente do alvo original, que, como discuti acima, é não-aspirado. A ideia é de que nós vimos pré-equipados com respostas naturais a dificuldades motoras e a de que nós precisamos reaprender não é um mistério, nem é isso tudo estranho aos princípios básicos da Gramática Cognitiva (NATHAN, 2007, p. 617-618, tradução nossa).

Esse pré-equipamento relacionado, de alguma forma, à noção de prototipicidade, sobre o qual fala o autor, é assunto fundamental em nosso trabalho e será o tema da próxima seção. Concluímos por hora que a FC, na voz de Nathan, está ancorada nos pressupostos cognitivos vinculados à teoria dos protótipos.

2. 2 Bases corpóreas do fonema

Partindo da importante premissa de que a linguagem é também determinada pela nossa experiência corporal, continuaremos a expor descobertas e contribuições de Nathan para a teoria fonológica – inserida no arcabouço da LC –, direcionando a atenção para o papel do nosso corpo na produção dos sons ditos segmentais para, em um momento mais adiante, falarmos dos suprasegmentais. Encaramos o fonema como uma categoria que ajuda a compor o aparato gramatical conceptualmente motivado e essa motivação é também de base corpórea.

Iniciaremos nossos argumentos reportando dados apresentados por Nathan acerca da organização dos aspectos fonéticos e fonológicos das línguas para estabelecermos uma explicação sobre a motivação física desses fatos e fenômenos linguísticos. Por exemplo, uma língua que tem muitas consoantes permite palavras pequenas, mas requer bastante precisão em cada consoante, visto que consoantes similares poderiam ser confundidas. Da mesma maneira, línguas tendem a evitar pontos de articulação muito semelhantes. Por exemplo, poucas línguas contrastam oclusivas dentais e alveolares. Na comunicação do dia-a-dia, as pessoas muitas vezes têm pressa e costumam evitar grandes esforços, de tal forma que palavras grandes implicam aborrecimento e menos eficácia. Sabe-se que palavras menores são mais eficientes. Um antigo princípio chamado Lei de Zipf mostra que há uma relação inversa entre o comprimento da palavra e sua frequência – palavras mais frequentes são menores. Tal lei, que utiliza fórmulas estatísticas, tem como escopo estudar a probabilidade de distribuição dos dados. As línguas são impulsionadas a adotar muitas consoantes e vogais diferentes. Assim como queremos que os sons sejam diferentes uns dos outros, não desejamos que as diferenças sejam tão grandes a ponto de exigirem movimentos articulatorios muito complexos, por exemplo, em situações nas quais tentamos avisar alguém sobre um perigo eminente (NATHAN, 2008).

Tendo em vista esses argumentos, corrobora-se que a prototipicidade é preferencial a partir da experiência física. Segundo Nathan, há significativas regiões na boca onde pequenas diferenças no ponto de articulação são relativamente inaudíveis, e outros locais onde um pequeno movimento faz uma grande diferença acústica. Esses fatos sobre o trato vocal colaboram para que os pontos primários de articulação tendam

a ser bilabial, coronal (também dental ou alveolar) e velar. Isso porque em várias línguas do mundo se encontram consoantes produzidas em cada um desses três locais. Similarmente, vogais são mais altas e mais audíveis à medida que a boca se abre mais, o que nos conduz a criar expectativas acerca de que cada língua tenha uma vogal aberta e não-arredondada [a].

Nathan (2007) faz uma analogia com a prática de ensino e aprendizado do caratê. Os iniciantes nesse esporte parecem apresentar resistência a movimentos que requerem duas atividades diferentes com os braços ao mesmo tempo. Mesmo após numerosas repetições, reverterem as mãos para movimentos não ensinados, que eles nunca tinham visto antes, mas que são mais naturais no sentido de que o corpo humano tem preferências naturais para movimentos dos membros que seguem um padrão de simetria bilateral. Falhas espontâneas similares ocorrem no aprendizado da dança quando um passo assimétrico é exigido. A emergência de padrões motivados naturalmente faz parte de todo aprendizado humano de habilidades motoras, e a produção da fala não está isenta de integrar esse quadro. Os processos naturais da Fonologia Natural são nada mais que instanciações perceptuais e articulatórias de tais falhas. Deprendemos, então, que nosso trato vocal possibilita e, ao mesmo tempo, restringe nossos movimentos na produção da fala.

Queremos sustentar nesse trabalho que, durante a produção do som, há uma licença física, corpórea, e uma motivação cultural atuando em conjunto. São muitos os autores que, de alguma forma, sustentam nossa alegação. É o caso de Valliima-Blum (2005), que cita Demolin (2002) e faz menção ao fato de que o número médio de sons segmentais entre as línguas é de 33. Tal informação aponta para uma certa limitação a esse respeito, ou seja, embora línguas se diferenciem minimamente quanto a esse número, há uma certa universalidade em se dizer que essa quantidade não é ultrapassada. Valliima-Blum assume que o número de fonemas distintivos de uma língua é necessário para a expressão de contrastes semânticos no funcionamento linguístico da mente humana. Utilizando uma progressão aritmética para contabilizar a produtividade desses 33 fonemas em uma língua na produção de palavras e sentenças, a autora chega a um número bastante expressivo, o que comprova a não necessidade cognitiva de haver mais do que o número encontrado na maioria das línguas. O comportamento vocal humano tende ao universal em vista da estrutura da anatomia e psicologia que produzem o som (NATHAN, 2007). O aparelho fonador apresenta limitações como, por exemplo, articular um som em que a língua toque a ponta do nariz.

Isso faz com que o conjunto de sons possíveis de ocorrer nas línguas naturais seja limitado (MORI, 2001).

Apostando em uma influência direta de nossas experiências físicas na nossa manifestação linguística, ou mais propriamente, na nossa Gramática, MacNeilage (1998) e MacNeilage e Davis (2000), *apud* Nathan (2007), discutem que uma estrutura silábica é uma adaptação do comportamento de mastigar, relacionado ao comportamento inato de sugar. Tal associação se dá pelo ritmo alternado de abrir e fechar o trato vocal. Isso pode ilustrar como nos servimos do nosso corpo e interagimos nossas necessidades com nossas capacidades físicas.

Princípios gestálticos explicam por que segmentos de alta sonoridade como vogais tendem a servir como núcleo das sílabas. *Onset* é o segmento inicial ou ataque, não inclui a vogal. *Rhyme* é dividido em duas partes menores, a saber, o núcleo, que representa o elemento mais sonoro e indispensável na sílaba; e coda, que inclui qualquer consoante ou sequência de consoantes seguindo o núcleo. Nathan (2007) afirma que a sílaba prototípica é a sílaba simples CV (consoante vogal). Uma vez que, em consoantes vozeadas, o vozeamento emerge na vogal, espera-se que consoantes desvozeadas sejam melhores consoantes – demarcadoras de cada batida. O ponto mais alto da batida deve vir imediatamente depois do *onset*, havendo o “desvozeamento” do som na próxima batida.

Tratando do nível suprasegmental e mais especificamente do assunto que analisaremos, Reich (2009) afirma que curvas entoacionais tendem a mostrar um declínio fonético devido ao desgaste de pressão subglotal ao falar. Assim, o início da fala é marcado, por exemplo, pela ascensão dos contornos melódicos; já o término da fala, pelo declínio. O desgaste da pressão subglotal pode explicar a tendência de um universal linguístico, visto que é encontrado em várias línguas esse movimento e essa configuração melódicos. Ao longo do trabalho, mais precisamente na discussão dos dados, descobriremos se no nosso caso essa tendência se confirma. Logo, o que estamos expondo corrobora a visão de que existe uma licença física que monitora nosso comportamento com relação ao som da fala.

Quanto à motivação cultural que acreditamos estar envolvida também nesse processo, é importante destacar que as diferentes culturas particularizam as restrições de caráter biológico. O modelo sociocognitivista revê a dicotomia entre fundacionismo e relativismo. Conforme o primeiro paradigma, os sistemas conceituais humanos ancoram-se em uma base fixa e universal. Já para o segundo, os conceitos humanos não

têm qualquer lastro de universal, sendo variáveis e contingentes e estando sempre assujeitados ao tempo, ao lugar, à sociedade, à história, às inclinações e às convicções individuais (MARTINS, 2002). Diante desse quadro, os sociocognitivistas adotam uma perspectiva híbrida, sendo fortalecida a versão do relativismo.

Um bom exemplo para entender como podem estar os dois paradigmas em conformidade é o estudo acerca das cores. Pesquisas mostram que diferentes línguas do mundo tendem a selecionar suas categorias de cores a partir de um conjunto universal de cores focais. As cores focais são aquelas que o sistema visual humano é especialmente preparado para responder, tais como vermelho e verde, azul e amarelo. Além disso, elas são lexicalizadas primeiramente (BERLIN E KAY, 1969 *apud* TAYLOR, 2006).

A pesquisa das cores contribuiu decisivamente para a teoria dos protótipos. Estudos mostram que determinadas línguas possuem, por exemplo, somente dois nomes para distinguir as cores. Isso levou à conclusão de que fatores culturais determinam a quantidade de nomes designadores de cores em uma língua. Pode-se dizer que há uma relação direta do número de artefatos que necessitam ser nomeados quanto à cor em cada cultura com o repertório dos nomes para esse tipo de designação. Atrelada aos estudos da categorização humana em geral, a cognição cromática representa um avanço e uma derrubada no cognitivismo clássico, uma vez que o cognitivismo não-hegemônico entende o conceito de cor como corporificado, determinado pela biologia humana, por mecanismos cognitivos, bem como por questões culturais.

Postulamos sim uma interferência cultural na realização dos fonemas e também das nossas curvas entoacionais. Enfim, partimos desde o início dos pressupostos de que nossa cognição é cultural, assim como nossa Gramática é motivada conceptualmente a partir de experiências físicas (LAKKOF; JOHSON, 1999).

A apresentação dos postulados de Nathan a respeito do fonema como categoria básica e o entendimento de que os sons da fala têm bases corpóreas preparam o terreno teórico para a adoção de outra possibilidade analítica acerca da fonologia. Ao se admitir que os fonemas prototípicos sofrem alterações substanciais ao entrarem em uso no jogo simbólico, ganha espaço a noção de construção como emparelhamento entre forma (som) e sentido, cujas afetações mútuas são ditadas a partir de uma visão de linguagem como uma série de ações conjuntas e coordenadas.

2.3 Fonema em e como construção

Como já foi mencionado, encaramos a linguagem como uma rede de símbolos, cujas unidades simbólicas são construções linguísticas, ou seja, pareamentos entre forma e sentido. Reivindicamos a inclusão de informações fonéticas, fonológicas e prosódicas no pólo formal das representações das construções.

Tomadas essas posições e dando continuidade à apresentação agora da segunda linha de investigação em fonologia situada na perspectiva funcionalista, acreditamos que explicações relacionadas às estruturas linguísticas devem incluir aspectos que envolvem o dinamismo e a interação do uso. Estamos falando de processos que se dão, por exemplo, na aquisição contínua da linguagem. E é por esse motivo que comentaremos a seguir algumas ideias de autores que seguem essa linha de estudos, como os teóricos Bybee (1994) e Tomasello (2003).

Liderando outro ramo de investigação em fonologia, centrado nos modelos baseado no uso, Bybee (1994) pleiteia uma perspectiva funcional e cognitiva para o fenômeno. Ao reconhecer, na época, as escassas investigações nessa área em detrimento da força das abordagens morfossintáticas, a autora preconiza que as estruturas fonológicas também são moldadas pelo uso. A Fonologia Natural é também citada em Bybee – assim como em Nathan – por ser considerada uma versão do funcionalismo. Na visão funcionalista, a gramática é um artefato de processos comunicativos, sendo as propriedades da língua natural moldadas pela linguagem em uso.

A fonologia à luz dos pressupostos funcionalistas implica:

- Consider the substance of phonology rather than Just the structure (...) For phonology the substance is largely phonetic, but generative phonology has also allowed the consideration of morphology and lexicon in the conditioning of “phonological” rules.
- Consider the uses to which phonological elements are put (...) Very little study has been devoted to the distribution of phonological elements in texts (...) the frequency of segments affects their phonetic shape and evolution.
- Consider subphonemic detail and variation conditioned lexically, morphologically and socially (...) the study of detail in phonology will reveal important facts that bear on our understanding of how language is really processed and what structures have empirical validity.
- Attend to exceptions and marginal cases, for they can be valuable sources of information about the nature of processing and representation.

- Reconsider what Langacker (1987) calls the “rule-list fallacy”. Our thinking and analyses need not be restricted to only two options – either an element occurs in a list or it is generated by rule (...) lexical elements (words or phrases) consist of actual phonetic content that is modified as these elements are used (BYBEE, 1994, p. 287, tradução nossa).

- Considerar a substância da Fonologia ao invés de somente a estrutura (...) Para a Fonologia, a substância é amplamente fonética, mas a Fonologia Gerativa também tem considerado a Morfologia e o Léxico no condicionamento de regras “fonológicas”;

- Considerar os usos nos quais os elementos fonológicos estão inseridos (...) Poucos estudos têm se dedicado à distribuição de elementos fonológicos em textos (...) a frequência dos segmentos nos textos afeta seu formato e sua evolução fonética;

- Considerar detalhamento e variação subfonêmicos condicionados lexical, morfológico e socialmente (...) o estudo que contemple esses detalhes revelará fatos importantes sobre nosso entendimento de como a língua é realmente processada e quais estruturas possuem validade empírica;

- Atender a exceções e casos marginais, para que eles possam ser fontes valiosas de informação sobre a natureza do processamento e da representação;

- Reconsiderar o que Langacker (1987) chama de “*rule-list fallacy*”. Nossa análise não deve ser restrita a duas opções: ou que um elemento ocorre em uma lista ou que ele é gerado por uma regra (...) elementos lexicais (palavras ou sintagmas) consistem de um conteúdo fonético real que é modificado conforme esses elementos são usados (BYBEE, 1994, p. 287).

Dialogando inclusive com a Fonologia Articulatória (que será tratada posteriormente neste trabalho, quando estivermos lidando com o ritmo da fala e com a teoria que encara a linguagem como um sistema dinâmico), Bybee afirma que processos fonéticos podem ser vistos como redução ou sobreposição de gestos articulatórios. As propriedades desses processos relacionam-se a propriedades gerais dos gestos motores e a características particulares do sistema articulatório.

Mudanças fonéticas propagam-se gradualmente através do léxico e tendem a ser mais avançadas em palavras mais frequentes. Itens mais frequentes tendem a ser mais conservadores em mudanças não-fonéticas e mais inovadores nas fonéticas (BYBEE, 1994).

No âmbito fonético, palavras ou unidades lexicais maiores são perceptos auditivos associados com padrões de comandos motores. O que sabemos sobre outras unidades perceptuais armazenadas sugere que detalhes não são descartados, mas preservados em uma extensão maior. Citando Langacker (1987), a autora aponta que se regularidades fonéticas nas palavras são descartadas, o resíduo não seria “reconhecível

como uma entidade coerente plausivelmente atribuída à autonomia cognitiva” (BYBEE, 1994, p. 292). O fato de haver mudanças graduais, porém não definitivas, na língua em uso, não nega a realidade ou sistematicidade dos padrões de co-articulação.

O papel da frequência tem sido negligenciado na fonologia. Entretanto, a frequência é importante uma vez que palavras mais frequentes serão usadas em posição de menos acento e ênfase e então são mais suscetíveis a processos de redução. Palavras mais frequentes têm mais oportunidades de sofrer processos do que as menos frequentes.

O processamento de palavras envolve uma classificação e uma combinação com outras imagens similares armazenadas. Isso significa que elementos foneticamente similares de palavras serão associados com outros, e padrões comuns são organizados dentro de esquemas. Esses esquemas representam padrões motores que são usados na produção real. Como em qualquer tarefa motora que se realiza com grande velocidade e fluência, padrões repetidos facilitam a produção (BYBEE, 1994). Assumimos, portanto, que as palavras – como um todo – formam a base da categorização. O armazenamento é holístico.

Problematizando a questão acerca de como falantes armazenam os sons, as teorias baseadas no uso postulam, então, que falantes constroem categorias de fonemas com base em instâncias numerosas de eventos de fala reais armazenados, de tal maneira que aspectos comuns surgem naturalmente das similaridades entre imagens acústicas ou eventos articulatorios. Para esses teóricos, regras não podem ser estabelecidas pelos falantes sem um armazenamento anterior de um grande número de instâncias individuais para permitir uma extração de regularidades na forma de um esquema que as instâncias suportam. A visão que emerge é a de uma maciça rede na qual estruturas com variados graus de entrincheiramento e diferentes graus de representação da abstração são ligados juntos em relações de categorização, composição e simbolização (LANGACKER, 2000 *apud* NATHAN, 2007).

Segundo Bybee (1994), palavras são armazenadas como instâncias individuais, e falantes desenvolvem generalizações a partir de similaridades entre pedaços de palavras armazenadas, ou seja, entre morfemas, unidade a partir da qual julgamos que há um emparelhamento com o fonema, estabelecendo um processo de integração construcional. Se nos atentamos, por exemplo, para o desempenho de crianças que estão aprendendo a escrever, torna-se interessante reparar como elas, em geral, não apresentam dificuldade em soletrar palavras com letras em comum, apesar de tais letras

representarem sons diferentes (ex: o “s” em casa e sapo). Isso somente acontece porque o processo é *online* e o falante está lidando com unidades reais, e não isoladas. O ato de soletrar, por mais que pareça isolado, engloba um contexto mínimo, a palavra.

Não podemos deixar de ratificar a visão de que as mudanças e as variações são feitas no contexto da produção da fala, e uma análise de tal fenômeno deve ser, por isso, mono-estratal. Assim, o ambiente em que o fonema está inserido determinará muitas variações. Interagindo, pois, tais argumentos com a proposta construcional, concluímos que, na construção, um fonema interage com outros, sendo que ele afeta o ambiente em que está contido e, da mesma maneira, o ambiente o afeta também. Um exemplo disso é a alteração do fonema /s/ para /z/ quando integra a construção “as asas” em português. Nesse caso, um som surdo torna-se sonoro por força da vogal que inicia a palavra “asas”.

Conforme Nathan (2007), outra razão para se acreditar que a variação alofônica é computada durante a fala advém dos “erros” de fala. Muitas pesquisas demonstram que “erros” de fala normalmente ocorrem no nível fonêmico e que sons são sempre ajustados a suas novas posições. Tratando-se da aquisição da linguagem, crianças podem produzir palavras de um jeito que nunca ouviram antes, como “eu fazi”, em vez de “eu fiz”. Por conta de dados como esse, pode-se inferir que elas não simplesmente repetem o que ouvem, mas armazenam alvos corretamente e os processam na testagem de suas habilidades linguísticas. Assim sendo, fonemas são ajustados em tempo real aos seus contextos. Retomando Baudouin (1972), já encontrávamos a ideia de que percebemos e armazenamos um som de uma forma, mas ajustamos essa forma quando realmente falamos.

A teoria clássica, conforme Nathan (2007), nega a possibilidade de sobreposição de fonemas, porquanto ela dependia da visão aristotélica de classificação na qual um som não poderia simultaneamente pertencer a duas diferentes categorias de uma vez. Ocorre que o acesso lexical interage com a percepção fonêmica de complexas maneiras. Em inglês americano, os fonemas /t/ e /d/ podem se realizar como *flaps* em posição intervocálica. Isso mostra a sobreposição em uma área das categorias desses fonemas que não são completamente distintas. Como exemplos, são citadas as palavras *butter*, *Betty*, *electricity* e *rider*, *validity*, *grading*.

Existem evidências de que falantes armazenam essas categorias da maneira “como eles diriam”, e não como meros “gravadores”, como sustenta Taylor (2006). Por outro lado, pesquisas defendem que palavras são aprendidas a partir da reprodução de

aspectos prosódicos anteriormente percebidos. É possível postular, entretanto, que ambos os pensamentos se convergem, visto que a reprodução está sujeita a minuciosos aspectos biológicos e culturais de cada ser humano. A interferência da estrutura do corpo é importante nessa “reprodução”. O corpo interage na coordenação das ações.

Todavia, nosso foco não é estreitar tanto a discussão acerca de armazenamento do som, pois acreditamos que muitos outros fatores podem estar envolvidos em tal processo. Palavras, sílabas e fonemas/alofones passam pelo mesmo problema: a variabilidade acústica. Eles podem ser pronunciados de ilimitadas maneiras de acordo com o contexto linguístico da unidade, o ambiente fonético imediato, seu lugar dentro de um contorno entoacional etc., assim como de acordo com as propriedades dependentes dos falantes (dialeto, gênero, idade, propriedades específicas do trato vocal do falante etc.) (TAYLOR, 2006). Por isso, assumimos que os sons são sujeitos a efeitos de prototipia, ou seja, os vários tipos de sons podem ser categorizados de diferentes maneiras nas línguas do mundo, apesar do trato vocal ser praticamente o mesmo entre os seres humanos.

Pelo que se pode perceber, muitas são as opiniões acerca dos fatos relacionados aos sons da fala. Alguns apresentam evidências de que falantes podem ouvir e armazenar detalhados aspectos fonéticos, embora não possuam acesso metalinguístico a esse tipo de informação. Esta é a posição dos gramáticos cognitivos e dos fonologistas experimentais. Outros farão objeções a essa visão, apostando em uma interferência maior por parte do falante. As teorias baseadas no uso situam a aquisição da linguagem como um caso de armazenamento de um grande número e instâncias individuais, a partir das quais serão extraídas generalizações. Bybee (1994) afirma que as crianças armazenam as unidades como as palavras, e até outras unidades maiores, exatamente como são pronunciadas, ou seja, com rico detalhamento fonético. Essa pode ser considerada uma visão bastante concreta do processamento fonológico.

É importante deixar claro que assumimos um compromisso com a proposta da Gramática Cognitiva, que se opõe ao espectro estruturalista de que fonemas somente são definidos em oposição dentro de um sistema. Os fonemas, assim sendo, passam a ser definidos pelo modo de percepção, armazenamento e produção na língua. Consoante com a Fonologia Natural, divergimos também da perspectiva gerativista, pois acreditamos que fonemas não se tratam de listas de traços. Isso quer dizer que não definimos os fonemas por contraste, nem por especificações abstratas dos sons. Pensaremos o fonema na visão construcionista, tentando aproximá-lo mais da ideia de

sons reais, sem negar que sejam mentais. É importante destacar o entendimento de processos no modo pelo qual falantes produzem e percebem o som, evitando a reivindicação de simples registros, como cópias perfeitas. Em outros termos, entende-se que a visão processual é crucial para este trabalho porque prevê a integração dessas categorias abstratas com os ambientes em que os fonemas são realizados. Quando nos referimos à percepção, reafirmamos que há a intenção do outro:

a phoneme...is na underlying *intention* shared by the speaker and the listener (who are always “two in one”). The shared knowledge of intentions guarantees communication between the speaker and the listener within a given language, even if the actually pronounced forms diverge substantially from what is intended (...) In other words, phonemes are fully specified, pronounceable percepts (DZIUBALSKA-KOLACZYK *apud* NATHAN, 2009. p. 5).

um fonema (...) é uma *intenção* subjacente compartilhada pelo falante e pelo ouvinte (que são sempre “dois em um”). O conhecimento compartilhado de intenções garante a comunicação entre o falante e o ouvinte dentro de uma dada língua, mesmo se as formas efetivamente pronunciadas diverjam substancialmente do que é pretendido (...) Em outras palavras, fonemas são perceptos totalmente especificados e pronunciáveis (DZIUBALSKA-KOLACZYK *apud* NATHAN, 2009. p. 5, tradução nossa).

Isso resulta no fato de que falantes atuam automática - e possivelmente - inconscientemente com suas intenções na produção e percepção da fala. O fonema deve ser entendido como uma unidade mental, que envolve uma atividade cognitiva. O que aprendemos e usamos são morfemas. E é a partir dessa mesma noção de uso que pretendemos entender o fonema. Assumiremos que os fonemas são instanciados pelos morfemas, estando essas duas unidades diretamente em pareamento na construção linguística.

Comunicamo-nos e reconhecemos os sons que falamos e ouvimos porque assumimos olhares e percepções holísticas. A linguagem é construção, é pareamento de forma e sentido. Nossa cognição é cultural e há um caráter partilhado do nosso conhecimento em jogo na interação linguística. A noção de fonemas como intenções (recuperada em Baudouin de Courtenay), de certa maneira, interage com a noção de caráter partilhado da construção de sentido. A significação, segundo essa visão, é o resultado de uma atividade conjunta que presume cooperação e consentimento. Assim, voltando nosso olhar para a questão fonológica, procede dizer que, ao ouvirmos o outro,

estamos pensando, mesmo que inconscientemente, na intenção que pode estar subjacente aos sons escutados.

Os eventos sonoros em geral, sendo eles segmentais ou suprasegmentais, compondo o processo da significação, estão relacionados ao conhecimento de símbolos, os quais podem ser culturalmente reconhecidos. Nessa perspectiva, é imprescindível que se faça menção às ideias de Tomasello (2003) sobre o efeito catraca de acumulação de cultura. Como sustenta o psicólogo, a transmissão social ou cultural do conhecimento permite que cada organismo poupe muito tempo e esforço, para não falar de riscos, na exploração do conhecimento e das habilidades já existentes dos co-específicos. Por isso, o efeito é cumulativo, pela acumulação das tradições e dos artefatos culturais dos seres humanos ao longo dos anos. Assim, garante-se que a ontogênese cognitiva humana ocorra num meio que, em qualquer tempo, reúna toda a sabedoria coletiva de um grupo.

Na ontogênese individual, os seres humanos lidam desde cedo com suas habilidades de aprendizagem para adquirir as construções linguísticas que suas culturas particulares criaram ao longo do tempo histórico pelo processo de sociogênese. A linguagem natural é, pois, uma instituição social simbolicamente incorporada que surgiu historicamente a partir de atividades sociocomunicativas preexistentes. As crianças humanas estão biologicamente preparadas de várias maneiras para adquirir uma língua natural, ou seja, possuem aptidões sociocognitivas e fonoauditivas básicas. Isso porque seres humanos normais e saudáveis nascem com o mesmo tipo de mecanismo respiratório, sistema perceptual etc. À medida que a criança segue dominando os símbolos linguísticos de sua cultura, ela adquire a capacidade de adotar simultaneamente múltiplos pontos de vista sobre uma mesma situação perceptual (TOMASELLO, 2003).

Durante os primeiros estágios da aquisição da linguagem, é possível que palavras sejam aprendidas, armazenadas e recuperadas como “todos”, “conjuntos” fonológicos. À medida que o léxico aumenta, algum tipo de análise interna do tamanho das palavras torna-se necessária. Então, pedaços de uma palavra podem ser identificados com pedaços de outras palavras. Esses pedaços, por sua vez, podem se quebrar dentro de unidades ainda menores (TAYLOR, 2006).

No plano suprasegmental, resultados de estudos sugerem que, em quatro dias após o nascimento, os bebês já têm adquirido (presumidamente, isso ocorre durante os últimos meses de gravidez) a habilidade de distinguir a prosódia da sua língua nativa da de outras línguas. Inclusive, estudos bastante recentes anunciam que, desde seus

primeiros dias de vida, os bebês choram no idioma materno, já que, ao emitirem seus primeiros sons, levam a marca da língua de seus pais. Ao chorar, as crianças preferem os sons típicos do idioma que ouviram no último trimestre de gestação (WERMKE, 2009). No experimento, bebês franceses tenderam a chorar em um tom ascendente, enquanto os alemães faziam em tom descendente. Os cientistas argumentam que, na origem dessas diferenças, poderá estar a tentativa das crianças em criar laços com a mãe, imitando-as. Dessa maneira, os bebês percebem o conteúdo emocional das mensagens mediante a entoação. Tal evidência corrobora o fato de existirem características específicas no sistema entoacional de cada língua (HIRST; DI CRISTO, 1998).

Há dois comportamentos sociais que podem sugerir que os bebês humanos não são apenas sociais como outros primatas, mas, antes, “ultra-sociais”: pouco depois de nascerem, eles entabulam “proto-conversas” com quem cuida deles (olhar, tocar e vocalizar – emoções básicas); os neonatos humanos imitam alguns movimentos corporais dos adultos, sobretudo alguns movimentos da boca e da cabeça (TOMASELLO, 2003).

Há unidades linguísticas, aprendidas primeiramente, que são consideradas próximas do que é universal, e parecem ser diretamente percebidas como tal. Podemos postular que há unidades prosódicas as quais possuem proeminência perceptual imediata no nível básico. Reconhecemos tipos de melodias da fala assim como, segundo Nathan (2008), prontamente reconhecemos um cão ou um gato. Enfim, categorizamos o mundo com base no modo como interagimos com ele. E isso se aplica aos sons da fala.

Ao desenvolver seus argumentos acerca do conhecimento simbólico da língua, Goldberg (1995), com base em Fillmore (1997), defende uma visão de que os significados são tipicamente definidos com relação a algum *frame* ou cena particular, que pode ser altamente estruturado. Para ilustrar, a autora explica que a diferença central entre *ceiling* e *roof* do inglês reside nos *frames* distintos a que cada termo diz respeito (*ceiling* se refere ao interior de uma construção, enquanto *roof*, ao exterior).

O estudo de *frames* associa-se à visão holística das unidades linguísticas, assumida pela GC, na qual o todo é maior que a soma das partes que o integram. Além disso, o todo, conforme o olhar gestáltico, é cognitivamente mais simples. Dessa maneira, o tratamento da composicionalidade é assumido na sua versão mais fraca.

Assim, entendemos que, no jogo simbólico da língua, os falantes recorrentemente acionam seu conhecimento enciclopédico, ou seus *frames* armazenados

na memória. Defendemos ainda que, tratando-se de prosódia, o mesmo acontece quando o falante adquire, através da experiência, vários modelos em que uma ou outra configuração prosódica são assumidas. As molduras prosódicas que desempenham um importante papel na produção de sentido é um tipo de *frame* vocal – noção que vai ao encontro da noção de *frame* exposta acima.

De acordo com Tomasello (2003), recordar determinados objetos, co-específicos, eventos e todos os outros aspectos da experiência pessoal – e, em alguns casos, antecipar experiências futuras baseando-se nessas recordações – é a condição *sine qua non* da cognição. O mesmo autor, acerca do olhar holístico sobre novas construções, afirma que, ao mesmo tempo que estão adquirindo suas primeiras palavras, as crianças também estão adquirindo construções linguísticas mais complexas, certos tipos de *gestalts* linguísticas.

Uma vez que as construções são unidades básicas do sistema linguístico e que é possível que haja sistematicidade entre as construções, Goldberg (1995, p. 67-68)¹ sugere quatro princípios organizadores da língua e que estão atrelados à unidade construcional. São eles:

- I. **O princípio da motivação maximizada:** Se uma construção A é relacionada sintaticamente a uma construção B, então o sistema da construção A é motivado a ponto de ser relacionado semanticamente a uma construção B. Tal motivação é maximizada.

- II. **O princípio da não-sinonímia:** Se duas construções são sintaticamente distintas, elas devem ser semântica ou pragmaticamente distintas. Aspectos pragmáticos incluem estrutura informacional (tópico e foco), além de aspectos estilísticos como o tipo de registro.

Corolário A: Se duas construções são sintaticamente distintas e semanticamente sinônimas, então elas não devem ser pragmaticamente sinônimas.

¹ Na realidade, Goldberg (1995) recupera em outros autores a base de tais princípios. Observações de sutis diferenças no significado já haviam levado Bolinger (1968) a concluir que uma diferença na forma sintática leva a uma diferença no sentido. Além disso, a mesma hipótese – Princípio da não-sinonímia das formas gramaticais – foi formulada por Givón (1985), Kirsner (1985), Langacker (1985), Clark (1987) e Wierzbicka (1988). E este princípio tornou-se a hipótese de trabalho de Goldberg.

Corolário B: Se duas construções são sintaticamente distintas e pragmaticamente sinônimas, então elas não devem ser semanticamente sinônimas.

- III. **O princípio do poder expressivo maximizado:** o inventário de construções é maximizado para propósitos ou fins comunicativos.
- IV. **O princípio da economia maximizada:** o número de construções distintas é minimizado tanto quanto possível, dado o princípio III.

Por meio de todos esses princípios constatamos que, sendo a forma alterada, o sentido se transforma conseqüentemente, devido ao pareamento existente na língua entre esses dois pólos (forma e sentido). Através dos princípios III e IV, especialmente, percebemos que a gramática permite ao mesmo tempo tanto maximizar as possíveis construções da língua quanto restringi-las, visto que há uma tendência em se fazer generalizações e simplificações na língua.

Lembramos que, embora a autora concentre o trabalho citado em construções de estrutura argumental, ela não restringe a esses casos os princípios comentados, mas se refere à organização da língua como um todo. Visto que a LC assume que as unidades simbólicas são gerenciadas pelos mesmos princípios independentemente de graus de complexidade (fonema, morfema, palavra, sentença, enunciado etc.) podemos nos amparar paralelamente no modelo de análise de Goldberg, voltado para estrutura argumental. Dessa maneira, o presente trabalho defende uma reconsideração do Princípio da Não-Sinonímia (Goldberg, 1995), pois postulamos a inserção, no emparelhamento forma/sentido, de unidades construcionais fonológicas/prosódicas.

Assim, a nova roupagem que atribuímos ao Princípio da Não-Sinonímia é:

Princípio de Não-Sinonímia: se duas construções são sintática e/ou **fonologicamente** distintas, elas devem ser semântica e/ou pragmaticamente distintas. (ROCHA, 2004, p. 101)

O termo “fonologicamente”, pois, estende-se das unidades fonêmicas distintas – os fonemas – até os traços suprasegmentais ou prosódicos.

Podemos retomar Välimaa-Blum (2005), dizendo que os fonemas por si mesmos não são independentes, mas ocorrem em combinação com outros elementos para formar unidades simbólicas. Para a autora, esses sons estão o tempo todo associados a morfemas. Assim sendo, falantes, segundo ela, não possuem representação mental de

fonemas e alofones descartados de suas unidades significativas. Parece ficar claro, então, que a posição da linguista é a de que sons não têm uma existência autônoma na mente dos falantes. Além disso, ela postula o conceito de fonema e alofone do seguinte modo: o fonema é uma classe de sons foneticamente similares que falantes tomam como sendo a mesma entidade, principalmente porque todos os co-alofones geram contrastes de semântica idêntica.

Acreditamos que uma contribuição importante do trabalho de Välimaa-Blum é a de que, assumindo uma perspectiva construcional, é impossível separar fonologia de morfologia. Segundo a autora, a interação de fonemas com significados formam a base da corporificação na fonologia. Se morfologia e sintaxe podem ser consideradas como um *continuum*, como assumido na GC, prosseguimos não acreditando em módulos independentes e separados da língua. Isolar a fonologia da expressão do significado gera uma visão de mente como uma máquina precária (VÄLIMAA-BLUM, 2005). Assim sendo, podemos nos firmar na ideia de que para a LC, e mais especificamente a teoria da GC, não há semântica sem sintaxe, sem fonologia ou sem morfologia, isto é, não há sentido sem a base formal da língua. Qualquer item lexical é um símbolo, em virtude de emparelhar significado e som.

É importante enfatizar o valor pragmático que age diretamente no funcionamento das unidades construcionais. Considerando exemplos no nível da sentença, diferentes estruturas sintáticas podem expressar o mesmo conteúdo proposicional. Abaixo, o mesmo verbo e os mesmos sintagmas nominais ocorrem em diferentes construções, mas nem todos expressam a mesma leitura pragmática:

- a. Alice descascou este abacaxi.
- b. Alice descascou este abacaxi?

A distinção entre “a” e “b”, licenciada por aspectos melódicos, marca a primeira como afirmação, e a segunda como interrogação. Essa licença nos permite tratar “a” e “b” como construções distintas com finalidades comunicativas também distintas. Os efeitos perlocutórios de cada uma delas podem se diferenciar mesmo com a estrutura “repetida”. Ou seja, o enunciado “a”, por exemplo, pode se configurar como um comentário simples (Alice descascou este abacaxi) ou como uma ordem, do tipo: se Alice descascou este abacaxi, descasque o outro. Possivelmente, a prosódia vai variar segundo tais versões pragmáticas de estrutura similar. Esses enunciados apresentam um mesmo conteúdo proposicional que têm a ver com Alice ter descascado o abacaxi. O

que as diferentes formas fazem é dar diferentes pontos de vista pragmáticos, com o apoio da fonologia e, em especial, da prosódia.

Para finalizar esta seção, gostaríamos de frisar qual é o pareamento que existe quando nos referimos à construção fonológica. Considerando a licença metafórica, nossa fala “transporta” nosso pensamento, “carrega” o que está escondido em nossas mentes em formas de construções, como palavras e sentenças. O ouvinte, por sua vez, conhecendo a associação entre os eventos sonoros e seu sentido, pode “concretizar” a comunicação. Afinal, estamos tratando da materialização do som, via recursos cognitivos e corpóreos. É interessante, à luz de Saussure, pensar que, de alguma maneira, forma (som) e sentido (ideia) estão sempre atreladas um a outro:

O som é o instrumento do pensamento.

Dois elementos que entram em jogo no funcionamento da Língua: as ideias e os sons.

A Língua representa o papel intermediário entre o pensamento e o som.

Uma ideia se fixa em um som e uma sequência fônica se constitui como significante de uma ideia (SAUSSURE 1974, p. 38).

É curioso também mencionar que Nathan (2008) registra a existência de estudos bastante antigos em fonologia. É o caso do trabalho do linguista indiano Pāṇini, que descreveu a gramática do sânscrito, uma língua indiana falada nos anos 1000 A.C. O livro contempla a fonologia completa do sânscrito e, com algum treinamento, pode ser lido e usado nos dias de hoje. O propósito do gramático não era puramente científico. Ele e, presumivelmente, as pessoas com as quais trabalhava e as que interpretaram sua obra queriam preservar a pronúncia “correta” do sânscrito, língua que fundamentava a literatura da religião hindu. Acreditava-se que as orações só seriam eficazes se pronunciadas de acordo com o padrão fonológico sagrado. Recuperamos essa observação de Nathan por considerarmos um forte argumento de amparo à premissa de que sons, de alguma forma, estão em pareamento com significados.

Enfim, a partir das perspectivas do conhecimento fonológico baseado em experiências corpóreas e dos modelos baseados no uso, para nós, fonemas atuam no jogo simbólico da língua em conjunto com outras unidades e participam de maneira decisiva na construção de sentido. A partir de agora, subfocalizamos a questão do fonema nos concentraremos na prosódia também como decisiva nesse emparelhamento forma/significado do qual nos ocupamos.

2.4 Prosódia: percurso e relevância

Assim como anteriormente uma breve revisita ao conceito de fonema foi realizada, aproveitamos agora para situarmos, de modo panorâmico, a prosódia na teoria fonológica desde estudos iniciais. No século XVI e XVII, já se falava em entoação nos manuais de pronúncia correta ou ortoepia. Relacionava-se prosódia à pontuação, aos graus de pausa, às questões de marcadores de exclamação etc. Com o advento da Linguística Moderna, os estudos em prosódia ainda não haviam adquirido a relevância merecida:

De Saussure a Trubetzkoy e de Martinet a Bloomfield, passando por Hjelmslev, Jakobson e Sapir, entre outros grandes nomes, os estudos fonológicos de cunho estruturalista investiram “pesado” na descrição dos segmentos que constituem o *continuum* sonoro, sendo destinada, em contrapartida, pouca (ou nenhuma) atenção à análise e à descrição dos suprasegmentos (...). É certo que tanto o Estruturalismo quanto o Gerativismo se referem aos chamados “traços suprasegmentais” dos sons da fala, mas quando o fazem, (a) ora lhes conferem *status* marginal ou secundário, (b) ora renegam seu lugar no âmbito da Teoria Linguística. (GONÇALVES, 1997, p. 64-65)

No Estruturalismo, a prosódia é basicamente considerada como parte do estudo das línguas tonais. Assim sendo, a análise restringe-se à função distintiva desempenhada pelo tom. Mas, nesse caso, fatos prosódicos eram equivalentes a fonemas, diferindo significados em uma língua. Para Martinet (1969), *apud* Gonçalves (1997), os fatos prosódicos não tinham valor nas relações de presença *versus* ausência, não podendo ser comutados ou segmentados.

Com o Gerativismo, instalou-se uma perspectiva um pouco mais processual de investigação e, conforme diz Anderson (1985), *apud* Gonçalves (*op. cit.*), o modelo era o de representações e de regras contra o da Fonologia Estruturalista, que era estritamente de representações. Inseridos na lista de traços fonéticos, proposta por Chomsky e Halle (1968), estão os traços prosódicos acento, tom e duração. Embora considerados não-lineares, esses traços foram concebidos como extrínsecos aos segmentos, mais restritamente de vogais, e foram tratados no interior de uma mesma matriz fonológica.

Enfim, não são propostas quaisquer formalizações para o controle ou para a manipulação quer das propriedades prosódicas fisicamente manifestadas, quer dos

fenômenos suprasegmentais. A prosódia, durante esse período, encontra-se relegada a segundo plano na investigação linguística.

Como aponta Gonçalves, a necessidade de se investir nos fenômenos suprasegmentais (que se pode chamar de fonologia-fundo em contraparte dos segmentos, fonologia-figura) fez com que o Gerativismo, mais recentemente, postulasse uma série de subteorias conhecidas genericamente como Fonologia Não-Linear: Fonologia Autossegmental, Métrica, Lexical e Prosódica rompem com a tradição anterior uma vez que analisam fenômenos além do segmento. A abordagem Autossegmental trata do tom e da hierarquização dos traços; já a Métrica e a Prosódica descrevem o acento e o ritmo; e a Lexical, por sua vez, postula a interação entre os componentes da gramática como um todo.

A prosódia, então, começa a receber mais atenção e importância para se entender o fenômeno geral da linguagem. Estudos em fonologia passam a incluir a investigação prosódica, passando a fonologia-fundo a ocupar lugar de destaque. Nas palavras de Gonçalves (1997), a prosódia pode oferecer a chave da interpretação. Não é difícil convencer alguém que a entoação é, muitas vezes, o que delimita o entendimento daquilo que ouvimos. Por exemplo, podemos ouvir após um evento musical o seguinte comentário: “que maravilha de banda”. No entanto, se o enunciado for proferido em tom enfadonho, não teremos dúvidas de que se trata de uma apreciação negativa acerca do espetáculo.

A prosódia, segundo Couper-Kuhlen (1986), desempenha as seguintes funções:

- **Informacional:** fatos prosódicos contribuem para contrastar informação nova com informação dada, sinalizando a estrutura informacional de enunciados. Normalmente, a combinação do *pitch* e do volume (parâmetros prosódicos cujos conceitos serão tratados adiante) evidenciam a informação nova;

- **Gramatical:** fatos prosódicos alteram a configuração gramatical. Por exemplo, o contorno prosódico diferencia um enunciado assertivo de uma pergunta. Construções ambíguas, muitas vezes, são solucionadas a partir de uma ênfase prosódica em determinado constituinte;

- **Illocucionária:** fatos prosódicos sinalizam a força intencional de um enunciado em dado contexto. É o caso em que, dependendo da entoação, pode estar sendo feito um pedido ou uma ordem;

- **Atitudinal**: trata-se da atitude prosódica com relação ao enunciado e ao contexto, podendo realçar, positiva ou negativamente, o valor de um termo expresso na sentença;

- **Indexical (ou Identificadora)**: a prosódia pode identificar características geográficas, sociais e individuais do falante.

Em Gonçalves (1997), tratando-se da função gramatical da entoação, são descritos como fenômenos linguísticos que necessitam da prosódia para portar significados: conexão sintática, especialmente em enunciados estruturalmente ambíguos; construções sintáticas com elementos deslocados ou topicalizados; ênfase, que possibilita, conforme Cagliari (1992), adjetivar um substantivo ou adverbializar um verbo, um adjetivo ou mesmo um advérbio; coesão frástica e coesão interfrástica, com especial relevo à distinção entre adjetivas restritivas e explicativas; pontuação, que requer, por exemplo, o conhecimento da natureza das pausas; e estabelecimento de constituintes sintáticos, tendo em vista a função segmentadora da prosódia.

Quanto à função ilocucionária, Gonçalves (1997) elenca temas e problemas da pragmática e da conversação que se estruturam em sintonia com elementos prosódicos. Tais temas e problemas são: a inferência, que pode ter a entoação como elemento relevante; a pressuposição, em que a localização da saliência acentual caracteriza o foco do enunciado em oposição aos pressupostos argumentativos do discurso, pronunciados, via de regra, em tom baixo; a modalidade, ligada ao conteúdo proposicional da sentença, permitindo distinguir pontos-de-vista do falante sobre o conteúdo veiculado; os atos de fala, especialmente os atos de fala indiretos requerem do falante uma prosódia particular.

Ao expormos as várias funções que desempenha a prosódia na língua, procuramos difundir a ideia central de que informações sintáticas, semânticas, fonológicas e pragmáticas convergem, todas elas, para a prosódia.

Segundo Moraes (1998), uma maneira de se conceituar prosódia é admitir suas origens musicais, referindo-se à parte da fonética/fonologia que se ocupa de elementos comuns à música e à linguagem, elementos esses que são intrínsecos à fala. Conforme Hirst e Di Cristo (1998), prosódia, tal como o termo suprasegmental, consiste em um sistema não-lexical, no qual características fonéticas específicas são relacionadas a parâmetros prosódicos acústicos. Os elementos considerados comuns entre música e fala são:

- **Altura melódica (*pitch*):** variação do tom (mais agudo ou mais grave). Tem como correlato acústico a frequência fundamental (doravante F0) da onda sonora, isto é, o número de vezes que as pregas vocais completam um ciclo de vibrações durante um segundo. Quanto maior o número de ciclos de vibrações das partículas, mais “alta” é a F0 e mais agudo é o som. Esses ciclos são controlados pelos músculos da laringe que determinam a tensão nas pregas vocais, como também por forças aerodinâmicas do sistema respiratório sublaríngeo (LUCENTE, 2008). De um ponto de vista articulatorio, quanto mais delgadas são as pregas vocais, maior número de vibrações. A F0, então, é medida em Hertz (Hz) e a sua percepção pelos falantes é chamada de *pitch*, ou seja, é a interpretação pelo cérebro do fenômeno físico;

- **Volume sonoro:** forte ou fraco, associa-se ao parâmetro físico da intensidade acústica, com dimensão vertical. Pode ser diferenciado também por mais alto ou mais baixo; no entanto, é importante ressaltar que são esses os termos técnicos para a F0. Decorre da amplitude da onda sonora (distância entre a pressão zero e a pressão máxima da onda). Relaciona-se à energia transportada pelas partículas e, conseqüentemente, à sensação auditiva do som;

- **Duração:** com dimensão horizontal, diz respeito ao tempo de articulação de um som, o que varia conforme à velocidade da elocução e influencia o ritmo de cada língua (MATEUS, 2004).

F0 e *pitch*, bem como intensidade e volume, e duração física e percebida representam a relação entre aspectos fonéticos e fonológicos, ou físicos e psicofísicos (LUCENTE, 2008). Quanto a técnicas de registro de *pitch*, Gussenhoven (2004) afirma que ouvintes não dispõem de conceptualizações e vocabulários apropriados para reportar suas sensações, e são comumente incapazes de dizer até mesmo se uma dada mudança de *pitch* representa um aclave ou declive (subida ou queda). Uma simplificação na análise ocorre quando as partes do sinal da fala são divididas em seções de 30 milésimos de segundo, e julgamentos estáticos do *pitch* são obtidos para cada uma das seções, registrados em uma escala de baixo a alto.

Os laboratórios para esse procedimento (julgamento de *pitch*) são ainda instáveis no modo impreciso pelo qual as pessoas reportam suas sensações de *pitch* – uma inconveniência que pode somente ser superada com a ajuda de um cuidadoso desenho experimental permitindo cálculos médios entre as experimentações. Um modo preciso para se medir a vibração das pregas vocais envolve dois eletrodos colocados na pele ao lado da laringe, de forma que as pregas vocais encostem neles. Todavia, tal método só

pode ser utilizado em condições laboratoriais. Escalas seguras usadas em medições usualmente são: de 75 a 400 hz para a fala de homens; de 100 a 600 hz para a fala feminina; e crianças podem produzir mais de 600 hz. É importante comentar que a lacuna de estudos em prosódia muito se deve à dificuldade de analisar, perceber, transcrever e medir as informações entoacionais.

A articulação dos segmentos é um fator que interfere na vibração das pregas vocais. Como ilustra Gussenhoven (2004), não há vibração, por exemplo, durante a produção de glotais fechadas ou consoantes desvozeadas como [p, x]. Oclusivas como [b, d, g] podem impedir a passagem do ar necessária para manter as pregas vocais vibrando. Assim, os contornos de *pitch* diferem dependendo das consoantes. Por outro lado, vogais altas serão pronunciadas com taxas de vibração mais elevadas. O componente associado a essa correlação é a frequência intrínseca. Há diferenças também em sílabas acentuadas e não-acentuadas. O autor ainda comenta que, no inglês, os finais dos enunciados em consoantes sonoras (como m, r, l, w) ou vogais podem terminar com uma reversão da F0 na última parte do enunciado. Seria razoável assumir que isso ocorre devido ao relaxamento dos músculos de controle da frequência de vibração das pregas vocais.

Para um exemplo de como há uma interação entre o que é físico, o que é produzido, e o que é percebido no que tange ao som, recuperamos aqui a fala de Välimaa-Blum (2005), que, ao tratar do acento nas palavras, propõe a seguinte tabela:

A. Correlato fisiológico	B. Correlato perceptual	C. Correlato acústico
a. Duração dos gestos articulatorios	Duração do som	Tempo do sinal acústico
b. Fonação	<i>Pitch</i>	F0
c. Esforço físico	Volume	Intensidade

A autora, ao utilizar essa tabela, explica os aspectos básicos do acento. Na coluna A, encontramos o evento fisiológico ocorrido durante a produção da sílaba acentuada. Na coluna B, como consequência do evento fisiológico, o sistema perceptual

recebe e interpreta o som. Na C, temos o correlato acústico gerado pelo evento de A e recebido em B, ou seja, a onda sonora.

Há ainda outros fenômenos prosódicos, bem como propriedades da dinâmica da voz, mas que nesse momento não serão descritos. Segundo Barbosa (2006), como o parâmetro fonético-acústico que controla diretamente a sensação de *pitch* é a F0 (o correlato acústico da frequência de vibração das pregas vocais), os estudos da entoação privilegiam a análise da curva de F0 ao longo dos enunciados.

É interessante como o conceito de entoação é abordado de diferentes maneiras, conforme diferentes autores. Neste trabalho, focamo-nos na ideia de entoação relacionada aos três parâmetros acústicos e suas sensações, principalmente a sensação do *pitch*. Além disso, esclarecemos que estamos tomando, ao empregar “prosódia” e “suprasssegmento”, um termo pelo outro, apesar de sabermos que o segundo pode ser empregado de várias maneiras.

Quando falamos de prosódia, podemos estender o assunto a vários domínios: morfemas, sílabas, palavras, sentenças, enunciados etc. Por exemplo, uma sílaba pode tornar-se mais proeminente em uma palavra e sentença quando está relacionada à combinação de parâmetros prosódicos como *pitch*, volume e duração. Da mesma forma, a prosódia pode distinguir orações relativas restritivas, pode comunicar ênfases e atitudes dos falantes; enfim, pode estabelecer aspectos expressivos do discurso que, muitas vezes, não são claramente transferidos através da escrita no papel.

Um detalhe interessante associado à prosódia é que ela pode também mapear a identidade de grupos sociais na fala. Recordamo-nos agora de um curioso exemplo oferecido por Cruttenden (1997) que apresenta aspectos sociolinguísticos da variação entoacional. Tais aspectos incluem os dialetais, os de classe, estilo, sexo etc. É bastante claro que, conforme a situação, os movimentos prosódicos podem variar. Podemos pensar na fala de políticos, de jornalistas, de apresentadores de programas de televisão e na comparação desses com os apresentadores de rádio, por exemplo. Para um novo tópico, na fala destes últimos, há uma tendência predominante de movimentos de acentos prosódicos. Não é difícil aceitar ademais que, em um país de uma extensão tão grande como o Brasil, haja variações dialetais fortemente marcadas pela prosódia. Em menor alcance, pensemos no estado de Minas Gerais e seus diferentes dialetos. A prosódia, certamente, não deixa de assumir uma importante contribuição nessas mudanças.

A prosódia interfere, de alguma forma, na identificação lexical, expressando proeminência, ênfase, fronteiras, não-finalidade etc., assim como atua em um grande número de significados vagamente compreendidos de estruturas de diálogos, atos de fala etc. (HIRST, 2005).

Estuda-se com frequência se há aspectos universais relacionados à prosódia. Hirst (2005) comenta que a tendência tida como universal de se associar movimento de declive de *pitch*, com declarações, e aclives finais, com perguntas, é contradita por algumas línguas que não usam sistematicamente uma subida final para as interrogativas.

Ohala (1994) comenta sobre evidências experimentais que mostram que, em geral, podemos encontrar similaridades entre culturas quanto à produção de *pitch* que sinaliza afeto, intenção ou emoção. Bolinger (1978) anteriormente já havia anunciado que uma *F0* alta ou em subida está associada a estados ou sentimentos como submissão, polidez e insegurança; assim como uma *F0* baixa ou em descida, à autoritarismo, agressão e confiança. Para se chegar a essas conclusões, muito se apoiou em dados de fala coletados de rádio, televisão etc. No entanto, contra-exemplos puderam ser detectados, os quais colocam em cheque tais afirmações. Scherer (1973), *apud* Ohala (1994), apresentou exemplos de *F0* alta atrelada a demonstrações de confiança. Diferentes processos experimentais revelaram que *F0* baixa faz a voz parecer mais dominante e se liga à ideia de algo grande. Já o tom alto alude ao que é pequeno, diminutivo, familiar, próximo, estreito.

Uma correlação sistemática entre *F0* e significados como esses é encontrada nos sinais vocais de outras espécies. No ambiente dos animais, por exemplo, os sons produzidos por um agressor confiante (ou um que pareça dessa maneira) tipicamente apresentam uma *F0* baixa. Já o choro de indivíduos submissos ou assustados tem uma *F0* alta, como podemos perceber em latidos de cães. É interessante ver como os animais, quando se valem dos gestos corporais na comunicação, para ameaçar o adversário, usam de recursos como aumentar o tamanho, espichando-se e alongando-se, para parecerem mais fortes e robustos. A *F0* da voz pode também, indiretamente, transportar uma impressão de tamanho, o que ilustra a correlação corpo e emissão de som. A *F0* é inversamente relacionada à massa da membrana que vibra (corda vocal), que é relacionada com a massa corporal como um todo. Para dar a impressão de ser maior e mais perigoso, o animal deve produzir uma vocalização com uma *F0* mais baixa possível. Por outro lado, para parecer pequeno e não ameaçador, é preciso uma *F0* alta.

Essa correlação entre F0 e significados entre as espécies é conhecida na literatura como código de frequência.

Pensando agora a partir do ponto de vista informacional, no caso das perguntas e assertivas, quem pergunta está pedindo por ajuda e quem afirma está se mostrando mais auto-suficiente (sabemos que isso não pode ser aplicado a todos os casos). A produção da F0 é um gesto que acompanha ou sobrepõe a mensagem linguística, a fim de realçar, elaborar ou, em muitos casos, contradizer seu significado. A descida ou subida de F0 pode ocorrer paralelamente, em alguns falantes, à descida ou subida das sobranças. O referente da palavra pode também ser sinalizado pelo som (pequeno ou grande).

E mais uma vez, mencionamos aqui a nossa hipótese de que, conforme Reich (2009), curvas entoacionais tendem a mostrar um declínio fonético devido ao desgaste de pressão subglotal ao falar. Veremos ainda quais associações podemos sugerir conforme os resultados do nosso estudo prospectivo da F0 nos dados. Adiante, trataremos especificamente de uma teoria que contribuiu, de forma significativa, para o desenvolvimento das ideias do presente trabalho.

2.5 Sistemas dinâmicos e cognição prosódica

Palavras como “ajuste”, “controle” e “coordenação” são importantes para refletirmos acerca de como funciona nossa fala. Ajustamos, controlamos e coordenamos nossas ações relacionadas, por exemplo, à glote ou às pregas vocais para produzirmos contrastes sonoros que geram sentido. Acreditamos que nossas pregas vocais produzem um tipo específico de ação que vamos perceber como o *pitch*. Estas são ideias importantes para interagirmos com a teoria sobre a qual seguiremos falando.

A proposta central de Raczaszek-Leonardi e Kelso (2007) é caracterizar a linguagem como um sistema dinâmico. Nessa perspectiva, explicações das estruturas linguísticas devem incluir aspectos dinâmicos envolvidos no uso da língua, como os processos ocorridos na comunicação, aquisição e evolução da linguagem. Acredita-se que informações do sistema biológico são úteis para a investigação da linguagem. É possível, com a leitura desses e outros autores como Barbosa (2006) e Meireles (2007), realizarmos relações que nos levam diretamente a muitos dos postulados condizentes com as premissas da LC. O que procuramos cumprir na primeira parte desses pressupostos teóricos é exatamente a tarefa de expor os aspectos fonéticos e fonológicos da língua a princípios de categorização, uso, restrições físicas, questões culturais, de aquisição e evolução da linguagem. Somos movidos, então, a uma missão de acoplar teorias e perspectivas que, embora a princípio distintas, carregam objetivos e preceitos compatíveis.

Na composição da teoria que encara a linguagem como um sistema dinâmico, são comentados alguns trabalhos relacionados aos sistemas biológicos, como os de Patee (1997). Na explicação de algumas características dos organismos vivos, é bastante comentado o conceito de controle. Tal conceito envolve alguns graus limitados de liberdade, deixando outros não restritos, resultando em uma ação coordenada, porém flexível.

Raczaszek-Leonardi e Kelso (2007) ilustram, de maneira simplificada, como se dá uma ação coordenada: basta olharmos para nossa mão e fazermos movimentos ondulatórios com nossos dedos. Trata-se de um sistema com alguns graus de liberdade restritos pela estrutura física dos nossos ossos, juntas e músculos. Quando uma ação é realizada, como pegar uma caneca, alguns movimentos têm que atuar juntos de uma

maneira organizada espacial e temporalmente. É importante compreender que algumas limitações impostas aos graus de liberdade das mãos não são especificadas de uma só vez e para sempre. A ação e a relação com os graus de liberdade têm de ser adaptadas, *on-line*, a demandas reais do ambiente, tais como as dimensões e os formatos da caneca, a temperatura do líquido contido na caneca etc. Outro fato importante ainda é a especificação da ação para formar diferentes situações. Assim, alguém pode segurar uma caneca de diferentes modos: por exemplo, com a palma da mão virada para fora, segurando uma caneta etc. Assim sendo, a flexibilidade da ação está diretamente ligada aos graus de liberdade.

Os organismos são partes ativas do ambiente, e a informação que eles extraem é especificada pelas demandas de sobrevivência e controle no ambiente. As mudanças no sistema podem ser fonte de informação quanto à natureza das forças que sustentam a estrutura da língua. Os autores defendem que, fora de contexto, símbolos não têm significado. Um símbolo “usa” todos os fatores presentes na situação, e o que um símbolo “faz” é sempre relevante a um momento particular. A não-determinação prévia do significado “transportado” pelo símbolo revela o poder da linguagem, pois remete à fonte da sua eficiência (BARWISE; PERRY, 1983 *apud* RACZASZEK-LEONARDI; KELSO, 2007). Em suma, entender um símbolo deve implicar um processo dinâmico.

Enfatiza-se que a história evolucionária dos símbolos e estruturas linguísticas pode não ser direta: eles podem ser “emprestados” para desempenhar uma função diferente ou podem aparecer como subprodutos de outras adaptações (RACZASZEK-LEONARDI e KELSO, *op. cit.*).

O fato de que expressões linguísticas podem, em princípio, ser descritas por um sistema de regras não significa necessariamente que essas regras existam em outro lugar, e sim nos olhos e mente do observador. Comportamo-nos de acordo com regras sem nem mesmo conhecê-las. O exemplo dado é o de um motorista que, ao se aproximar de um cruzamento pode parar, em virtude da sua visibilidade restrita, querendo estar seguro de que ninguém vem vindo. Ao observador, esse comportamento pode parecer como se o motorista estivesse obedecendo a uma placa de “pare”. Contudo, é possível ver que atribuir o comportamento de obedecer à regra falha em apontar unicamente a causa real e a função real de parar, confundindo isso com obedecer a uma regra ao invés de assegurar segurança e sobrevivência. Portanto, podemos falar em regularidade sem precisarmos estabelecer e fixar regras.

A ideia de “coordenação” está, de alguma forma, presente no texto de Clark (2000). Um resumo de como esse autor apresenta seus argumentos a favor da visão de linguagem como atividade em que as pessoas, de maneira coordenada, fazem as coisas será exposto a seguir. O uso da linguagem é realmente uma forma de ação conjunta, que é aquela ação levada a cabo por um grupo de pessoas agindo em coordenação uma com a outra. Como exemplo simples, pense em duas pessoas dançando valsa, remando em uma canoa, executando um dueto de piano ou fazendo amor. Quando João e Maria dançam, cada um dos dois se move no salão de uma maneira especial. No entanto, a dança é diferente da soma das suas ações individuais. Imagine-os dançando os mesmos passos em salas separadas ou em momentos diferentes. A dança é a ação conjunta que emerge à medida que João e Maria executam seus passos individuais em coordenação, como um casal. Fazer coisas com a linguagem é, da mesma maneira, diferente da soma de um falante falando e de um ouvinte ouvindo. Trata-se da ação conjunta que emerge quando falantes e ouvintes – ou escritores e leitores – desempenham suas ações individuais em coordenação, como um conjunto. O uso da linguagem, portanto, incorpora tanto processos individuais quanto processos sociais. Falantes e ouvintes, escritores e leitores, devem executar ações na capacidade de indivíduos se quiserem ter sucesso em seu uso da linguagem. No entanto, devem também trabalhar juntos, como participantes nas unidades sociais que Clark tem chamado de conjuntos. João e Maria tanto desempenham ações individuais, movendo seus corpos, braços e pernas; quanto desempenham ações conjuntas, coordenando esses movimentos, à medida que criam a dança. Em alguns campos, o uso da linguagem tem sido estudado como se fosse inteiramente um processo individual, como se ele coubesse totalmente dentro das Ciências Cognitivas – Psicologia Cognitiva, Linguística, Ciência da Computação, Filosofia. Em outros campos, ele tem sido estudado como se fosse um processo inteiramente social, como se ele estivesse inteiramente dentro das Ciências Sociais – Psicologia Social, Sociologia, Sociolinguística, Antropologia. De fato, já podemos admitir que ele pertence a ambos, sendo as ações conjuntas construídas sobre ações individuais.

Prosseguindo com os exemplos de Clark, quando José se dirige a Ana, ele é o falante e ela o destinatário. José está falando com o propósito de fazer com que Ana entenda e com que ela venha a agir com base em tal entendimento. No entanto, ele sabe que não poderá ser bem sucedido a não ser que ela execute as suas próprias ações. Ela deve prestar atenção nele, ouvir suas palavras, perceber seus gestos e tentar entender o

que ele diz no exato momento em que ele está falando. Ana sabe disso tudo. Assim, José e Ana não agem independentemente. Não se trata apenas de que eles executam ações levando em conta um ao outro, como também de que eles coordenam essas ações um com o outro (CLARK, 2000). O sinal de José consiste de sua elocução, seus gestos, suas expressões faciais, seu direcionamento de seu olhar e, talvez, de outras ações; e Ana identifica essa composição ao entender o que ele quer sinalizar. Além disso, eles não vão conseguir coordenar significado e entendimento sem referência à sua base comum.

Muitas ações conjuntas apresentam participantes fazendo coisas dessemelhantes. Um motorista que se aproxima de um cruzamento coordena-se com o pedestre que tenta atravessar a rua. Uma bailarina dançando coordena-se com a orquestra que a acompanha (CLARK, 2000). Falar e ouvir não são independentes uma da outra. Ao contrário, são ações partícipes, como as partes de um dueto, e o uso da linguagem que elas criam é uma ação conjunta, como o próprio dueto. Enfim, no uso da linguagem, é importante não confundir produtos previstos com produtos emergentes. Muitas das regularidades que se presumem como previstas ou pretendidas apenas emergiram.

No trabalho de Kelso, é feita uma menção aos trabalhos de Lakoff e Johnson (1999) e Langacker (1982), para quem regras que governam co-ocorrências de símbolos são descrições de regularidades encontradas em habilidades cognitivas mais gerais para estruturar o mundo. Tal visão pode ser considerada compatível com a proposta de postular a linguagem como um sistema dinâmico e encontramos aí já sinalizada uma grande aproximação entre as ideias do autor mencionado e dos autores em LC.

A perspectiva da Teoria dos Sistemas Dinâmicos (doravante TSD) despontou com estudos sobre coordenação e controle motor. A abordagem seguiu inserida em domínios usualmente considerados como mais cognitivos, tais como os de percepção, ação-percepção, reconhecimento de padrões, aprendizado, desenvolvimento motor e cognitivo e coordenação interpessoal. Enfim, o objetivo é identificar mudanças subjacentes no sistema de padrões perceptuais, cognitivos e comportamentais relativamente estáveis. Essa teoria visa a produzir um modelamento que tenha recurso para testar sistematicamente previsões relacionadas à estabilização e à mudança de comportamento em vários níveis de descrição. A ideia é a de que podemos aprender mais sobre um sistema quando o observamos em situações de mudança, isto é, quando seus padrões perdem a estabilidade e novos padrões emergem.

Relações gramaticais, nessa visão, são moldadas dinamicamente por relações entre atratores², expressando dependências presentes nas cenas perceptuais (PETITOT, 1995 *apud* RACZASZEK-LEONARDI; KELSO, 2007)). O trabalho de Goldberg (1991) é citado dentre os estudos em semântica de relações gramaticais. Têm sido realizados estudos relacionando pistas prosódicas e significados de sentenças. O trabalho de Tomasello também é lembrado por chamar a atenção para o papel da linguagem em cenas de atenção conjunta e para a percepção do outro no ambiente social.

A diferença cultural aponta para estruturas linguísticas distintas, conforme estudos dessa abordagem. Símbolos não são estáticos, e sim flexíveis. As pressões para controle interpessoal de um lado, e as capacidades individuais perceptuais de outro representam as duas forças que fazem da linguagem um sistema auto-organizado.

A teoria que prevê a linguagem como um sistema dinâmico tem como questão-guia como os humanos percebem, aprendem e coordenam comportamentos complexos. Temos, então, a acomodação do cérebro, mente e comportamento em um só programa de pesquisa: o dos sistemas dinâmicos auto-organizados. Tal auto-organização refere-se à formação espontânea de padrões e a mudanças de padrões em sistemas abertos que não estão no estado de equilíbrio. Nessa abordagem, não são colocadas em evidência dicotomias como competência/desempenho ou língua/fala. Além disso, forma e conteúdo estão intrinsecamente ligados e não podem ser separados, sendo os símbolos formados dinamicamente por atividades cerebrais, não-inatos e variantes no tempo.

Meireles (2007), apresentando de forma bastante clara a teoria dos sistemas dinâmicos, dá o exemplo da narração do jogo de futebol, ilustrando como um sistema pode ser alterado. A transmissão das jogadas, pois, é realizada por meio do recurso sonoro. O atrator é algum evento singular no campo, o qual motiva a narração do locutor e estabelece momentos de mudança de estratégia ou estilo na narração, como um grande drible, uma jogada rápida, um cruzamento de longa distância etc. Temos uma bifurcação no sistema quando ocorrem mudanças bruscas de estratégia de narração. Por exemplo, ao ocorrer determinado evento, como um gol, ocorre uma mudança qualitativa do atrator. Tem-se, dessa maneira, uma bifurcação da amplitude do movimento da mandíbula, resultando em um primeiro momento, na paralisação do movimento (goooooooooooooooool). O deslocamento da mandíbula do locutor representa, portanto, um

² Atratores são tendências ou preferências do sistema quando sujeito à perturbação externa.

efeito coletivo ou cooperativo do sistema. Movimentos motores como esse podem ser explicados segundo a auto-organização da teoria dos sistemas dinâmicos.

Um conceito bastante usado nessa teoria é o de sinergias. Trata-se de unidades funcionais que agem flexível e temporariamente para o desempenho de uma determinada tarefa. Ainda reportando a fala de Meireles, para produzirmos uma simples sílaba como [ba], fazemos uso de aproximadamente 36 músculos. Um experimento, no sentido de perturbar a produção silábica, permite observar a reorganização instantânea dos articuladores. O resultado alcançado é o de que assim que a mandíbula é bloqueada por alguns milissegundos em dado momento da sílaba, os lábios inferiores e superiores imediatamente compensam a perda do movimento, embora não tenha havido nenhuma movimentação no movimento da língua. A cooperação se mostra, enfim, bastante flexível.

Um ponto de inquestionável importância é como os falantes humanos, a partir de um conjunto enorme de possibilidades, conseguiram obter um padrão sonoro distintivo e bem formado? A discussão crucial aqui gira em torno da existência de padrões na língua. As estruturas coordenadas são definidas dinamicamente e unificam ambos os níveis articulatórios abstratos e os de produção da fala. A forma mais comum para descoberta de padrões na TSD é a utilização de um elemento perturbador do sistema para se observar a emergência de regularidades. No caso da fala, o perturbador mais comum é a variação da taxa de elocução. Constatou-se, por exemplo, que o padrão CV é mais estável que VC e mais recorrente nas línguas do mundo e mais utilizado pelas crianças na produção das primeiras sílabas. Alterando a taxa de elocução na produção da sílaba “ip”, o som produzido acaba se modificando para “pi”.

Meireles (2007) mostra a contribuição da interlocução da TSD com a Fonologia Articulatória. Tal vertente de estudo representa, segundo o autor, a aplicação da TSD à fonologia. Uma ponte entre as estruturas fonológicas e a estrutura física da linguagem é feita com a introdução da noção de gesto linguístico. Assim, são trazidos à tona aspectos mentais e físicos da produção linguística, como o controle da fala. Os gestos são organizados entre si, através de coordenações temporais, chamadas de constelações gestuais, para formar as estruturas conhecidas tradicionalmente como segmentos, sílabas, palavras e enunciados maiores. Os itens lexicais são contrastados pela presença ou ausência de determinado gesto (“há” vs “dá” – presença de um segmento / “s” vs “sim” – traço da nasalidade). A unidade mínima de pesquisa passa a ser, então, o gesto articulatório, e não o traço distintivo.

Comentando acerca dos processos relacionados aos sons da fala, Meireles afirma que, apesar de o controle motor da fala ser, em sua maior parte, um processo cognitivo, ele sempre resulta no deslocamento de objetos físicos reais (língua, mandíbula etc). Produção e percepção da fala podem ser considerados processos semelhantes.

Podemos vislumbrar pontos de interseção entre a TSD e a LC, pois ambas buscam a integração da linguagem com aspectos sociais, psicológicos, biológicos e culturais. Um dado interessante do trabalho de Meireles é o estudo do ritmo, ao considerar a organização do movimento humano como, por exemplo, o ritmo do coração e da respiração. Com isso, o autor sugere como os ritmos biológicos têm uma grande repercussão na vida cotidiana. Nas palavras dele, a criação de agrupamentos rítmicos é uma característica fundamental de nossa percepção no mundo, lembrando que restrições são inerentes ao aparato fisiológico e perceptual do falante.

A TSD é capaz de descrever e explicar dados da filogenia e ontogenia da produção e percepção da fala, conforme Madureira comenta na apresentação do livro de Plínio Barbosa, *Incursões em torno do ritmo da fala* (2006). Para explicar os mecanismos que subjazem à produção do ritmo da fala, é preciso levar em conta a interação entre o indivíduo e o ambiente que o cerca (formado por objetos e outros indivíduos). Essa interação, vista como essencialmente dinâmica, diz respeito a mudanças de estado que co-evoluem: a mudança que o falante/ouvinte provoca em seu ambiente, e aquela que o ambiente provoca no mesmo falante/ouvinte. O autor completa dizendo que a interação dinâmica entre as partes do sistema que subjaz o ritmo da fala faz com que o sistema evolua adaptativamente numa certa direção através de um princípio de auto-organização. E ainda, como qualquer sistema natural, adaptar-se consiste em buscar estratégias para a realização da tarefa, mesmo que ocorram modificações nas condições iniciais.

Barbosa cita Macneilage (1998), cuja explicação ontológica e filogenética prevê que a propriedade organizacional da fala é caracterizada pelo movimento cíclico e alternante de abertura e fechamento da boca, movimento que especifica, a cada ciclo, a sílaba. Essa pode ser derivada de ciclos de oscilação mandibular presentes nos humanos desde o balbucio. Tal desenvolvimento teria se dado quando oscilações mandibulares relacionadas à ingestão (atos de mastigar, sugar e lambe) tomaram significado comunicativo, como cliques realizados com os lábios e a língua, também presentes em primatas não-humanos. A partir do movimento de abertura e fechamento mandibulares, da ordem da grandeza da sílaba, deve ter se desenvolvido, por necessidade

comunicativa, a independente diferenciação dos segmentos, que se dá aos poucos, mesmo na aquisição da linguagem na criança.

Outro aspecto destacado por Barbosa (2006) é o de que a linguagem e a ação, através de um processo evolutivo, possibilitam que o cérebro reconheça novos gestos como compostos de ações conhecidas. O exemplo que Roy e Arbib (2005) oferecem sobre esse aspecto é o da extração e aplicação de regras sintáticas em palavras conhecidas derivadas de um sistema de espelho, reconhecendo e gerando novas sentenças. Essa menção que Barbosa nos faz remete à questão dos neurônios espelho.

Esses neurônios foram inicialmente identificados na Itália, por volta dos anos 1980 e 1990, quando pesquisadores realizavam testes em cérebros de macacos. Foi constatado que, quando um sujeito observa outro fazendo uma ação, o primeiro está automaticamente simulando a ação. Uma vez que a ação e a simulação utilizam o mesmo substrato neural, os mesmos neurônios são ativados durante a ação e a observação - os neurônios espelho são a explicação para isso. Experimentos comprovaram que macacos, assim como humanos, podem inferir o objetivo de uma ação, mesmo quando a informação visual é incompleta. Esse tipo de inferência pode ser explicado como resultado da simulação da ação por um grupo de neurônios espelho. Os neurônios espelho estão atrelados a uma questão inclusive cultural, visto que envolvem o reconhecimento e a simulação da ação alheia. Isso envolve questões mais complexas, como as da exemplaridade, do direito, da empatia etc.

Tais bases neurobiológicas licenciam o ser humano a travar experiências no sentido de que oferecem equipamento para lidar com a simulação. As metáforas se vinculam diretamente a isso visto que se fundam na experiência e promovem um mapeamento simulador entre domínios mentais distintos. Desse modo, mostraremos brevemente a importância de se atentar para o papel da metáfora como algo sistemático e coerente na nossa experiência e conhecimento conceituais. Pretendemos, durante a análise dos dados, mapear como se dá a cena interativa de uma coletiva de imprensa, pelo menos no que diz respeito à configuração prosódica.

Lakoff (1987) define três tipos de metáforas, a saber, as orientacionais, que fazem referências a orientações lineares (MENOS É PARA BAIXO, MAIS É PARA CIMA); as ontológicas, que projetam características de coisas ou seres sobre algo que não tem essas características inerentemente (A MENTE É UM RECIPIENTE); e as estruturais, as quais estruturam um tipo de experiência em termos de um outro tipo de experiência (COMPREENDER É VIVER). Vale ressaltar que a metáfora “infiltrada”

na cena a ser analisada (a coletiva de imprensa) classifica-se como estrutural. O conceito de guerra estaria por trás da maneira como as pessoas percebem, pensam, agem, se comunicam, enfim, experienciam uma discussão.

Na projeção metafórica, há um mapeamento que liga o domínio-fonte ao domínio-alvo. De acordo com Lakoff e Johnson (1980), a metáfora pode ser considerada como mais ou menos convencional à medida que é automática, livre de esforço e, geralmente, estabelecida como um modo de pensar entre os membros de uma comunidade linguística. Além disso, a metáfora é conceitualmente indispensável ou básica à medida que dispensá-la implica mudar o modo de pensar. Sobre a universalidade e variabilidade da metáfora, Kövecses (2005), *apud* Feltes (2007), sugere:

- (a) experiências universais não necessariamente geram metáforas universais;
- (b) a experiência corpórea pode ser usada seletivamente na criação de metáforas;
- (c) a experiência corpórea pode ser sobrepujada por processos culturais e cognitivos;
- (d) metáforas primárias não são necessariamente universais;
- (e) metáforas complexas podem ser potencial ou parcialmente universais;
- (f) e metáforas podem não ser necessariamente baseadas na experiência corpórea, pois muitas são baseadas em fatores culturais e processos cognitivos de vários tipos (FELTES, 2007, p. 153).

Kövecses entende a metáfora como um fenômeno linguístico, conceptual, sociocultural, neural e corporal. Outros autores como Feldman, Narayanan e Bailey afirmam que, como as experiências corpóreas no mundo são universais, então as metáforas primárias correspondentes são universalmente adquiridas. Não significando que elas sejam inatas, essas metáforas são aprendidas como resultado de um mapeamento conceptual imediato através de conexões neurais, através do nosso movimento de constante percepção no mundo.

De antemão, acreditamos que poderemos estabelecer correspondências entre o domínio da troca de turnos (interação) na coletiva de imprensa e o domínio da guerra:

Domínio 1	Domínio 2
Coletiva de imprensa	Guerra

Entrevistadores	Guerreiros que atacam
Entrevistado	Guerreiro que está sendo atacado
Palavras	Armas
Perguntar	Atacar
Responder	Defender
Conflito de opiniões	Diferentes posições dos beligerantes

Especificadamente com relação à configuração melódica, as correspondências são as seguintes:

aclive da F0	Atacar
declive da F0	Render-se

Na verdade, postulamos o domínio da guerra como domínio-fonte no âmbito da coletiva de imprensa, uma vez que é mais físico e menos abstrato, e oferece uma base experiencial que motiva o alvo específico (discussão, debate, entrevista), sendo que a experiência corpórea resulta em conexões neurais entre áreas do cérebro. Ao falarmos de experiência corpórea, faz-se necessário um apelo aos gestos articulatórios que realizamos ao produzirmos sons. Esses gestos, retomando agora a ideia de coordenação, são como ações associadas a tarefas de mastigar, sugar etc.

Para ilustrar isso, mencionamos a seguir uma imagem em raio x de uma mulher pronunciando a palavra *både*, em dinamarquês; “ambos”, em português. A sequência é um excerto de um filme registrado no Hospital Danderyd, de Estocolmo, em março de 1997. Trata-se do registro de movimentos articulatórios da produção do som, com auxílio de equipamento moderno, que permite a visualização do funcionamento do trato vocal e suas partes (faringe, mandíbula, língua, lábios etc.). As filmagens são de excelente qualidade, o que facilita a interpretação fonética, porém, aos serem transpostos para o papel, os movimentos são congelados. Disponibilizamos o link do site (http://www.ling.su.se/STAFF/ericsson/projects/xray_info.htm), pois demasiadamente interessante observar essa questão gestual, que, por ser imperceptível sem o auxílio do referido equipamento, parece não existir.



Imagem 1 Mulher pronunciando “både” (ambos), em dinamarquês
Fonte: http://www.ling.su.se/STAFF/ericdotter/projects/xray_info.htm

Prosseguindo na discussão sobre metáfora, um dado importante é que o domínio-fonte pode aplicar-se a vários alvos, bem como um alvo pode ligar-se a várias fontes. Pensemos, assim, que o domínio da guerra pode ser aplicado também a domínios como VIDA, dentre outros.

A correspondência entre domínios-fonte e domínios-alvo pode originar expressões metafóricas, sendo o dado linguístico derivado da conexão de dois domínios conceptuais. Daí explica-se a ocorrência de expressões como “x atacou y com os argumentos de que...” ou “os argumentos de defesa” etc. O número de expressões linguísticas que sinalizam dada metáfora é diretamente proporcional à produtividade da mesma.

Metáforas podem ser realizadas através de formas não-linguísticas. As posições em que as pessoas costumam ficar, uma contra a outra, e gestos específicos que são realizados durante discussões, com mãos e braços, são realizações da metáfora DISCUSSÃO É GUERRA.

Dessa maneira, sugerimos que a metáfora DISCUSSÃO É GUERRA, na qual simulamos uma discussão como guerra, torna-se presente na conceptualização do que ocorre na coletiva de imprensa, gênero escolhido para ser analisado. Lakoff e Johnson (1980) mostram que as pessoas frequentemente usam linguagem metafórica. Além disso, a linguagem metafórica não inclui somente o modo como as pessoas falam, mas também refletem exatamente o modo como as pessoas pensam e agem. Os autores oferecem uma teoria do significado baseada na ideia de conhecimento corporificado.

Sobre a metáfora específica a qual tomamos como hipótese de trabalho, Lakoff e Johnson afirmam que conceptualizamos uma discussão (um debate) por meio da metáfora DISCUSSÃO É GUERRA sem que para isso tenhamos que ter experiência

pessoal da guerra, mas porque temos dela imagens mentais ricas. Segundo Silva (1997, p. 8): “como realizações linguísticas desta metáfora conceptual, atacamos ou defendemos determinada ideia ou argumento, tomamos posições e utilizamos estratégias, atacamos cada ponto fraco da argumentação de alguém, demolimos a argumentação do outro etc”. Ressaltamos que, embora consigamos estabelecer esses mapeamentos, considerando entrevistado e entrevistadores em posição de “adversários”, veremos com os dados que nem sempre as projeções são postas de modo tão acirrado como uma guerra, de fato, pressupõe, mostrando o quão flexível e dinâmica a cena interacional pode ser.

3. Procedimentos metodológicos

Uma breve explicação inicial de como se dividem as bases metodológicas torna-se necessária. Anunciaremos inicialmente alguns aspectos teóricos que nos ajudam a fundamentar nossa metodologia. Nosso estudo, além de alcançar dimensões fonéticas, pretende atentar para o aspecto pragmático e interacional do objeto. Por isso, julgamos relevante uma explanação sucinta de questões relacionadas à troca de turnos.

Em seguida, trataremos especificadamente do gênero escolhido para análise, descrevendo informações pertinentes quanto a cada uma das entrevistas selecionadas. E, por último, comentaremos nossos materiais e procedimentos, principalmente no que diz respeito ao uso do programa computacional de leitura acústica.

3.1. A cena interacional e a troca de turnos: “Quando um burro fala, o outro abaixa a orelha”.

Clark (2000) elenca os diferentes tipos de cenários a que podemos fazer referência quando tratamos do uso da linguagem: cenários pessoais (livre troca de turnos entre dois ou mais participantes), cenários não-pessoais (monólogos), cenários institucionais (limitados por regras institucionais), cenários prescritivos (palavras estabelecidas de antemão), cenários ficcionais (interpretação), cenários mediados (presença de intermediários), e cenários privados (pessoas falam em nome próprio, sem se dirigirem a mais ninguém, “apenas para mim”).

Sobre o fato de a conversa face-a-face ser o uso básico da linguagem, o mesmo autor afirma que, para um cenário de uso da linguagem ser básico, ele deve ser universal às sociedades humanas. Isso elimina os cenários escritos, uma vez que sociedades inteiras, bem como grupos de sociedades letradas, dependem somente da palavra falada. Além disso, a maioria das línguas se desenvolveu antes da expansão do letramento. Podemos também eliminar os cenários falados que dependem de tecnologias como o rádio, a televisão, as gravações, uma vez que esses não são exatamente universais. A maioria das pessoas participa de cenários não-pessoais, institucionais e prescritivos apenas raramente, e ainda assim, a sua participação é restrita a certos papéis. A conversa face a face é, portanto, o cenário mais comum de todos e é também o cenário básico para a aquisição da linguagem por parte das crianças. Durante os seus primeiros dois ou três anos de vida, as crianças, em sociedades letradas ou iletradas, aprendem a sua língua quase somente em cenários conversacionais. Mesmo o que elas aprendem com os livros também se dá em cenários conversacionais, à medida que seus tutores lêem em voz alta e verificam se há entendimento. As crianças podem aprender linguagem com outros meios, mas, ao que consta, elas são incapazes de aprender a sua primeira língua somente com o rádio ou a televisão. Na escola, a linguagem dos colegas tem influência no dialeto adquirido, e isso também vem de cenários conversacionais. A conversação face-a-face é, portanto, o berço do uso da linguagem (CLARK, 2000). Assim sendo, os outros cenários prescindem do imediatismo, do meio ou do controle que caracterizam a conversa face a face, exigindo, portanto, técnicas ou práticas especiais.

Um dos motivos que nos levam a estudar prosódia relaciona-se à questão de como a variação e os detalhes fonéticos constroem sentido na fala diária. Acreditamos que a

prosódia diz respeito a um recurso que falantes utilizam para alcançar coisas e gerenciar o fluxo da fala. Segundo Local (2007), detalhes fonéticos codificam níveis da organização linguística de forma que falantes possam localizar o momento preciso de começar a falar. Tais detalhes contribuem também na produção e no entendimento do que está sendo dito, porque aquilo está sendo dito e que função isso tem.

Crystal (1992), de forma bastante leve e bem humorada, fala acerca de como se dão as trocas de turno e as interrupções na fala do dia-a-dia. O texto trata de um diálogo entre os personagens Ned e Ed sobre como, na conversa, altera-se a vez de falar dos participantes. O tom humorado se deve à maneira como um acaba interrompendo o que o outro tenta dizer, embora eles afirmem, ao longo do diálogo, que as pessoas não falam ao mesmo tempo. O texto mostra que, em muitos momentos, o ouvinte se apóia em pistas oferecidas pelo falante para que aquele complete a fala deste, ou mesmo tome o turno e inicie um novo assunto. É importante ressaltar que, além das informações prosódias (as que direcionamos mais atenção neste estudo), estão relacionadas às possíveis tomadas, às sobreposições e aos assaltos a turnos de fala, as informações semânticas, sintáticas e pragmáticas. Assim, a projeção de uma possível completude de turno pode envolver várias pistas textuais e contextuais.

Estudos da conversa mostram que há uma troca de turnos ordenada entre falantes. Para que isso ocorra suavemente, ouvintes devem monitorar a fala para localizarem o momento de começar a falar. Segundo Sacks, Schegloff e Jefferson (2003), a tomada de turnos é usada na ordenação de movimentos em jogos, na alocação de cargos políticos, na organização do tráfego em cruzamentos, no atendimento a clientes em estabelecimentos comerciais e na fala em entrevistas, reuniões, debates, cerimônias, conversas etc. Visto que é usada para organizar tipos de atividades bastante diferentes umas das outras, é de particular interesse ver como os sistemas de tomada de turnos em operação podem ser caracterizados em função de suas adaptações às propriedades dos tipos de atividades nas quais eles operam. Ainda os mesmos autores afirmam que, em qualquer conversa, observamos:

A troca de falante se repete, ou pelo menos ocorre.
Na grande maioria dos casos, fala um de cada vez.
Ocorrências de mais de um falante por vez são comuns, mas breves.
Transições sem intervalos e sem sobreposições são comuns.
A ordem dos turnos não é fixa, mas variável.
O tamanho dos turnos não é fixo, mas variável.
A extensão da conversa não é previamente especificada.

O que cada um diz não é previamente especificado.
A distribuição relativa dos turnos não é previamente especificada.
O número de participantes pode variar.
A fala pode ser contínua ou descontínua.
Técnicas de alocação de turno são obviamente usadas. Um falante corrente pode selecionar um falante seguinte ou as partes podem se auto-selecionar para começarem a falar.
Várias “unidades de construção de turnos” são empregadas; por exemplo, os turnos podem ser de modo projetado a “extensão de uma palavra” ou podem ter a extensão de uma sentença.
Mecanismos de reparo existem para lidar com erros e violações da tomada de turnos; por exemplo, se duas partes encontram-se falando ao mesmo tempo, uma delas irá parar prematuramente, reparando, assim, o problema. (SACKS; SCHEGLOFF; JEFFERSON, 2003. p. 14-15)

Certamente, é também muito importante para a organização da tomada de turnos aquilo que de algum modo se entende como produção do som (fonologia, entoação etc). Por exemplo, distinções entre o “que...?” (*what?*) como uma pergunta de uma só palavra e como o começo de uma construção sentencial (ou clausal ou sintagmática) são feitas não através da sintaxe, mas da entoação (SACKS, SCHEGLOFF e JEFFERSON, 2003). Vamos, a partir de agora, tentar entender como se dá a troca de turnos, especificamente, em uma coletiva de imprensa.

3.2. Contexto de pesquisa: O gênero coletiva de imprensa

Na coletiva de imprensa, uma assessoria convida jornalistas para a divulgação de informações, abrindo espaço para perguntas. Uma coletiva, como é simplesmente conhecida, pode ser agendada com o propósito de responder a questionamentos de vários repórteres de uma só vez. Além disso, trata-se de um evento que atrai atenção de vários meios de comunicação.

Na coletiva de imprensa, jornalistas disputam espaço e reivindicam o turno de fala para entrevistar a fonte da informação. Em encontros como esse, por conta da intensificação do embate pela palavra, supõe-se que os profissionais de imprensa acentuem suas habilidades de percepção prosódica para tentar a tomada de turno e conseguir dirigir perguntas aos entrevistados. Os turnos de fala se alternam entre entrevistado e jornalistas, sendo que a cada término de resposta novas questões são lançadas. A tendência, então, é a de o jornalista se auto-selecionar para lançar questões ao entrevistado.

A primeira coletiva de imprensa estudada no presente trabalho se refere a uma entrevista (agendada e com local pré-definido) da Agência Nacional de Aviação Civil (ANAC) sobre o acidente do voo 1907, da Gol, ocorrido em 29 de setembro de 2006. Os trechos selecionados são respostas do Brigadeiro Leite, representante do Comando da Aeronáutica. A duração deste vídeo é de 09:18 min. Tal entrevista, assim como as outras, pode ser acessada pelo site *You Tube* (<http://www.youtube.com/>), onde vários vídeos são veiculados mundialmente. O endereço eletrônico dessa entrevista específica é: (<http://www.youtube.com/watch?v=JOEJuMzMGPo>).

A segunda coletiva de imprensa selecionada para análise é a do jogador de futebol, Daniel Carvalho (<http://www.youtube.com/watch?v=wKD2SfEzISs>). O clube gaúcho do Internacional apresentou à mídia sua nova aquisição à época, o meia-atacante Daniel, na tarde de quarta-feira (30/07/08), na sala de imprensa do estádio Beira-Rio (Porto Alegre – RS). A gravação tem a duração de 06:55 min. Antecipamos que, para análise e exposição dos dados, aproveitamos somente um trecho de fala do jogador (que concluía sua resposta que antecedeu a tomada de turno por parte dos repórteres), em virtude da baixa qualidade do som para esse tipo de análise. Aproveitamos, então, para utilizar uma outra entrevista, também com um jogador de futebol em situação semelhante. Essa terceira entrevista foi com o zagueiro Álvaro, que veio do futebol

espanhol para jogar no Internacional (<http://www.youtube.com/watch?v=KJvY63KpmQQ>). Os trechos separados foram da coletiva que ocorreu na tarde de sexta-feira (08/08/08), também na sala de entrevistas do Beira-Rio; e a duração do trecho acessado foi de 02:07 min.

A quarta entrevista escolhida, também agendada, foi com a artista Ivete Sangalo (<http://www.youtube.com/watch?v=f3t8y9yy00Q>). A cantora baiana conta aos jornalistas, durante o vídeo, novidades a respeito do seu trio elétrico para o carnaval de 2009. O vídeo foi postado em 11 de fevereiro de 2009 e tem a duração de 22:48 min. Um fato que chama a atenção nessa interação é que, na verdade, as sobreposições existentes são de tomadas de turno que a cantora faz sobre a fala dos jornalistas, e não o contrário, como em alguns momentos ocorreu na entrevista sobre o acidente aéreo.

As quatro entrevistas, em arquivos de áudio, bem como os trechos selecionados para análise acústica estão disponíveis em CD-ROM anexo a este trabalho.

Enfim, nosso corpus é de fala semiespontânea e justamente por esse motivo pode ser de difícil tratamento, com os dados surgindo de forma diversa e imprevista. Contudo, acreditamos que somente nesse tipo de situação comunicativa conseguimos observar fenômenos pouco presentes, por exemplo, na fala de laboratório. Além disso, essa perspectiva se alia totalmente à visão da LC, de estudo da língua em uso e buscando seus aspectos recorrentes e produtivos.

Ressaltamos, porém, que embora reconhecendo a importância de se olhar para um número elevado de ocorrências, este trabalho tem uma pretensão inicial mais qualitativa do que quantitativa. É importante esclarecer isso, pois não queremos parecer destoantes da proposta teórica dos modelos baseados no uso, a que nos referimos nos pressupostos teóricos.

3.3. Materiais e procedimentos

Consideramos válido relatar, neste momento, como se deu todo o percurso de confecção desta dissertação. Uma primeira interação entre orientador e orientanda resultou na seguinte proposta:

O objetivo do trabalho é descrever aspectos prosódicos, cognitivos e interacionais envolvidos no gênero coletiva de imprensa agendada ou não-agendada, evento em que grupos de jornalistas disputam espaço e reivindicam o turno de fala para entrevistar a fonte de informação em locais pré-definidos ou improvisados. Em encontros como esse, por conta da intensificação do embate pela palavra, os profissionais de imprensa acentuam suas habilidades de percepção prosódica, ancorada em processos sociocognitivos, para tentar a tomada de turno e conseguir dirigir perguntas aos entrevistados. Quais seriam estratégias linguísticas envolvidas no momento em que o entrevistado intuitivamente autoriza a inserção da pergunta? Como e quando o jornalista percebe que pode rapidamente lançar uma questão? Por fim, pretende-se explicitar com isso quais são os recursos de linguagem adotados numa cena pontual em que há um acirramento da arena conversacional (ROCHA, 2008).

Diante disso, as primeiras empreitadas se deram no sentido de ir em direção aos dados. O que eles mostram? Há um padrão na configuração prosódica? De que maneira apresentar isso como evidências de processos sociocognitivos?

Nosso primeiro recurso foi, então, o programa PRAAT. Não deixamos de mencionar que a qualidade acústica é um aspecto de grande importância nos estudos prosódicos. Além disso, assumimos também que o programa, por mais bem desenvolvido que seja, é passível de erros.

Quanto ao processo de captação de som, realizamos o *download* das entrevistas, disponíveis no site *You Tube*, e convertemos o arquivo em gravação de mp4, através do programa *Total video converter*. Em seguida, editamos o som, selecionando as partes que julgamos necessárias para análise, por meio do editor sonoro para Windows *Wave Pad*. Finalmente, conduzimos as partes selecionadas ao tratamento no *software PRAAT*, desenvolvido por Paul Boersma e David Weenink (Universidade de Amsterdã). O PRAAT é um programa de *software* livre, disponível na URL [HTTP://www.praat.org](http://www.praat.org).

A análise dos resultados parte das seguintes etapas:

- i) Identificação dos momentos de transição de turno na entrevista;
- ii) Seleção da parte final da fala do entrevistado nesses momentos de transição;
- iii) Submissão dos dados selecionados ao tratamento acústico, a fim de investigar a medição dos parâmetros da frequência fundamental, através do PRAAT;
- iv) Checagem, segundo o gráfico oferecido pelo programa PRAAT, de contornos melódicos quando se encerra o turno do entrevistado;
- v) Checagem se há sobreposições nas falas do entrevistado e entrevistador;
- vi) Análise da correlação da F0 e da tomada de turno;
- vii) Sistematização e tratamento dos dados.

Partimos desse contato inicial com os dados para reivindicar o espaço da prosódia em uma teoria que, mesmo no que diz respeito à fonologia, ainda subfocaliza questões relacionadas aos processos de produção e recepção do som da fala. Fizemos então uma pesquisa teórica entre autores da LC para apresentarmos o estado da arte da FC. Assumimos as propostas da FC, e quanto a teorias em prosódia, buscamos também cumprir um apanhado geral do assunto. Sustentamo-nos na TSD para unir concepções teóricas e formular novas ideias acerca de prosódia no âmbito da LC. Munidos, enfim, do mencionado suporte bibliográfico, e expondo nossa trajetória metodológica, partimos para a análise, buscando dar conta dos dados de fala real e dos achados teóricos, apontando aspectos interacionais, prosódicos e sociocognitivos envolvidos no fenômeno estudado.

Diante do fato de que a interpretação dos gráficos gerados pelo PRAAT é, de certo modo, complexa, aproveitamos a exposição metodológica para oferecer instruções mínimas sobre sua leitura. Abaixo encontramos um exemplo de tela do PRAAT, resultado da análise acústica dos dados coletados para este trabalho. Utilizaremos o recurso da caixa de texto para descrever minimamente partes do gráfico a seguir no alinhamento de três camadas.

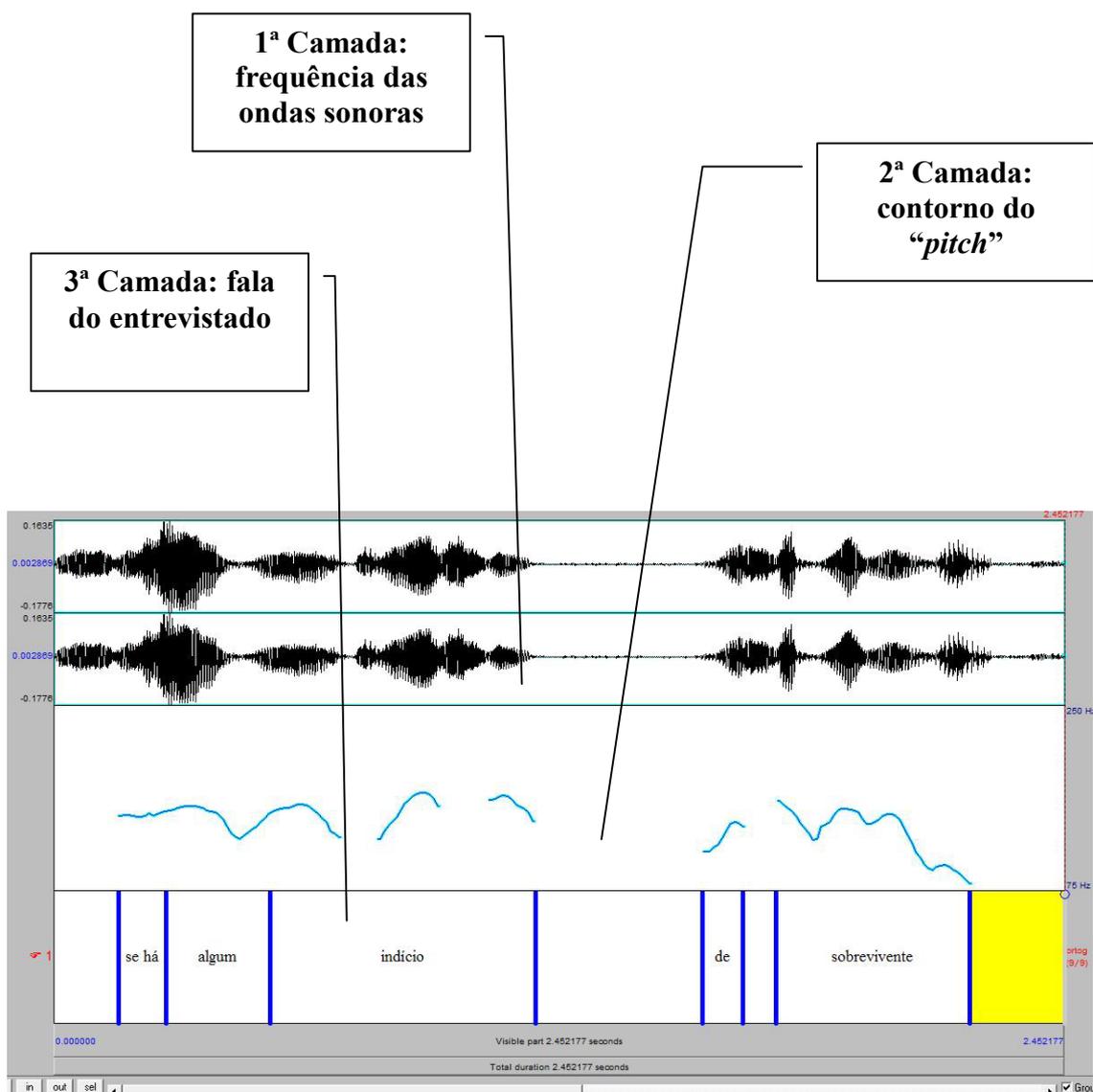


Gráfico 0 – Modelo

Lembramos, por fim, que os ajustes nas medidas de Hz foram realizados pelo programa PRAAT, para que as curvas pudessem ser visualizadas com melhor precisão e definição. A fala do brigadeiro e dos jogadores Daniel e Álvaro ficou ajustada entre 75 e 250 Hz, por se tratar da fala masculina; e embora a fala feminina seja habitualmente ajustada entre 75 e 500 hz, acertamos as medidas entre 75 e 400 hz para o caso da

entrevista de Ivete, visto que a voz dela é mais grave. Esse ajuste de Hz pode ser feito através da ferramenta *pitch settings* disponível no PRAAT.

3.4. Sistema de notação

Para transcrevermos as curvas encontradas, nada mais coerente e plausível do que nos servirmos de um sistema de notação que analisa e modela dinamicamente a fala. Estamos falando do sistema DaTo (*Dynamic Tones of Brazilian Portuguese*), proposto por Lucente (2008). A intenção é realizar uma notação em que sucessivos tons estejam interligados e não considerados distintos como em sistemas anteriores. Além disso, os outros métodos refletiam apenas o componente macroprosódico do sinal da fala, ficando os efeitos microprosódicos de fora, como as interferências na curva de F0 causadas por características inerentes de cada segmento.

Os contornos dinâmicos, como prevê a nova proposta, são divididos em ascendentes (*rising*), tendo como referência para o alinhamento com a sílaba tônica a posição mais alta (H - high), e descendentes (*falling*), tendo como referência para alinhamento a posição baixa (L - low).

Os contornos ascendentes são três:

- a) LH - marca focos³ em posições iniciais e intermediárias de enunciados declarativos. Em posições finais de enunciados interrogativos, este contorno também pode aparecer. O contorno configura-se pela subida da frequência, com a posição alta alinhada com a sílaba tônica;
- b) >LH - apresenta as mesmas características de LH, com a diferença de que neste há um atraso da subida de F0. Assim, a vogal tônica fica alinhada com o movimento de subida e não com o pico;
- c) HLH - neste contorno, há dois picos sucessivos de F0, sendo o primeiro alinhado à sílaba pré-tônica e o segundo, à sílaba tônica.

Quanto aos contornos descendentes, temos:

- a) HL - encontrado frequentemente em fronteira de sentenças declarativas. O contorno configura-se pela descida da frequência, com a posição baixa alinhada com a sílaba tônica, após a subida de F0 na sílaba pré-tônica;

³ “Foco” aqui sendo entendido como destaque ou ênfase a alguma parte da sentença motivada por uma situação particular do discurso.

- b) >HL – distingue-se de HL por representar um movimento de descida atrasado com relação à sílaba tônica. Neste caso, a vogal tônica ocorre em posição mais alta da descida de F0, e não no mesmo nível que a fronteira L;
- c) LHL - aparece em posição final de enunciados declarativos em alternância com o contorno LH. Trata-se de um movimento de subida e descida mais amplo do que acontece em LH. Há uma descida suave formando sucessivos degraus até alcançar a sílaba tônica.

4. Análise dos dados

Esta seção apresenta e descreve os gráficos resultantes da submissão dos dados no programa PRAAT. Assim como explicamos na metodologia, os arquivos completos em formato .wav foram segmentados em trechos menores contendo o exato momento da transição de turnos, ou melhor, anterior à tomada de turno por parte dos jornalistas.

Dividimos a análise em três partes, conforme o contexto das entrevistas que utilizamos, a saber:

- a) Entrevistado *versus* entrevistador
- b) Entrevistado *e* entrevistador
- c) Entrevistado “*entrevista*” entrevistador

Essa divisão tem como base o grau de acirramento interacional entre entrevistado e entrevistador, ora mais combativo, ora menos. Isso se relaciona à gravidade e à complexidade do assunto abordado na coletiva. Sob o domínio da metáfora DISCUSSÃO É GUERRA (LAKOFF; JOHNSON, 1980), tal *frame* interacional instancia cenas distintas em que a dinâmica face-a-face recorre tanto a momentos de imposição de forças discursivas do tipo ataque e defesa explícitos, quanto de negociação cordial de sentidos. No primeiro caso, as perguntas dos jornalistas soam como cobranças diretas de esclarecimentos, e a resposta do entrevistado sinaliza uma dura prestação de contas (entrevistado *versus* entrevistador). Já no segundo, esse jogo de forças é mitigado por uma temática não tão controversa, em que entrevistador e entrevistado mantém relações simétricas (entrevistado *e* entrevistador), que se emolduradas por uma postura descontraída podem até promover a inversão de papéis (entrevistado “*entrevista*” entrevistador).

Mostraremos, em todas as respostas dos entrevistados, dois gráficos, um com o recorte de uma unidade maior de análise, pois julgamos imprescindível olhar esse “todo” para que possamos responder se houve uma subida ou queda da F0 justamente em relação ao contexto anterior; e o outro gráfico revela, de maneira mais detalhada, os contornos melódicos do último momento da fala do entrevistado.

Buscamos responder, ao final da exposição de cada trecho, questões como: no momento que antecede a transição dos turnos, o contorno melódico é ascendente ou descendente? Observando uma unidade maior de análise, como podemos interpretar a

F0 do momento da transição em relação ao material que o precede? Quanto ao aspecto semântico e interacional, trata-se de uma ação conversacional completa dentro do contexto sequencial? E ao encerrar cada uma das três partes da análise, responderemos, quanto ao que ocorre no contexto específico da entrevista, quais projeções metafóricas podemos estabelecer? E o que esse movimento tem a nos dizer sobre o simbolismo do gesto sonoro?

No início de cada seção, apresentaremos breves reflexões acerca da construção de sentido do que discutimos, recorrendo a imagens de momentos da entrevista. Observaremos como, aliado ao gesto sonoro, o gesto físico aparente faz parte integrante do processo da atividade simbólica da linguagem. Olhar essas imagens, conforme nossa perspectiva de trabalho, torna-se relevante, pois acreditamos que nossa cognição é cultural e corporificada, partindo de nossas experiências físicas e capacidades sensório-motoras variadas.

Seguiremos, então, analisando os dados deixando claro que, embora não tenhamos a pretensão de chegar a conclusões definitivas sobre o assunto, pretendemos cumprir um estudo prospectivo, expondo resultados que possam mostrar caminhos interessantes.

A) Entrevistado *versus* entrevistador

O título que encabeça esta seção se justifica por conta de um gerenciamento interacional que leva em consideração o acirramento do debate entre entrevistado e entrevistador em função de uma temática controversa. Nesse sentido, os papéis entrevistado e entrevistador são muito bem marcados e personificam ações de adversários em guerra. Dessa maneira, os lugares relevantes de transição de turno são muito bem monitorados pelos participantes da cena, com vistas ao encaixe de perguntas e respostas. É um cenário em que as habilidades cognitivas atencionais dos jornalistas são redobradas em virtude do grande número de perguntas que surgem em ocasiões dramáticas e polêmicas. Desse modo, como o entrevistador, em questão de segundos ou milésimos de segundos, percebe que o momento oportuno para fazer a pergunta chegou?

A sinalização prosódica é um componente importante nesse caso, porquanto o entrevistado busca manter a palavra “evitando” novas perguntas e, possivelmente, novos confrontos; além de o entrevistador estar antenado na melodia da fala do entrevistado com o intuito de propor uma nova questão. É importante frisar que a tarefa central deste trabalho é analisar os movimentos de F0 que licenciam o *pitch* captado pelos participantes da cena durante a percepção de que o turno pode ser tomado. Ilustrando esse contexto de entrevista, temos a coletiva com o Brigadeiro Leite, representante do comando da Aeronáutica, sobre o acidente do vôo da Gol.

Como anunciado, copiamos algumas imagens do site You Tube para ilustrar o que discutimos no sentido de ofertar comentários sobre o espectro semiológico dos dados. Queremos mostrar, com isso, como nosso corpo também, além do som produzido pelo gesto vocal, oferece pistas para interpretarmos o sentido de uma cena interacional. Evidentemente, não se trata de uma análise aprofundada da imagem e sim do acréscimo de alguns comentários finais.



Imagem 2: o brigadeiro encerrando sua resposta à primeira pergunta dos jornalistas.

Fonte: <http://www.youtube.com/watch?v=JOEJuMzMGPo>

Queremos chamar atenção para o detalhe da quantidade de microfones em volta do brigadeiro, o que o faz parecer até acuado frente aos repórteres. Outro ponto de realce nessa entrevista foi que, em muitos momentos, tornou-se necessário um pedido de ordem e silêncio entre os presentes para que o entrevistado pudesse responder a questão. Como já foi mencionado, o clima dessa entrevista era tenso e notava-se muita exaltação por parte dos jornalistas. O brigadeiro, por sua vez, precisava manter o controle da difícil situação respondendo as questões de maneira a se proteger. Assim, os contornos ascendentes de F0 da fala dele se associam a movimentos defensivos, diferentemente das subidas de F0 na fala dos outros entrevistados.

A seguir, analisamos o primeiro exemplo que emblemiza as reflexões acima.

FRAGMENTO A.1:

“para que nós possamos identificar se há algum indício de sobrevivente”

Nesse primeiro fragmento, temos a marcação dos contornos de F0 do trecho final da resposta do brigadeiro à primeira pergunta ocorrida no evento. O falante, nesse momento específico, conclui sua resposta acerca da possibilidade de existirem sobreviventes ao acidente. Ele anuncia, com pesar, que a equipe de resgate não conseguiu completar o circuito em volta dos destroços “para que pudessem identificar se há algum indício de sobrevivente” (fragmento A.1).

Quanto ao aspecto semântico, a fala do brigadeiro parece tratar de uma resposta completa à questão, apresentando a idéia inicial (ainda não era possível comprovar os sobreviventes) e a justificativa (o resgate ainda não havia completado sua tarefa). Isso quer dizer que, no cenário da coletiva de imprensa (que categorizamos como relativamente prescritivo por conta das perguntas previamente formuladas), esse momento de transição pode ser interpretado como encerrando uma ação conversacional completa nesse contexto sequencial.

De outro modo, é preciso notar que a oração de finalidade do fragmento A.1, do ponto de vista das construções de estrutura argumental, configura-se como apenas uma parte de um período subordinado que exigiria a oração principal para ser entendida como uma construção gramatical de finalidade. Embora a oração principal não seja contemplada na análise acústica deste caso, ela compõe a explicação dos aspectos semânticos e interacionais da cena, como observado no parágrafo anterior.

O movimento prosódico de A.1, definido na unidade maior de análise, mostra o contorno descendente, como se vê a seguir:

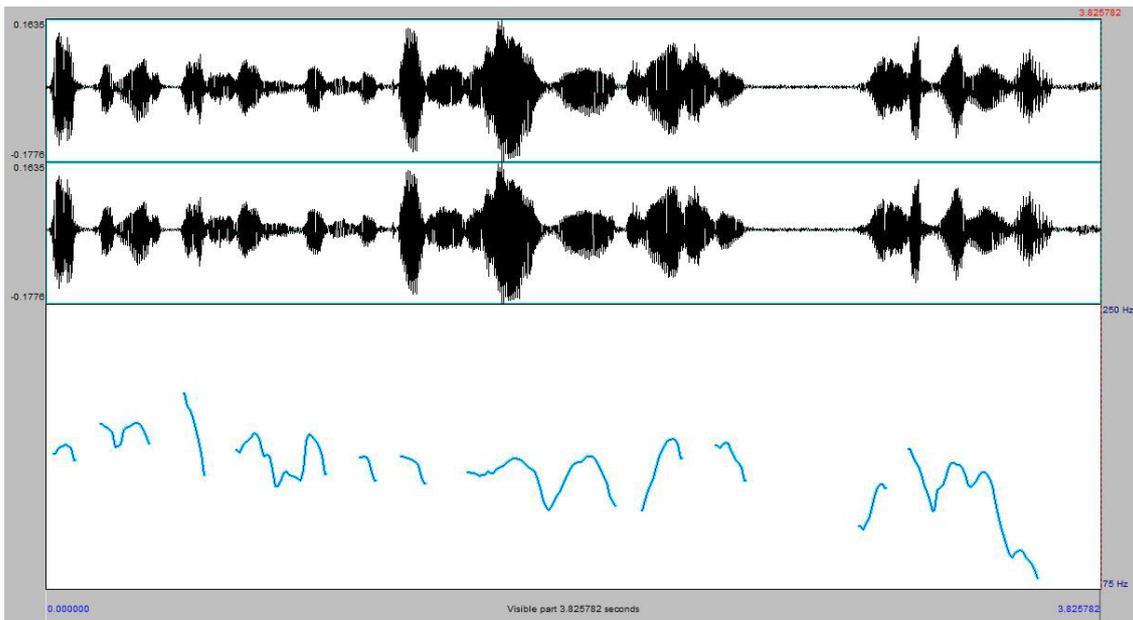


Gráfico 1 Fragmento A.1: “para que nós possamos identificar se há algum indício de sobrevivente”

Se definirmos uma unidade menor de análise, o gráfico apresenta a seguinte configuração:

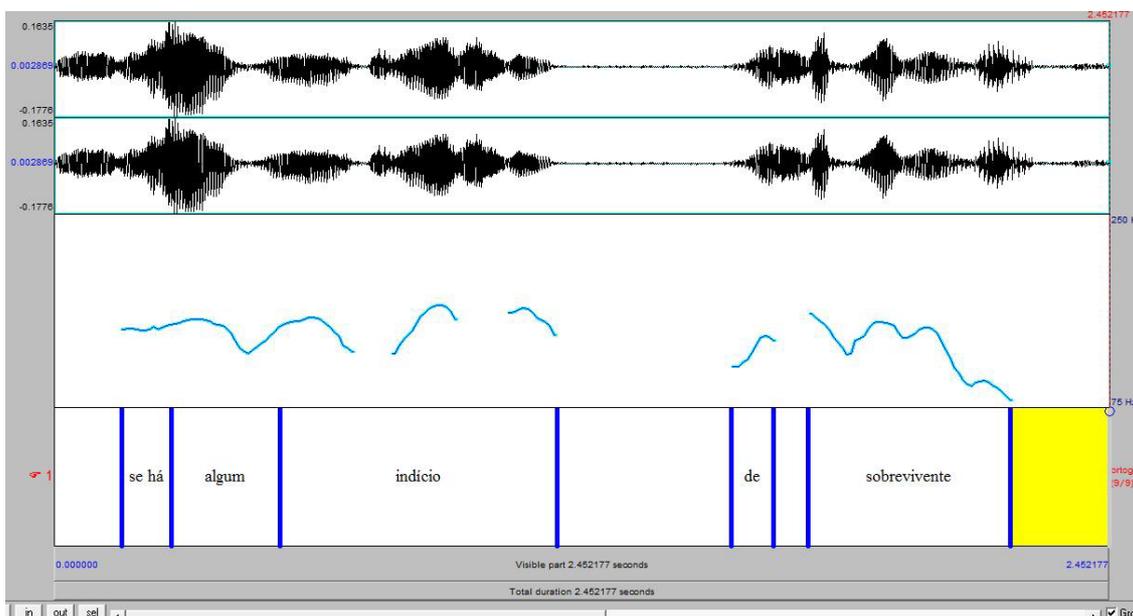


Gráfico 2 – Trecho: “se há algum indício de sobrevivente”

Como se observa, o gráfico 1 abarca um trecho maior que o gráfico 2. A razão disso é contemplar o contorno final da entrega do turno, integrado aos movimentos anteriores a ele. Dessa forma, o escopo melódico do enunciado é captado de modo mais amplo, tornando-se possível verificar se o movimento de descida acontece ao longo de

uma quantidade maior de material. Evitam-se, assim, generalizações equivocadas a partir de unidades de análise muito pequenas que não contemplam um todo mais abrangente. Por outro lado, a justificativa para a apresentação de um gráfico menor, nesse caso, é definir, conforme o sistema de notação selecionado, o contorno exato do lugar relevante de transição. O gráfico menor permite um exame mais local de aspectos que podem ser subfocalizados no gráfico maior.

Contrastando os gráficos mencionados acima, não se verificam informações aparentemente destoantes, já que o movimento descendente que sinaliza possibilidade de tomada de turno é bem perceptível no desenho dos contornos. Não há em ambos os gráficos uma subida que serviria como contra-exemplo à hipótese de trabalho de que o final do turno apresenta contorno de descida. O que se vê, analisando a F0 de “sobrevivente”, palavra que encerra o turno, é o contorno descendente HL. Tal contorno consiste no movimento de descida da frequência, estando alinhada a posição baixa à sílaba tônica, como se observa geralmente em fronteiras de sentenças declarativas. Traçando uma linha entre a altura melódica inicial em “para que” e “sobrevivente” (linha tracejada no gráfico 3 abaixo), perceberemos claramente o declínio de F0. Esse movimento, associado ao encerramento semântico do raciocínio de finalidade, faz do término do fragmento em questão o lugar exato para tomada do turno, onde ocorre a queda mais substancial da F0.

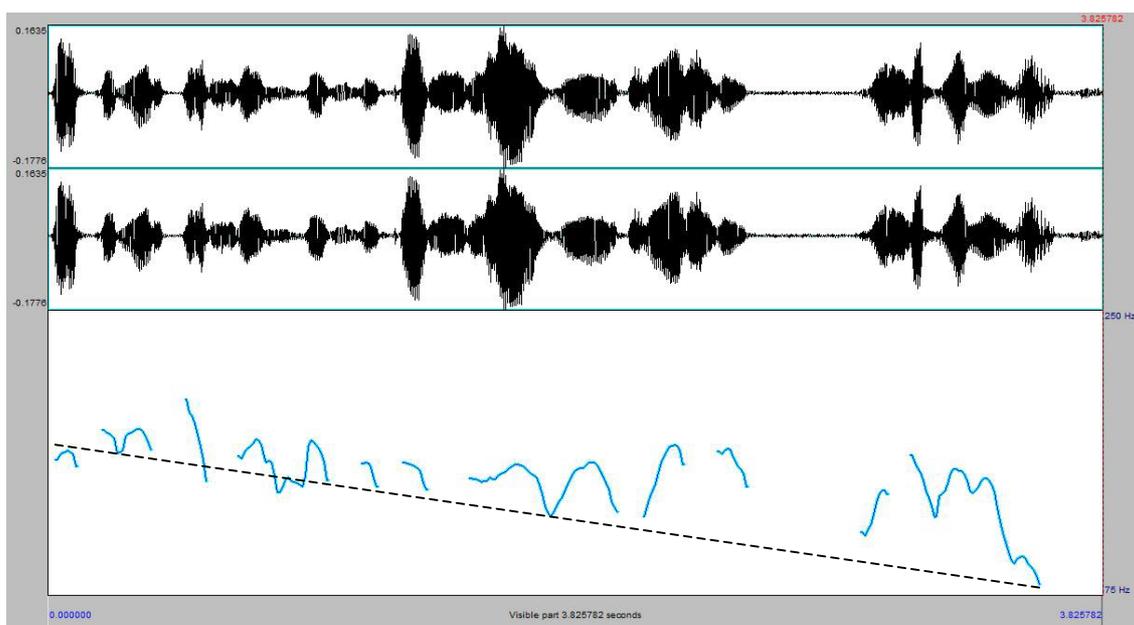


Gráfico 3 Fragmento A.1: “para que nós possamos identificar se há algum indicio de sobrevivente” (linha tracejada)

É preciso deixar claro que as informações obtidas por meio dos gráficos levam-nos a visualizar um jogo de subidas e descidas que, a princípio, poderiam nos guiar para o equívoco de considerar qualquer descida como lugar relevante de transição. Entretanto, o fechamento semântico de unidades construcionais contribui decisivamente para o interlocutor se sentir a vontade em reivindicar a palavra. É o que acontece após a emissão do fragmento A.1, quando o brigadeiro entrega o turno aos jornalistas. Esse é um caso em que a passagem de turno se dá de forma clara sem interrupções ou sobreposições, o que não ocorre no fragmento a seguir.

FRAGMENTO A.2:

“eu acho que é é é da sua competência declarar o luto [“mas”

(sobreposição dos jornalistas)] na minha na minha opinião”

Esta fala do brigadeiro diz respeito ao fato de o presidente Lula ter decretado luto oficial em consequência do acidente aéreo. Ao opinar sobre isso, o entrevistado acredita ser a fatalidade do acidente o motivo que deve ter levado o presidente a decretar luto oficial e acrescenta que a decisão de tomar essa medida compete exclusivamente ao próprio presidente.

Um detalhe interessante é que, até o momento em que o brigadeiro encerra o enunciado com a palavra “luto”, a ação conversacional parece estar completa, o que pode explicar a tentativa de tomada de turno por parte dos jornalistas e a consequente sobreposição de vozes inserida nesse limite. Concomitante à investida dos repórteres em lançar nova questão, o brigadeiro retoma seu turno (“na minha na minha opinião”) e modaliza sua fala, dizendo que tais alegações tratavam-se de uma opinião dele, esclarecendo que aquilo não era uma imposição ou verdade inquestionável, e sim seu pensamento sobre o assunto.

O gráfico 4 abaixo registra, em unidades mais focais, o que ocorre prosodicamente enquanto se dá esse jogo interacional. Observe o pareamento entre a queda da F0 e a palavra “luto”, situação em que o turno é quase roubado pelos jornalistas. A presença da conjunção adversativa “mas”, que se encaixa nesse instante preciso, anuncia uma suposta contra-argumentação, não efetivada pelo jornalista devido à retomada do turno por parte do entrevistado. Explicações mais detalhadas do gráfico a seguir podem ser observadas nas caixas de textos, visto que é complexo apontar textualmente para cada detalhe do gráfico, sem a utilização desse recurso.

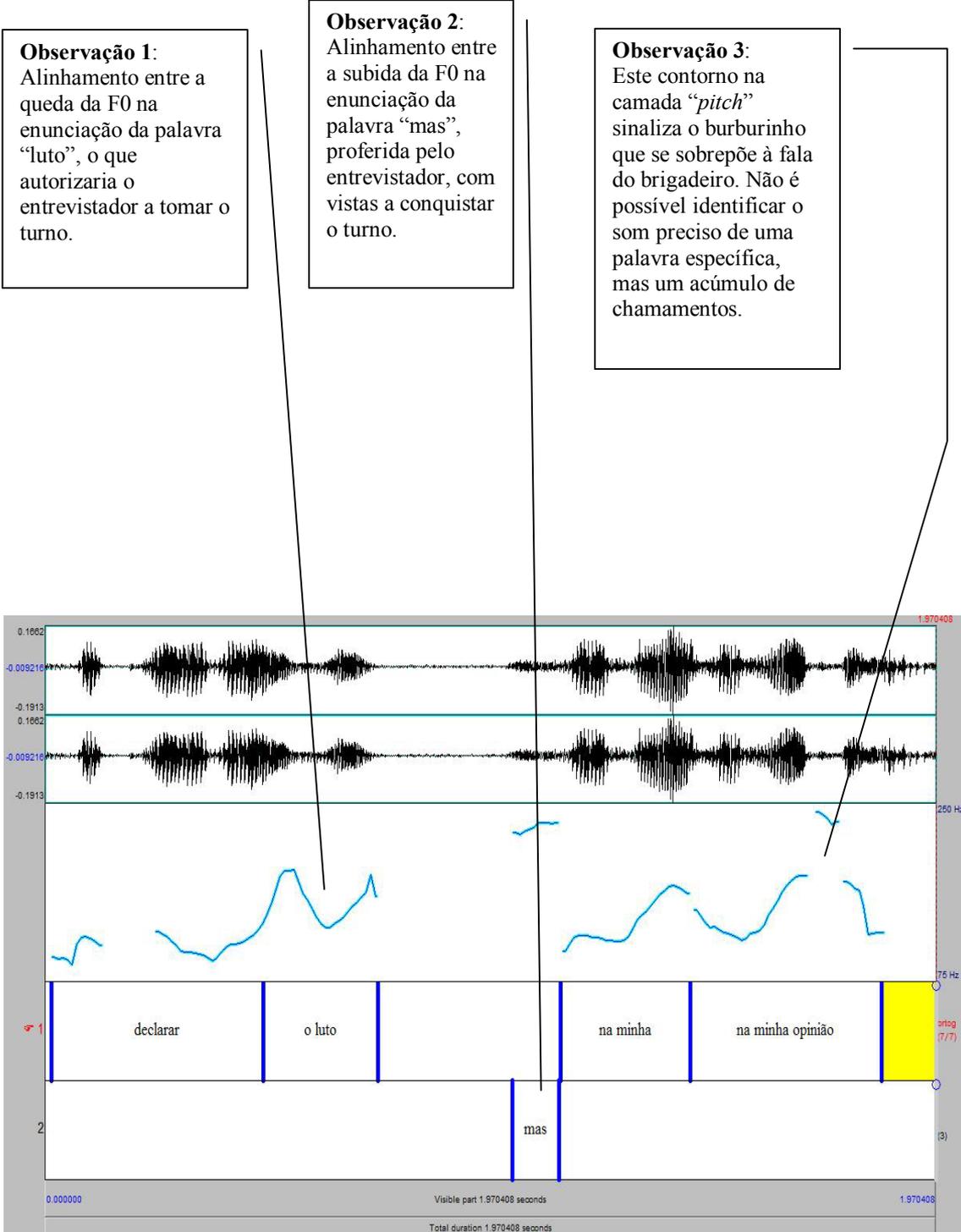


Gráfico 4 Trecho: “declarar o luto [mas] na minha na minha opinião”

No próximo gráfico, onde é contemplado o mesmo trecho, mas inserido em uma unidade maior, podemos ver como a competição da fala pode estar refletida na configuração dos contornos melódicos. São muitos os movimentos de subidas e

descidas que revelam estratégias por parte do falante em assegurar seus turnos de fala. Em muitos casos, ele precisa marcar que ainda não concluiu o que estava falando. É o que podemos ver abaixo:

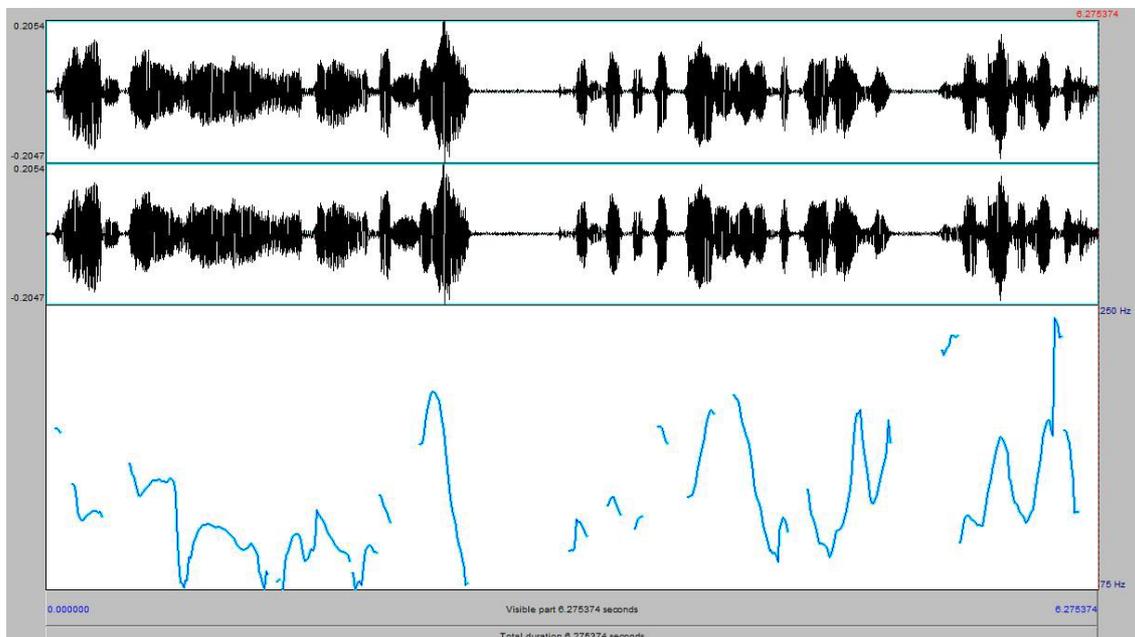


Gráfico 5 Fragmento A.2: eu acho que é é é da sua competência declarar o luto ["mas" (sobreposição dos jornalistas)] na minha na minha opinião

Nossa leitura desses dados é a de que estaria relacionada a esses movimentos uma estratégia do entrevistado em manter como seu o turno, em assegurar o piso conversacional e conseguir completar seu pensamento. Dois pontos, portanto, são importantes de se mencionar: um é a hesitação na fala do entrevistado, o que como aponta Marcuschi (1998), pode ser um convite à tomada de turno; o outro é a sobreposição de vozes, que se trata da fala durante o turno do outro. Nesse momento, a F0 sobe, marcando a competição pelo turno conversacional.

Em proporções um pouco menores, os mesmos movimentos de subidas e descidas são encontrados no próximo trecho.

FRAGMENTO A.3:

“então nós levamos três horas para localizar os destroços em função do que foi acompanhado pelos radares”

Nessa fala, o brigadeiro dá aos jornalistas mais informações acerca da tragédia. Segundo o entrevistado, a probabilidade de haver sobreviventes em um acidente da natureza do que ocorreu é muito pequena não só pelo acidente em si, como também pela área inóspita onde o avião caiu. A selva amazônica é de uma complexidade muito grande e trata-se de uma zona muito cerrada. Com esses argumentos, ele justifica a demora para se chegar ao local.

O brigadeiro, nessa parte da entrevista, oferece informações que não haviam sido solicitadas. A fim de complementar a resposta e, possivelmente, já responder alguma futura questão dos jornalistas, ele dá mais detalhes a respeito da dificuldade de se chegar ao local do acidente. Com isso, ele deseja atender os anseios das pessoas e mostrar que tem as informações sobre o caso. Assim, trata-se de uma ação conversacional completa que, certamente, contribui para o prosseguimento da entrevista, ou seja, o lançamento de uma nova questão.

O gráfico 6 ilustra o movimento prosódico do fechamento desse momento da entrevista. Após todas as explicações, o entrevistado conclui a resposta e o que vemos são contornos descendentes e ascendentes ao longo do trecho e a queda maior ocorre após a justificativa introduzida pelo “em função do que”. Novamente utilizaremos o recurso das caixas de texto para identificarmos melhor esses aspectos.

Observação 1:
Alinhamento entre a queda de F0 e a palavra “destroços”. Até esse momento, há contornos ascendentes mais marcados.

Observação 2:
Alinhamento entre a subida da F0 e a retomada da fala do brigadeiro após a pequena pausa, com a justificativa introduzida por “em função do...”. A partir daí, a queda mais substancial fica evidente.

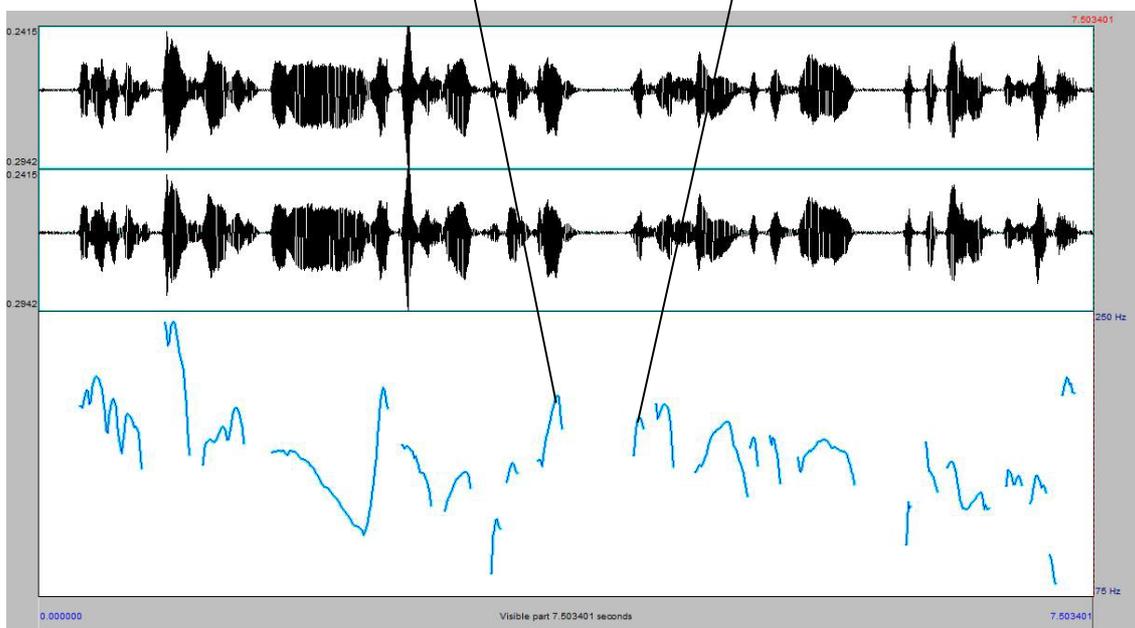


Gráfico 6 Fragmento A.3: então nós levamos três horas para localizar os destroços em função do que foi acompanhado pelos radares

No próximo gráfico, de maneira mais focalizada, visualizamos o mesmo momento:

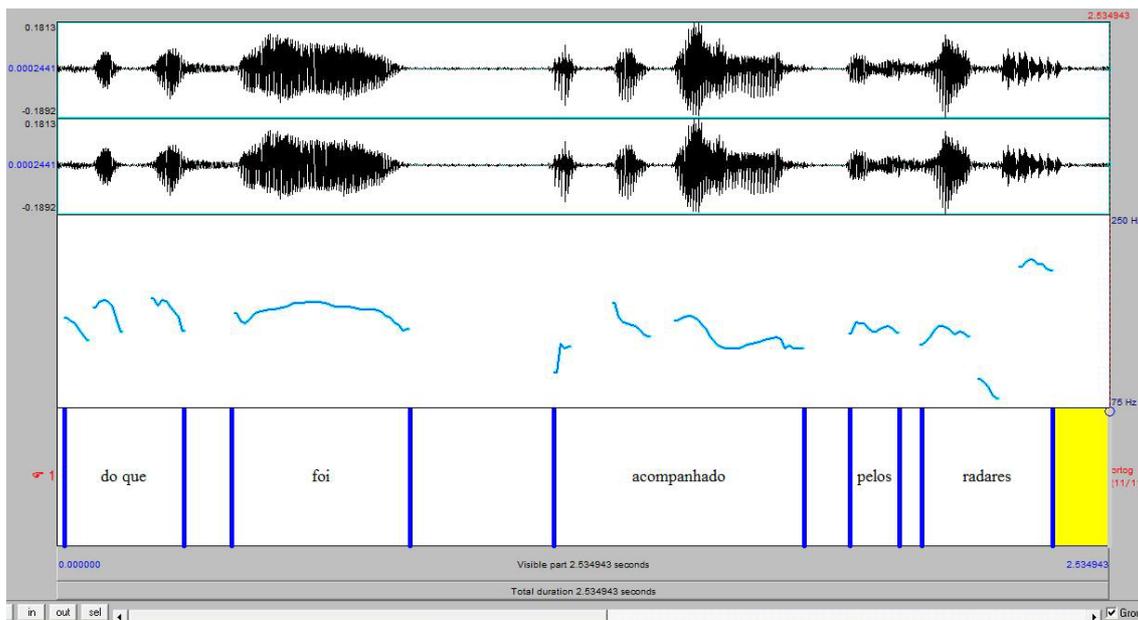


Gráfico 7 Trecho: “do que foi acompanhado pelos radares”

Novamente, fazemos questão de frisar a importância de se olhar para os dois gráficos para um exame mais cuidadoso dos contornos. Se não nos atentássemos para a queda de F0 no gráfico 6 – a partir do conector explicativo “em função do” (como indicado na caixa de texto) – e olhássemos somente para o gráfico 7, correríamos o risco de não considerar o lugar relevante de transição inserido em um movimento descendente.

Observamos, portanto, uma entrega de turno mais consentida no sentido de que não há contornos ascendentes que chamem atenção. Esse mesmo tipo de passagem de turno se dá no próximo trecho.

FRAGMENTO A.4:

“para que possamos trazer os corpos para o DF”

Nesse momento, o brigadeiro explica o procedimento para resgate dos corpos das vítimas do acidente. Segundo ele, uma infraestrutura foi montada na base aérea de Cachimbo, unidade da força aérea, para que os corpos pudessem ser recolhidos com o auxílio do Instituto Médico Legal do Mato Grosso e, com isso, o processo do traslado dos corpos ocorrer de forma mais rápida.

Os jornalistas, várias vezes, perguntam ao entrevistado se todas as informações estavam sendo transmitidas à família. Quando cobrado com relação a números precisos que indicassem o tempo para traslado dos corpos e o número de vítimas, o brigadeiro novamente fala da complexidade de se chegar ao local e a dificuldade de se localizar a aeronave. A resposta dele é semanticamente completa uma vez que ele expõe todos os procedimentos realizados para agilizar as buscas e o traslado dos corpos. Ocorre aqui, como no fragmento A.1, que para análise acústica o recorte deixou de fora a oração principal da construção; no entanto, mencionamos, para um exame semântico e interacional satisfatório, que a construção completa é: “nós poderemos minimizar o tempo para que possamos trazer os corpos para o DF”.

Para uma análise da sequência desse trecho, traçaremos novamente uma linha entre a primeira e a última marcação de contorno da F0, para checarmos se o movimento é ascendente ou descendente.

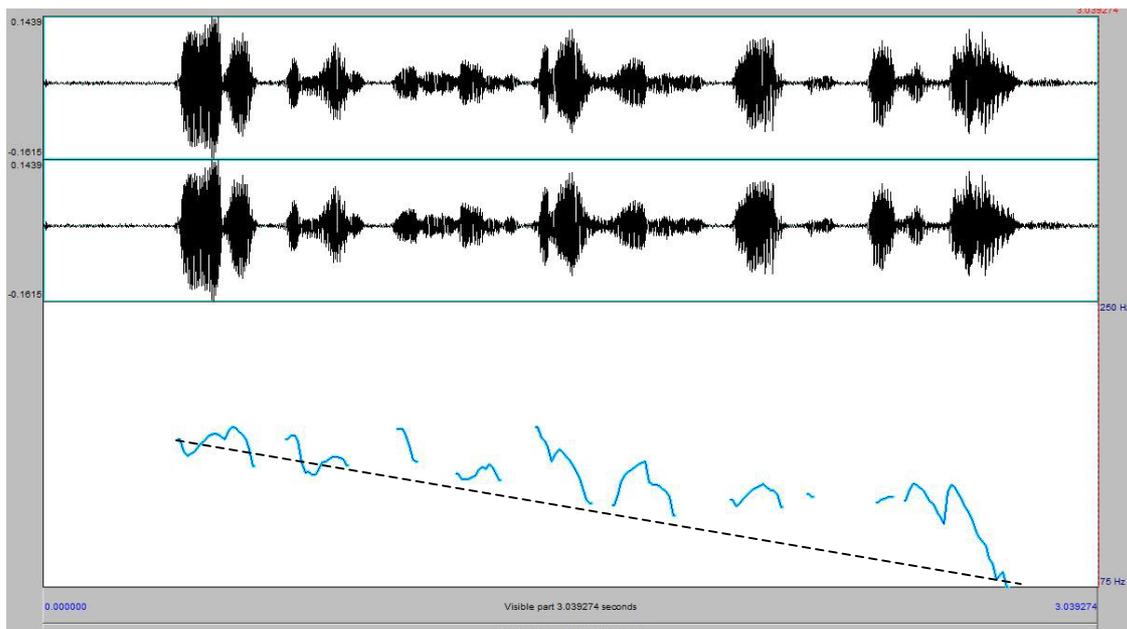


Gráfico 8 Fragmento A.4: “para que possamos trazer os corpos para o DF”(linha tracejada)

Notamos, então, que o contorno é descendente. Nossa leitura é a de que, após apresentar todos os argumentos e mostrar que os esforços por parte da ANAC estão sendo feitos, o brigadeiro entrega seu turno, e os jornalistas percebem que o momento de entrar em ação novamente passou a ser deles.

No gráfico a seguir, encontramos o contorno HL na conclusão da fala.

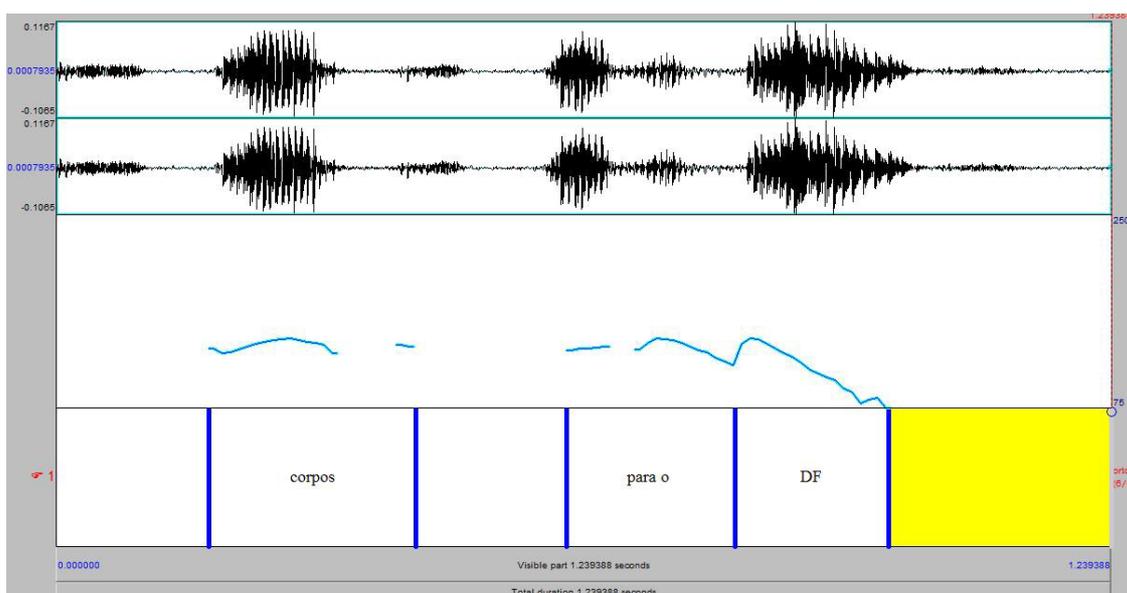


Gráfico 9 Trecho: corpos para o DF

Embora não marquemos, em todos os casos, o contorno dinâmico, conforme o sistema de transcrição DaTo, pois cumprimos a notação somente nos melhores exemplos, acreditamos que a recorrência do contorno acima nos permite esperar que haja um padrão de HL para fronteiras finais de sentenças declarativas no português. Discutiremos agora o último trecho selecionado da entrevista do brigadeiro.

FRAGMENTO A.5:

“vai ser em função do da quantidade de corpos que foram que forem encontrados no local”

A última fala selecionada do brigadeiro mais uma vez responde ao anseio dos jornalistas com relação à possibilidade de haver sobreviventes. O que o entrevistado explica é que se há um sobrevivente e ele toma um rumo diferente do tomado pelo pessoal das buscas é possível que consigam encontrá-lo somente um tempo depois. Isso dependeria do número de corpos encontrados no local.

Após essas explicações, ele encerra seu turno. Consideramos que a resposta se trata de uma ação conversacional completa. Nesse e em outros trechos, a posição do entrevistado parece ser realmente a de “brigar” para conseguir mostrar que ele, representando o cargo que assume, está tomando todas as medidas necessárias com relação ao acidente.

A F0 da fala do brigadeiro é marcada por contornos que, ora ascendentes, coincidem com a estratégia de mostrar que ainda quer falar; ora descendentes, coincidem com a entrega, a conclusão, a deixa para os jornalistas. Melhor do que descrever isso textualmente, o gráfico a seguir ilustra como esses contornos estão dispostos no fluxo da fala do entrevistado.

Observação 1:
Considerável subida de F0 quando ele diz: “vai ser em função”, marcando, então, que sua fala continuaria.

Observação 2:
Sobreposição de vozes, marcada pelo burburinho dos jornalistas em busca de mais perguntas. Observamos que essa sobreposição é, de certa maneira, licenciada pelo movimento de forte descida anterior.

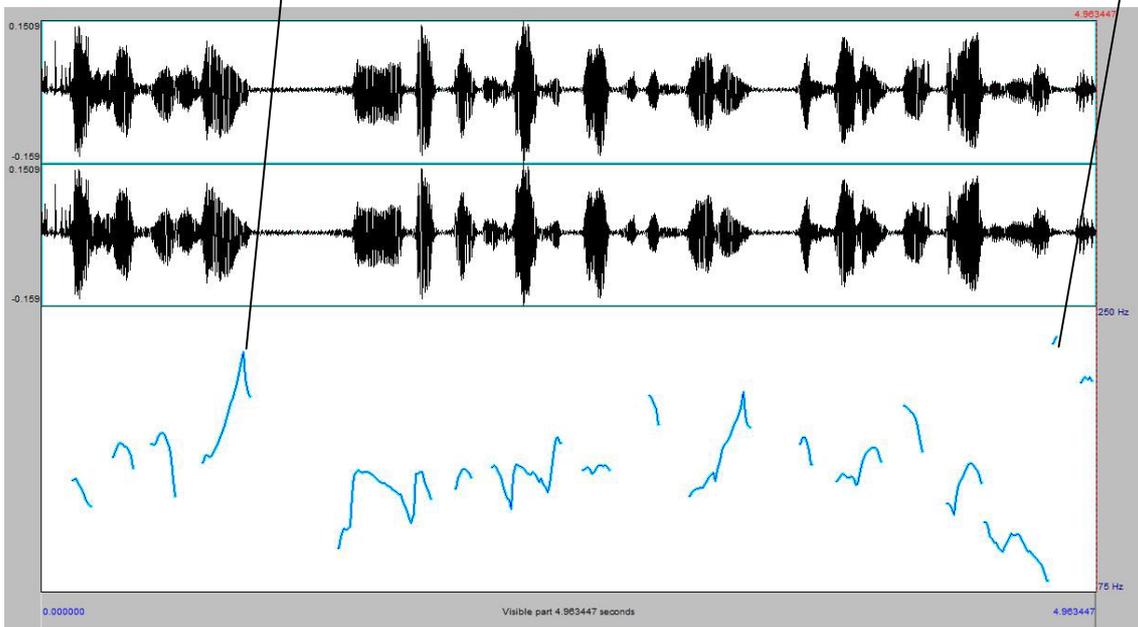


Gráfico 10 Fragmento A.5: “vai ser em função do da quantidade de corpos que foram que forem encontrados no local”

É interessante notar que a subida da F0 funciona tão bem como pista para o ouvinte de que a vez de falar prossegue que, mesmo com a pausa após o “vai ser em função”, os jornalistas não tentam roubar o turno interrompendo o brigadeiro. É evidente que a não completude semântica colabora para que isso aconteça. Contudo,

trata-se de um exemplo de como o contorno melódico funciona bem para demarcar esses limites de território na distribuição dos turnos de fala. Continuaremos com nossas caixas de texto para mais comentários sobre esse momento da fala do brigadeiro.

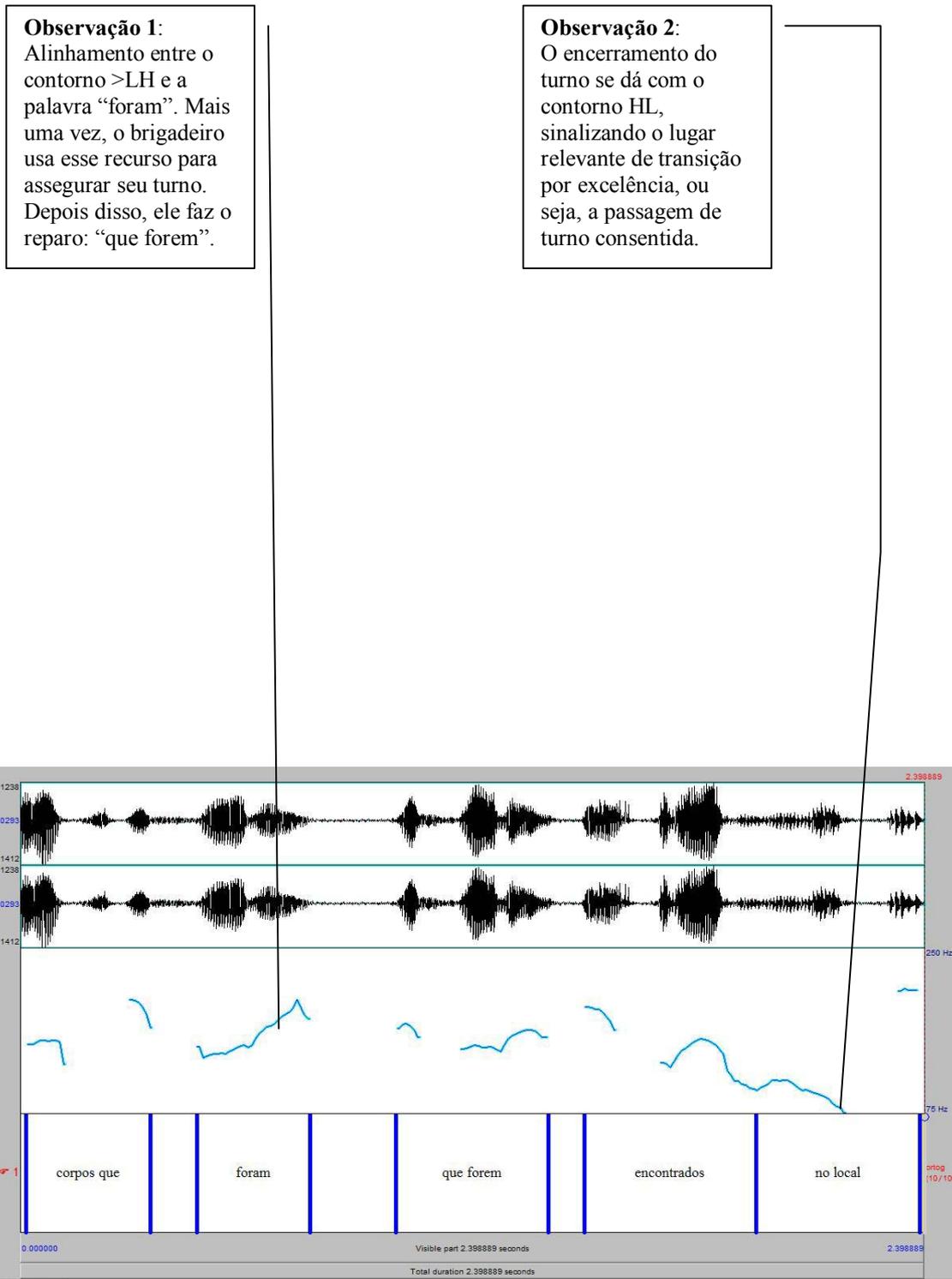


Gráfico 11 - Trecho: “vai ser em função do da quantidade de corpos que foram que forem encontrados no local”

Lembrando que o contorno >LH, mencionado acima, refere-se ao movimento ascendente, com um atraso na subida da F0, estando a vogal tônica alinhada com o movimento de subida e não com o pico.

Após respondermos as questões relacionadas com os aspectos semânticos, interacionais e prosódicos dos trechos selecionados para análise, seguiremos discutindo a contraparte cognitiva desses resultados. O que essa entrevista, de contexto tão acirrado e com tantos movimentos de subidas e descidas bastante acentuados de F0, aponta com relação ao movimento de luta e aos gestos físicos aparentes?

No que diz respeito à relação entrevistado *versus* entrevistador, na cena interacional analisada, o simbolismo dos gestos sonoros detectados está diretamente vinculado a aspectos cognitivos e corpóreos. Apostando-se veementemente na hipótese da corporificação da linguagem, são muito plausíveis algumas analogias entre os gestos produzidos pelo aparelho fonador e aqueles feitos por outras partes do corpo. Não é coincidência, quando se assume tal hipótese, que movimentos distintos das pregas vocais, entendidos como gestos sonoros, podem sinalizar diferentes unidades simbólicas. Uma vez que a metáfora é entendida como sendo presente em nossas ações e em nossos pensamentos, não é inviável deduzir que o badalar contundente das pregas vocais, sugerindo uma F0 alta, pode ser associado ao movimento dos braços na luta corporal. Quanto mais frequentes são os movimentos, maior é o poder de ataque, a resistência e a proteção dos combatentes. Quanto menos frequentes são os movimentos, maior é a chance de contra-ataque. O próprio desenho dos contornos melódicos da camada 2 sinaliza oscilações entre descidas e subidas, que parecem simular uma luta “sonora”.

A foto a seguir, extraída do site <http://eupodiatamatando.com/tag/esportes/> (acesso em 26 de janeiro de 2010), ilustra corporalmente essa alternância acústica, na qual o lutador da direita impõe sua força ao da esquerda, que ainda desfere golpes, resiste, mas está mais sujeito a se submeter ao mais forte. Analogicamente, é o que acontece com a “luta prosódica”, em que os golpes efetuados com o corpo são análogos aos “golpes” sonoros promovidos pela oscilação de F0 durante o embate pela palavra.



Foto 1 Golpes corporais análogos aos “golpes” sonoros

Fonte: <http://eupodiatamatando.com/tag/esportes/>

Isso vale para a prosódia verificada na primeira coletiva examinada, em que os ânimos estão muito acirrados dada a gravidade da situação. A descida da curva entoacional e o fechamento do raciocínio simbolizam a perda da palavra e o consequente e provisório recuo por parte do falante no embate discursivo. É o momento em que ele permite o golpe do adversário em forma de pergunta. A queda de F0, observada ao longo do todo do enunciado, é a base concreta para projeções de caráter metafórico do tipo: MAIS FREQUENTE É MAIS FORTE.

O ditado “água mole em pedra dura tanto bate até que fura” é muito significativo para emblematizar o poder da frequência e da persistência de ações, sejam elas de efeitos sonoros ou não, no resultado das interações do ser humano com o ambiente. Do ponto de vista linguístico, a frequência de construções é relevante para o armazenamento e fixação de padrões, como ensinam os modelos teóricos baseados no uso. Embora estejamos falando polissemicamente da palavra “frequência”, a base de sentido que agrega o entendimento popular e a abordagem científica é o corpo, mais precisamente a frequência produzida pelos movimentos das pregas vocais. Entrevistado *versus* entrevistador agrega o conjunto de instâncias de entrevistas coletivas nas quais a

luta pela palavra se dá de modo tão contundente que os participantes devem acionar seus conhecimentos mais finos e sofisticados para se manterem fortes na cena.

O limite do jogo de forças é dado também pela capacidade corpórea de respiração no momento da produção do enunciado. Nota-se claramente que, ao longo da fala, o ato de expirar é um importante componente aerodinâmico da F0, exercendo uma pressão subglotal que interfere no contorno melódico final. É o que dizem muitos autores em prosódia, mas por outro lado trata-se de uma outra evidência para o simbolismo vinculado a F0. Na maioria das vezes inconscientes, tanto o *pitch* quanto a percepção dos movimentos respiratórios dos interlocutores atuam na licença para tomada de turno, uma licença que está estruturalmente prevista no domínio da coletiva de imprensa. Já que se render, ou seja, passar a palavra, nessa instância é comum, pode ser que o entrevistado promova movimentos de F0 com o intuito de manter a palavra. Durante a entrevista, o brigadeiro mantém a F0 alta para evitar perguntas, que representariam ameaças às justificativas que ele expõe o tempo todo.

Prosseguiremos a análise com o segundo contexto de entrevistas estudado.

B) Entrevistado e entrevistador

O título dessa seção se difere claramente do título da seção anterior por conta do conectivo entre as palavras. Na primeira situação (entrevistado *versus* entrevistador), tínhamos uma situação concreta de confronto e assim foi categorizado o primeiro contexto das entrevistas.

Agora, o conectivo “e” ameniza a idéia de combate e sugere, senão uma união, pelo menos a imagem de duas forças seguindo em uma mesma direção. É o que ocorre nas duas próximas entrevistas. Diferente do primeiro caso, nenhum assunto polêmico gira em torno das coletivas de imprensa. Não há um interesse por parte dos jornalistas em pressionar o entrevistado na busca de informações e nem de cobrar ações dele. Ao contrário, a cena envolvida conta com entrevistado e entrevistadores que cumprem seus papéis e seguem uma espécie de protocolo. Os jogadores entrevistados respondem às perguntas e, conseqüentemente, promovem-se colocando em evidência suas experiências e suas novas posições no novo time. Os jornalistas conseguem suas informações para atender à curiosidade de seus leitores sobre a nova contratação do clube de futebol.

Nesse momento, a troca de turnos e a sucessão de perguntas e respostas não representam tão tipicamente adversários em guerra como em A (entrevistado *versus* entrevistador). De qualquer forma, analisaremos como essa nova cena está organizada, respondendo às mesmas questões, colocadas na primeira parte da análise. Como a prosódia atuará nesse caso, e como, em construção com as unidades prosódicas, operam as informações semânticas e interacionais?

As entrevistas que ilustram esse contexto são as coletivas com os jogadores Daniel e Álvaro. Justificamos que, devido à qualidade baixa dessas gravações de modo em geral, aproveitou-se para análise acústica apenas os trechos mais confiáveis. O eco e as interferências prejudicam a submissão de uma parte maior dessas entrevistas ao PRAAT.

As imagens copiadas para ilustrar as situações descritas são as seguintes:



Imagem 3 O jogador Daniel sorrindo quando brinca com os jornalistas sobre treinar com o pé direito já que ele e os demais jogadores do quadrado mágico são canhotos.

Fonte: <http://www.youtube.com/watch?v=wKD2SfEzISs>

Nessa imagem, Daniel se mostra descontraído e receptivo às perguntas dos jornalistas. Ele até brinca em uma das respostas, em que um dos entrevistadores chama de quadrado mágico o grupo formado por Daniel e mais três jogadores. Em nenhum momento, destaca-se algum tipo de exaltação ou investida mais abrupta dos jornalistas. Parece haver, enfim, uma luta “amigável” entre as partes e isso fica refletido na configuração melódica que não apresenta contornos de “guerra”.



Imagem 4 O jogador Álvaro encerrando sua resposta à primeira pergunta dos jornalistas.

Fonte: <http://www.youtube.com/watch?v=KJvY63KpmQQ>

Através dessa imagem, percebemos que, com Álvaro, a entrevista também ocorre de maneira leve, com perguntas e respostas pacíficas. O semblante do jogador é tranquilo e ele chega a sorrir também em alguns momentos da entrevista. O assunto envolve boas expectativas do jogador em voltar a atuar no Brasil. Assim como na entrevista de Daniel, não há uma disputa acirrada entre os participantes da cena. Basta ouvirmos as entrevistas para perceber que, com os jogadores, o “tom” da entrevista é de uma conversa amigável, um bate-papo descontraído, bastante distinto da situação com o brigadeiro.

A partir de agora, apresentaremos alguns trechos que puderam ser selecionados das entrevistas com os jogadores.

FRAGMENTO B.1:

“aí vai ser outra história outra novela e a gente vai esperar pra deixar pra dezembro tu resolver”

Daniel, o jogador-autor desse trecho, conta aos jornalistas sobre sua experiência de ter jogado no futebol russo e sobre suas expectativas em atuar no clube Internacional de Porto Alegre. Quando questionado sobre a possibilidade de permanecer no time gaúcho por mais de seis meses, o tempo combinado, ele responde que existe a possibilidade, pois ele pode querer ficar mais. No entanto, para o jogador, seria preciso aguardar até dezembro para tomar a decisão, pois só nesse mês já haveria os resultados dos campeonatos.

Quanto a tratar-se de uma ação conversacional completa ou não, podemos dizer que Daniel desenvolve tranquilamente seus argumentos sustentando sua opinião da necessidade de esperar o fim do ano para tomar a decisão de permanecer ou não no Internacional. Ele completa seu raciocínio e em nenhum momento a completude da sua fala parece ameaçada, um dado que evidencia o não-acirramento da interação.

A configuração prosódica do trecho em destaque mostra um desenho mais linear (quando comparado com os dados da primeira entrevista). Observamos somente uma única subida mais considerável de F0, como podemos conferir no gráfico abaixo:

Observação 1:
Alinhamento da única subida de F0 mais considerável com “outra história”.

Observação 2:
O encerramento do turno, ou seja, o lugar relevante de transição é marcado por essa pausa e em seguida a descida da F0.

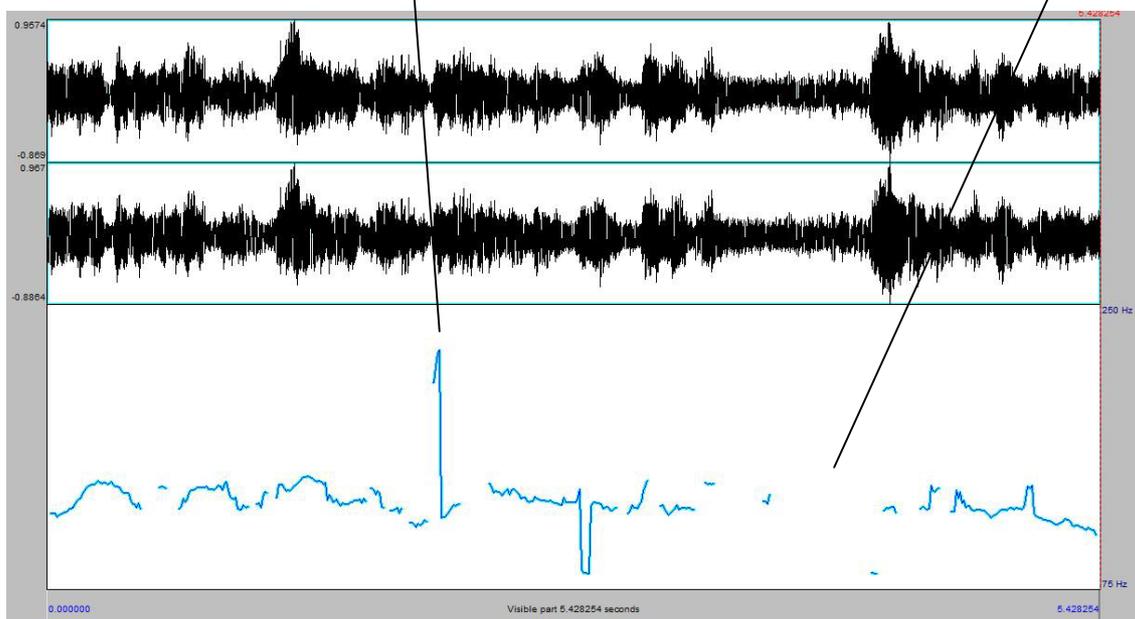


Gráfico 12 Fragmento B.1: “aí vai ser outra história outra novela e a gente vai esperar pra deixar pra dezembro tu resolver”

O mesmo trecho, mas recortado em uma unidade menor, segue abaixo:

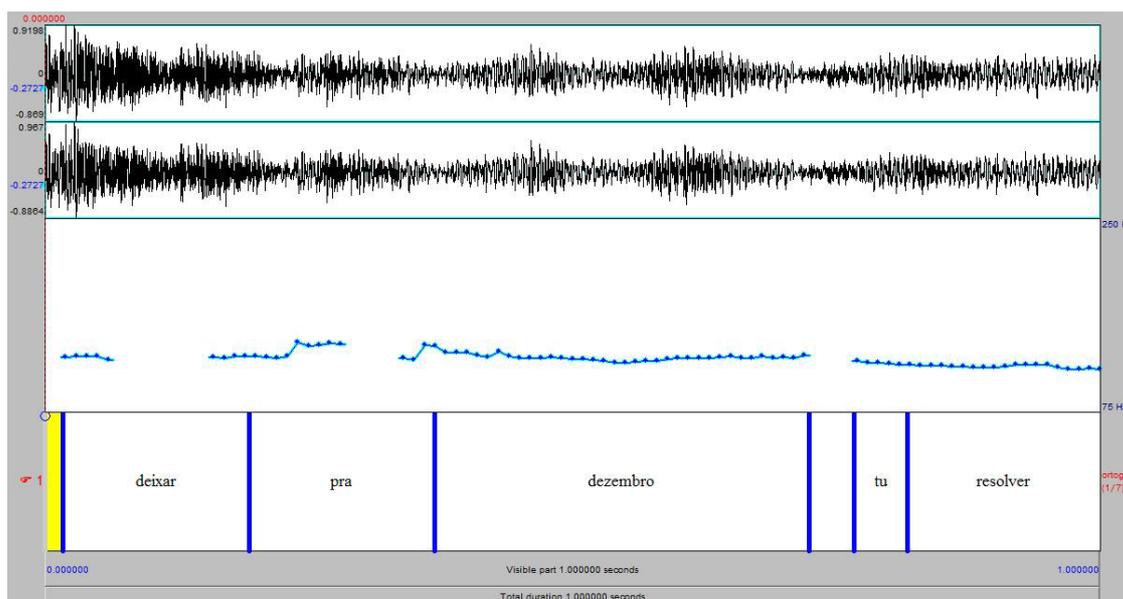


Gráfico 13 Trecho: “deixar pra dezembro tu resolver”

Como podemos ver, não há movimentos de subidas e descidas da forma como acontece na primeira entrevista. O gráfico 12 é imprescindível para a análise desse momento de transição, pois somente através dele percebemos que há, de fato, um contorno de descida no encerramento do turno. Essa percepção só é possível em virtude do material anterior que, em sequência, revela um contorno descendente. Como já explicado, a entrevista completa do jogador Daniel Carvalho, quanto ao aspecto da qualidade do som, apresenta problemas para uma análise acústica satisfatória. Por esse motivo, aproveitou-se o melhor trecho para expor no trabalho. Assim sendo, seguiremos com os trechos da fala da entrevista do outro jogador, Álvaro.

FRAGMENTO B.2:

“agora eu vou ter a oportunidade de de acompanhar de jogar e de viver o campeonato brasileiro que que esse era o meu desejo”

O contexto da entrevista do jogador Álvaro é bastante parecido com o da entrevista de Daniel. Na primeira resposta selecionada, ele comenta com os jornalistas que, quando estava na Europa, não pôde acompanhar bem o campeonato brasileiro, mas que, estando agora no Brasil, pretende participar de todas as etapas.

Semanticamente, a ação conversacional está completa e o jogador não enfrenta obstáculo nenhum que possa fazer com que ele não complete o sentido da sua resposta. Isso, certamente, proporciona uma passagem de turno consentida e sem sobreposição de vozes. Isso reforça a idéia de encontro sem acirramento, com pacífica negociação de sentidos.

Quanto à análise no PRAAT, o desenho resultante é o que segue abaixo:

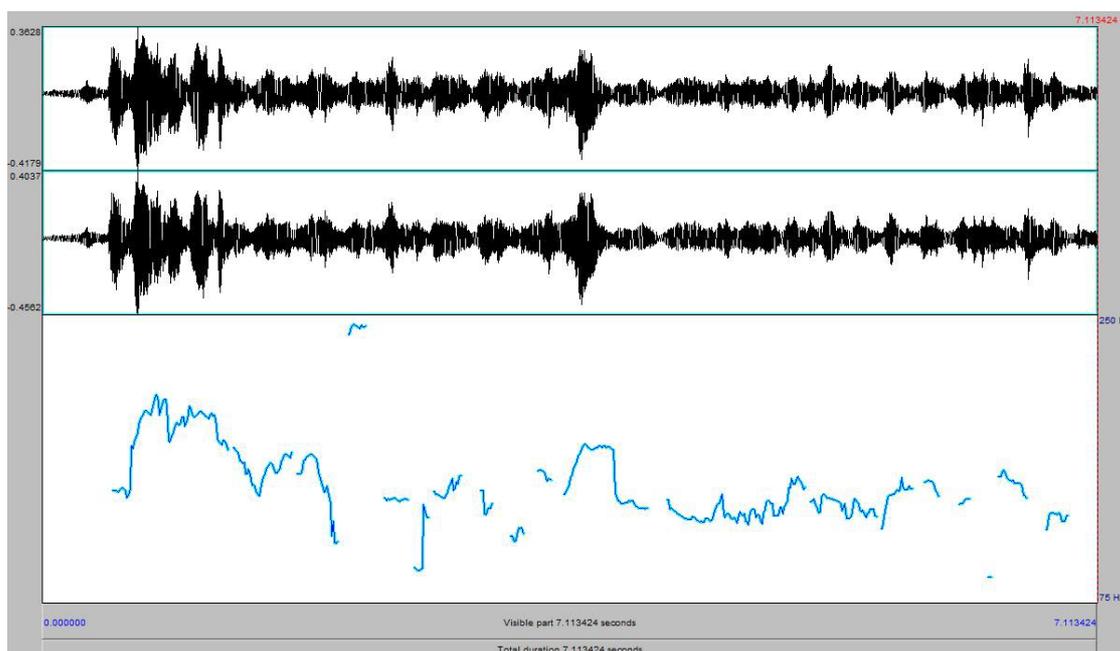


Gráfico 14 Fragmento B.1: “agora eu vou ter a oportunidade de de acompanhar de jogar e de viver o campeonato brasileiro que que esse era o meu desejo”

Já a unidade de análise recortada resulta no seguinte gráfico:

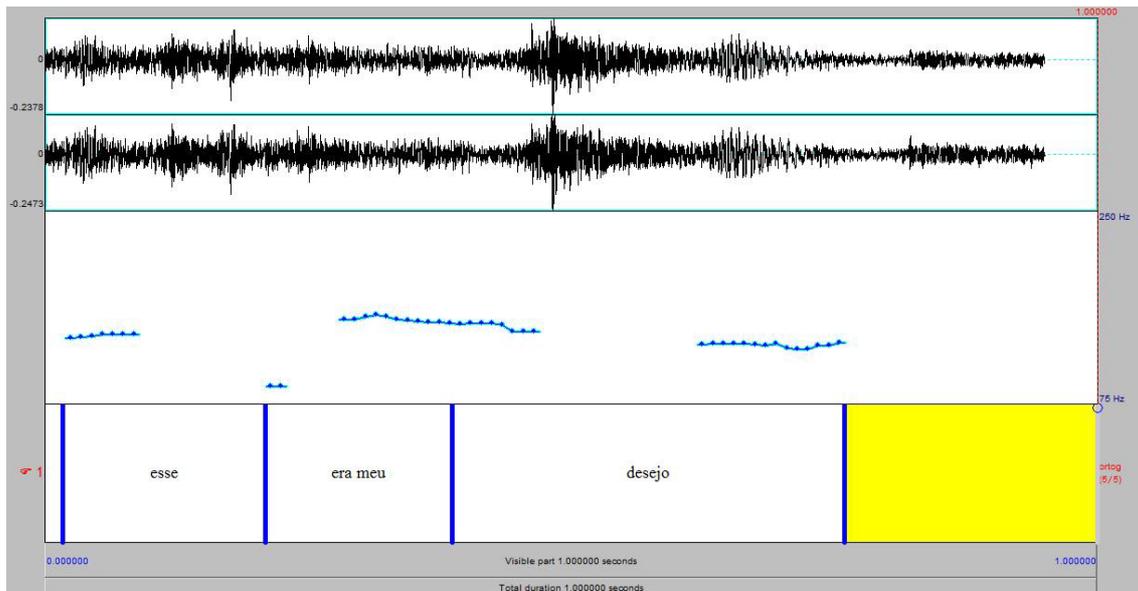


Gráfico 15 Trecho: “esse era o meu desejo”

Podemos identificar que a análise acústica gerou um movimento prosódico mais linear, com subidas e descidas muito leves. Esse seria um caso de uma passagem de turno “suave”. Quando comparamos os contornos de F0 do material antecedente ao encerramento, podemos constatar que houve uma queda que, possivelmente, confirma nossa hipótese de que o lugar relevante de transição é marcado por um contorno melódico descendente.

Seguiremos mostrando mais um exemplo da fala do jogador Álvaro, mas destacamos que, em nenhuma dessas três análises, utilizamos o sistema de transcrição adotado, justamente pelo motivo da baixa qualidade do som. Se insistíssemos em transcrevermos os contornos, poderíamos gerar resultados não muito confiáveis.

FRAGMENTO B.3:

“me encontro em em em condições de fazer um trabalho forte essas semanas e e e pra entrar em condições de jogo o mais rápido possível”

Neste trecho selecionado, Álvaro fala sobre seu estado físico. Para o jogador, suas condições são boas, em virtude do treinamento que fez para se preparar para a temporada. Assim sendo, a resposta é completa e não gera nenhuma polêmica ou mal-entendido que faça com que os jornalistas interrompam ou intervenham na resposta de Álvaro.

Apresentamos o gráfico a seguir:

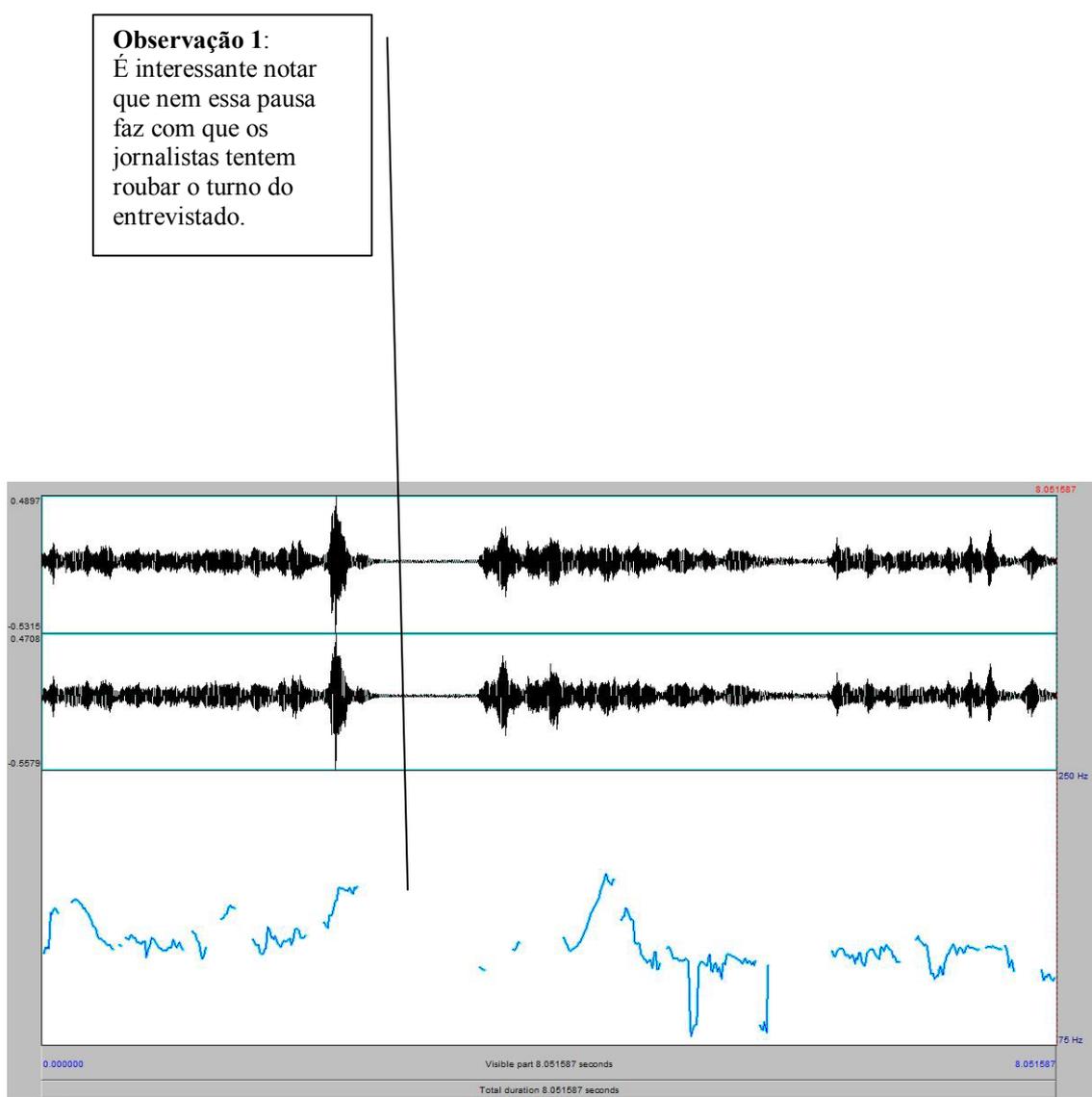


Gráfico 16 Fragmento B.3: “me encontro em em em condições de fazer um trabalho forte essas semanas e e e pra entrar em condições de jogo o mais rápido possível”

Assim como nos outros exemplos, a não-interrupção, mesmo com a pausa, decorre da pista prosódica anterior que oferece um movimento ascendente e da pista semântico-pragmática que sugere que a ação conversacional ainda não está completa. O mesmo trecho, analisado de maneira mais focal, é o que segue abaixo:

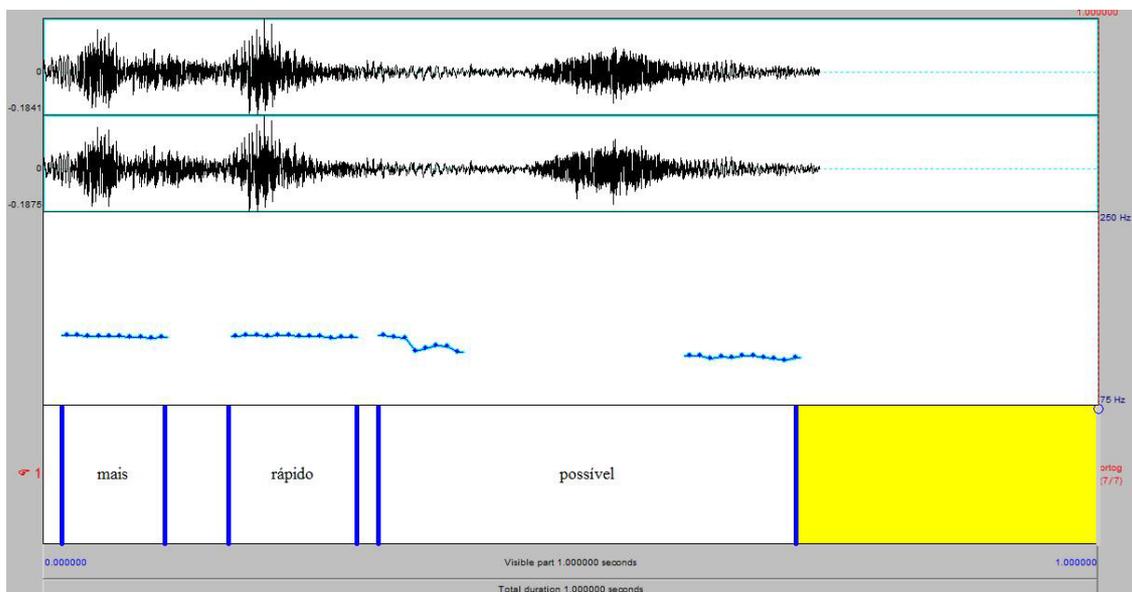


Gráfico 17 Trecho: “mais rápido possível”

Novamente, o que encontramos é uma configuração prosódica com movimentos de subidas e descidas muito leves, apresentando-se de maneira mais linear e alinhada. As implicações simbólicas revelam que maior linearidade de F0 sinaliza, pelo menos no contexto considerado, grau de acirramento interacional baixo, ou seja, não se verificam sobreposições, tentativas de tomada e manutenção de turno.

Assim como na outra seção, para encerrarmos, comentaremos um pouco do que a análise e a exposição dos dados podem nos revelar sobre a contraparte cognitiva desse movimento prosódico que atua na construção do sentido na interação, e mais especificamente, na troca de turnos, da coletiva de imprensa.

Procede dizer, também para esses casos, que o simbolismo dos gestos sonoros estão vinculados a aspectos cognitivos e corpóreos. Persistimos na hipótese de que podemos tecer analogias entre gestos produzidos pelo aparelho fonador e os realizados por outras partes do corpo. Nesse segundo contexto de entrevista, a metáfora DISCUSSÃO É GUERRA se faz presente, mas dessa vez, o movimento mais alinhado da F0 nega a existência de movimentos fortes de ataques com os braços na luta corporal.

Assim, como não há agressão nem resistência entre os combatentes, não há também defesa, e, com isso, a frequência dos movimentos cai (movimentos das pregas vocais). A luta sonora em “entrevistado *e* entrevistador” não atinge as oscilações de “entrevistado *versus* entrevistador”. A configuração da F0 se mantém alinhada, na maioria das vezes, simbolizando a união e não o embate entre entrevistado e entrevistador. Por isso, não predominam os golpes de ataque e defesa. A metáfora MAIS FREQUENTE É MAIS FORTE também pode ser aplicada para esses casos, visto que a F0 baixa vai ao encontro de movimentos menos frequentes, pois não há a necessidade de imposição contínua e repetitiva de forças.

Passaremos agora à análise do terceiro contexto de entrevista e observaremos em que ele se difere ou se aproxima dos demais.

C) Entrevistado “*entrevista*” entrevistador

O título desta seção pode causar, em uma primeira leitura, certo estranhamento ao leitor da dissertação. Isso porque ele anuncia uma inversão de papéis, até então, inesperada. Como pode um entrevistado entrevistar o entrevistador, se a função daquele é a de responder o que este pergunta?

Como mostraremos, é diferente o modo como a cantora Ivete Sangalo se deixa ser entrevistada, de tal forma que ela pode até se passar, em alguns momentos, por entrevistadora. Analisaremos como esse comportamento interfere no jogo interacional. Como essa cena se mostra, sobretudo, prosodicamente? Serão usadas, nesse contexto, as mesmas estratégias de “entrevistado *versus* entrevistador” ou “entrevistado *e* entrevistador”? As habilidades cognitivas que envolvem atenção dos jornalistas serão empregadas da mesma maneira?

A entrevistada, muitas vezes, detém a posse do turno conversacional, tornando difícil a missão dos repórteres de lançar novas questões. No entanto, selecionamos quatro trechos em que ela efetivamente encerra sua fala e a transição do turno para os jornalistas se concretiza.

A seguir, uma imagem da entrevista, a título de ilustração:



Imagem 5 Ivete respondendo a pergunta dos jornalistas sobre a crise na economia brasileira.

Fonte: <http://www.youtube.com/watch?v=f3t8y9yy00Q>

Embora não faça parte de um momento selecionado para análise acústica, essa imagem de Ivete reflete bem a maneira como ela se portou durante a entrevista, pois são muitos os movimentos e gestos que a cantora produz. Ela comanda, de forma bem humorada, a maior parte da interação, assaltando o turno dos jornalistas, ou seja, ela, muitas vezes, interrompe os jornalistas, ou demora a deixar que eles perguntem, pois retoma sua fala várias vezes. Ivete interage tanto com os jornalistas que, muitas vezes, devolve perguntas a eles. Ocorre também nessa entrevista que algumas tentativas de perguntas dos jornalistas acabam frustradas, tendo eles que repetirem a pergunta visto que as primeiras investidas passam despercebidas por Ivete.

A exposição de imagens das entrevistas é realizada a fim de acompanharmos como o que acontece “musicalmente” (contornos melódicos) e “internamente” (gestos do aparelho fonador) é demonstrado “visualmente” (imagens dos entrevistados na cena).

Vamos, então, ao primeiro exemplo para entendermos melhor o que está sendo dito.

FRAGMENTO C.1:

“o respeito mútuo que temos uns com os outros a prova disso é que como ela falou com a Bahia ninguém pode”

Ivete é perguntada quanto à rivalidade entre artistas baianos e quanto à competição das músicas no carnaval. A cantora responde que todos os artistas se comportam de maneira a respeitar um ao outro. Ela entende a dita rivalidade como foco de notícias, pois a mídia se sustenta desse tipo de assunto. Completa, ainda, dizendo que cantaria músicas de todos no carnaval, assim como todos vão tocar suas músicas, o que é importantíssimo para o sucesso dos artistas.

Como percebemos, ao longo da entrevista, Ivete vai além das respostas esperadas, estendendo, às vezes, o assunto. O domínio irreverente da interação faz com que a cantora responda às perguntas, mas também promova o desvio dos tópicos abordados pelos jornalistas. Em alguns momentos, as respostas são incompletas, contudo isso não é motivo de cobrança por parte dos entrevistadores. Em outras ocasiões, ela interrompe o jornalista no meio da pergunta já sinalizando entender a questão sem que ele a conclua. É nesse sentido que dizemos que Ivete “entrevista” o entrevistador.

Somente para observarmos como se dá prosodicamente essa outra possibilidade de interação, ou seja, a situação em que entrevistada assalta o turno do entrevistador, atentemo-nos para o gráfico abaixo, no qual Ivete interrompe o jornalista:

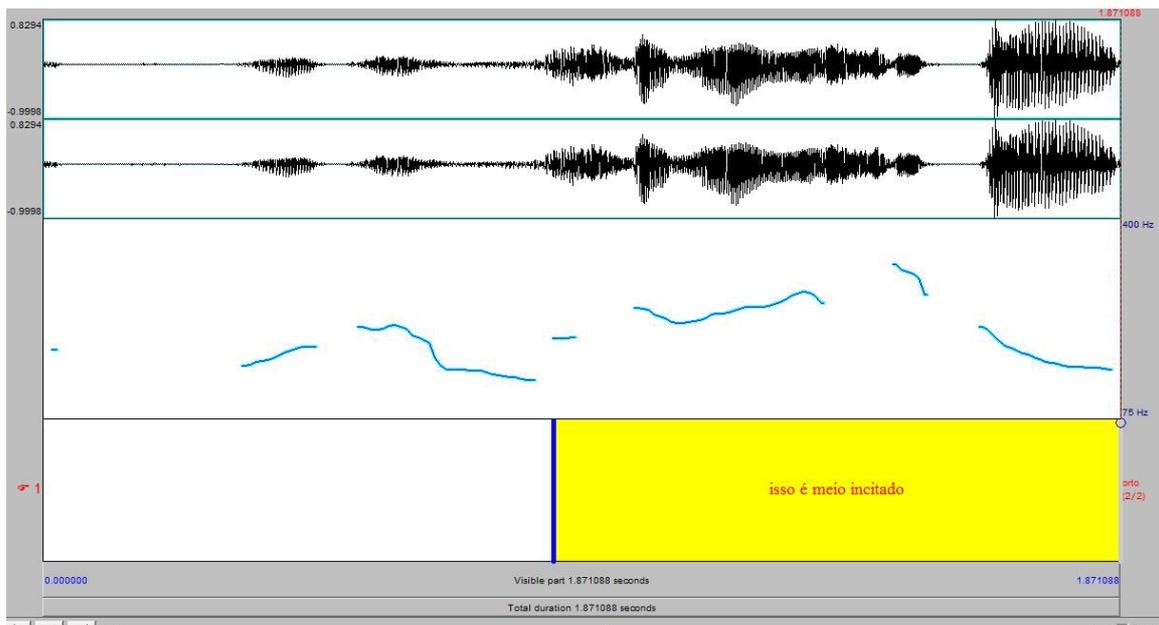


Gráfico 18 Trecho: “isso é meio incitado”

O momento no qual se insere o trecho acima é o seguinte: ao elaborar uma pergunta, o jornalista está afirmando que existe uma rivalidade entre os artistas: “(...) a imprensa, como os fãs também, eles criam uma rivalidade entre os artistas e entre e entre os músicos (...)”. Nesse momento, antes mesmo de completar sua questão, Ivete diz: “isso é meio incitado”. Como podemos ver no gráfico, antes do espaço marcado abaixo de amarelo, o movimento é de descida e a fala de Ivete apresenta uma F0 mais alta. Assim, percebemos que menos frequência possibilitou a tomada de turno por parte da entrevistada.

Voltando ao Fragmento C.1, a resposta da cantora se trata de uma ação conversacional completa. Quanto ao desenho prosódico, as curvas de F0 geradas pelo programa PRAAT mostram movimentos sucessivos de subidas e descidas, sendo que, ao final, o contorno é descendente.

Observação 1:

Nesse momento, ela parece estar encerrando sua fala: “o respeito mútuo que temos uns com os outros”.

Observação 2:

Em seguida, a F0 sobe para ela não perder o turno e completar sua resposta: “a prova disso é como ela falou...”.

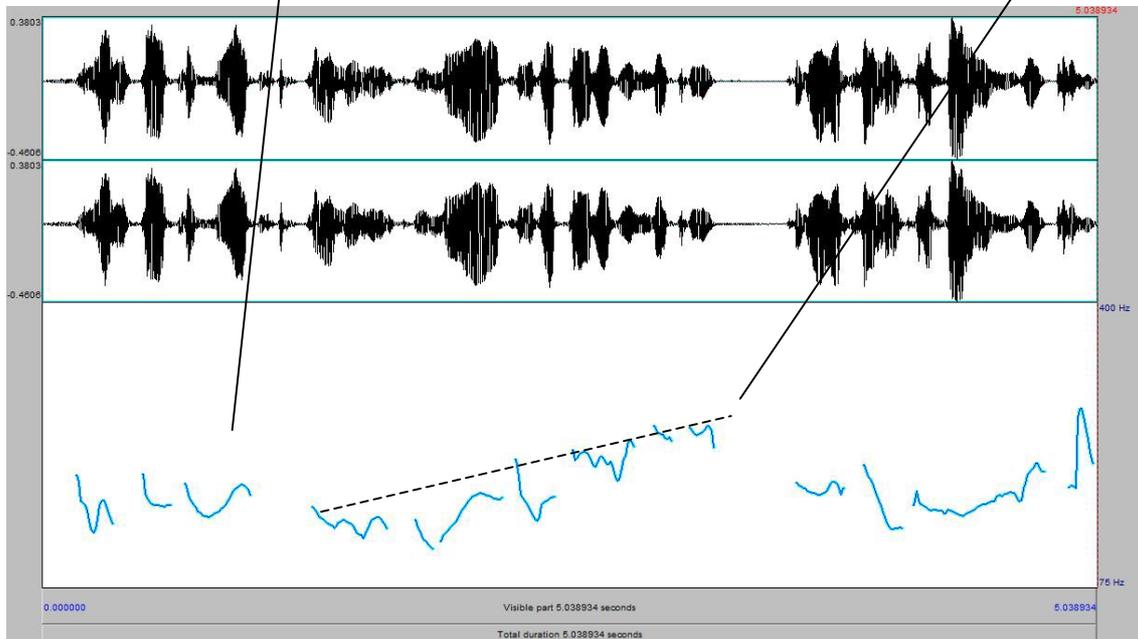


Gráfico 19 Fragmento C.1: “o respeito mútuo que temos uns com os outros a prova disso é que como ela falou com a Bahia ninguém pode”

O mesmo trecho segue abaixo em dimensões mais focais:

Observação 1:
Quando se dá,
efetivamente, a
transição do turno, o
contorno melódico é o
HL.

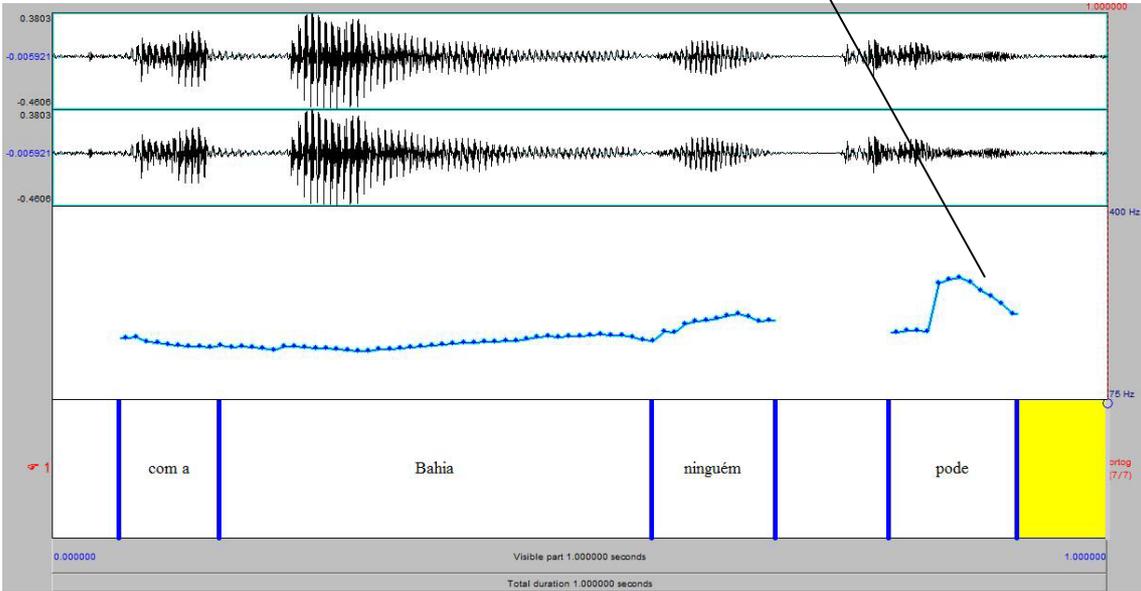


Gráfico 20 Trecho: “com a Bahia ninguém pode”

É interessante notar que o domínio do turno, de fato, é detido pela cantora. Os jornalistas só conseguem lançar nova questão quando, além de um fechamento total de caráter semântico do enunciado, dá-se também a queda da F0.

FRAGMENTO C.2:

“se você fizer a conta da passagem do tecido com as lantijola e as pedra as costura as agulha sai tudo em conta meu bem”

Neste momento, o tema é o investimento que a cantora faz em seu figurino. Ela acredita na importância da imagem; por isso, muitas vezes, viaja para fazer compras em Nova York. Segundo ela, fica mais barato e ainda é chique.

Mais uma vez, Ivete se alonga na resposta e interage muito com os participantes presentes no local da coletiva. Ela faz brincadeiras e fala de forma divertida. A cantora desenvolve completamente o tema de maneira a não deixar espaço para mal-entendidos por parte dos jornalistas e, conseqüentemente, interrupções ou tomadas de turno.

Prosodicamente, o movimento encontrado foi:

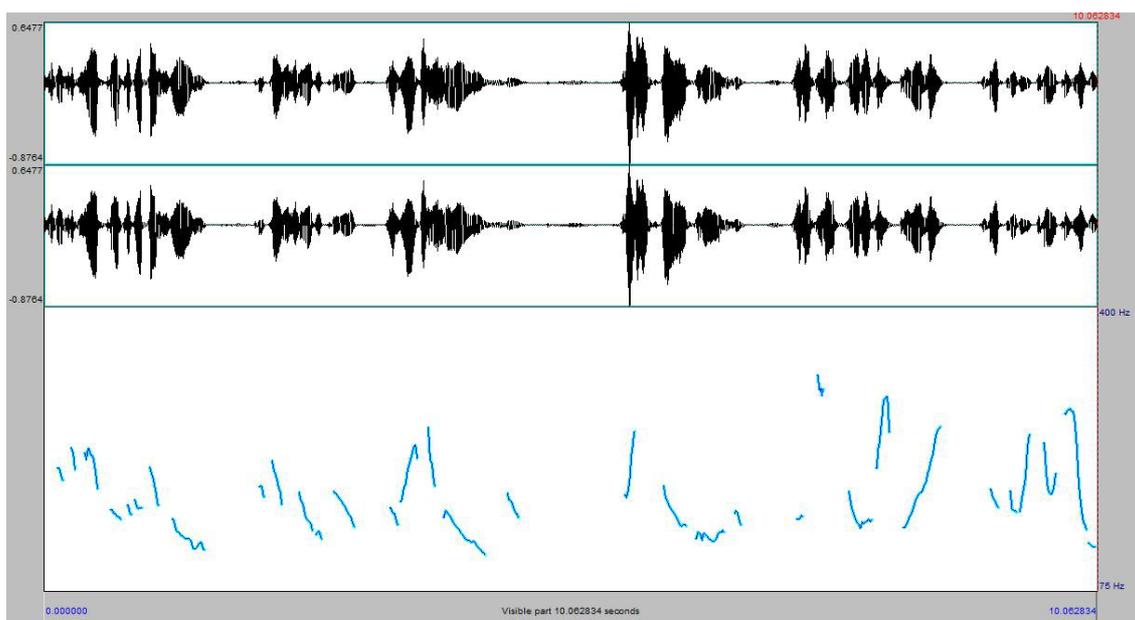


Gráfico 21 Fragmento C.2: “se você fizer a conta da passagem do tecido com as lantijola e as pedra as costura as agulha sai tudo em conta meu bem”

Notamos, no gráfico acima, uma configuração dinâmica com contornos ascendentes e descendentes. A sucessão de contornos ascendentes registrada se confirma na estratégia da cantora de não perder seu turno de fala.

O próximo gráfico mostra o mesmo trecho de modo a focalizar melhor o encerramento da fala:

Observação 1:
O lugar relevante de transição é caracterizado pelo contorno HL.

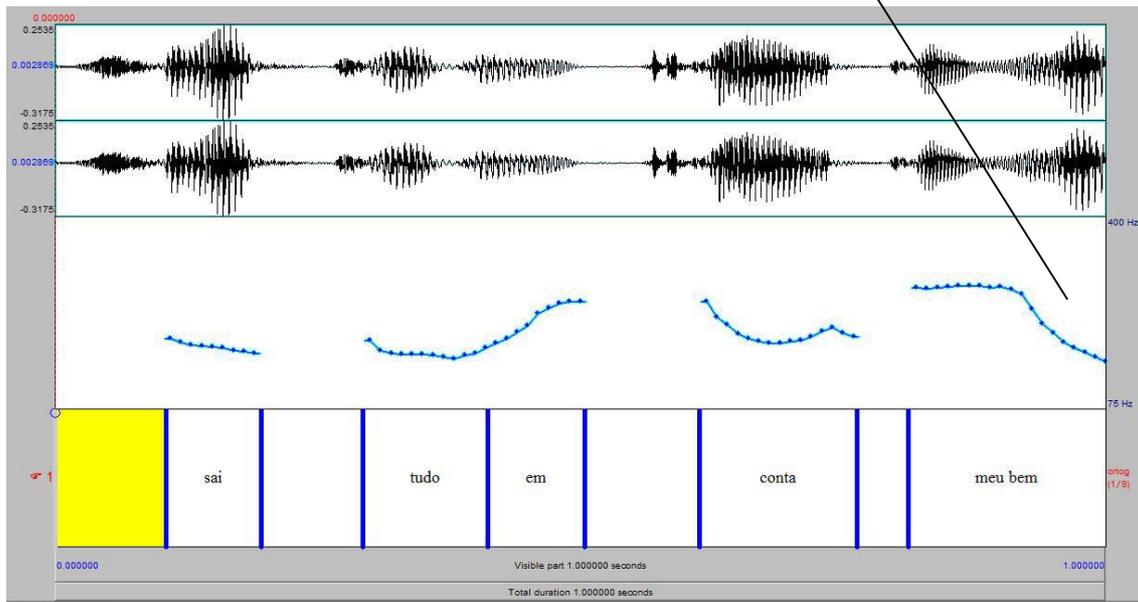


Gráfico 22 Trecho: “sai tudo em conta meu bem”

Vemos, então, que o encerramento que dá lugar ao novo turno coincide com um contorno descendente, isto é, a diminuição da F0. Mais uma vez, postulamos que a pista decisiva que guia o jornalista a perguntar é a entoação aliada com o sentido na interação. Só a partir dessas unidades em construção a negociação é realizada, e a conversa pode fluir. Analisemos o próximo exemplo da coletiva de Ivete Sangalo.

FRAGMENTO C.3:

“a relação do público de fora é muito intensa com o Brasil é maior do que a gente imagina do que a consciência da gente muito maior”

A pergunta feita para Ivete gira em torno da aceitação do carnaval baiano no exterior. Conforme responde a cantora, o público estrangeiro se identifica muito com a música e a cultura brasileiras. Segundo ela, a justificativa para essa recepção positiva é a de que nosso país é único e tem muita alegria a oferecer, o que atrai muito as pessoas.

Assim como nas outras respostas, Ivete não dá abertura para a tomada de turno, pois responde com clareza e desenvoltura a resposta. Não podemos deixar de mencionar que o tema sobre o qual se desenvolve a entrevista é o carnaval e o lançamento do trio elétrico da cantora. Na realidade, a coletiva de imprensa, para esse caso, serve como um meio de promover e divulgar o trabalho da artista.

O movimento prosódico encontrado é semelhante aos demais, sendo marcado por contornos ascendentes e descendentes e se encerrando com um HL. É o que veremos adiante:

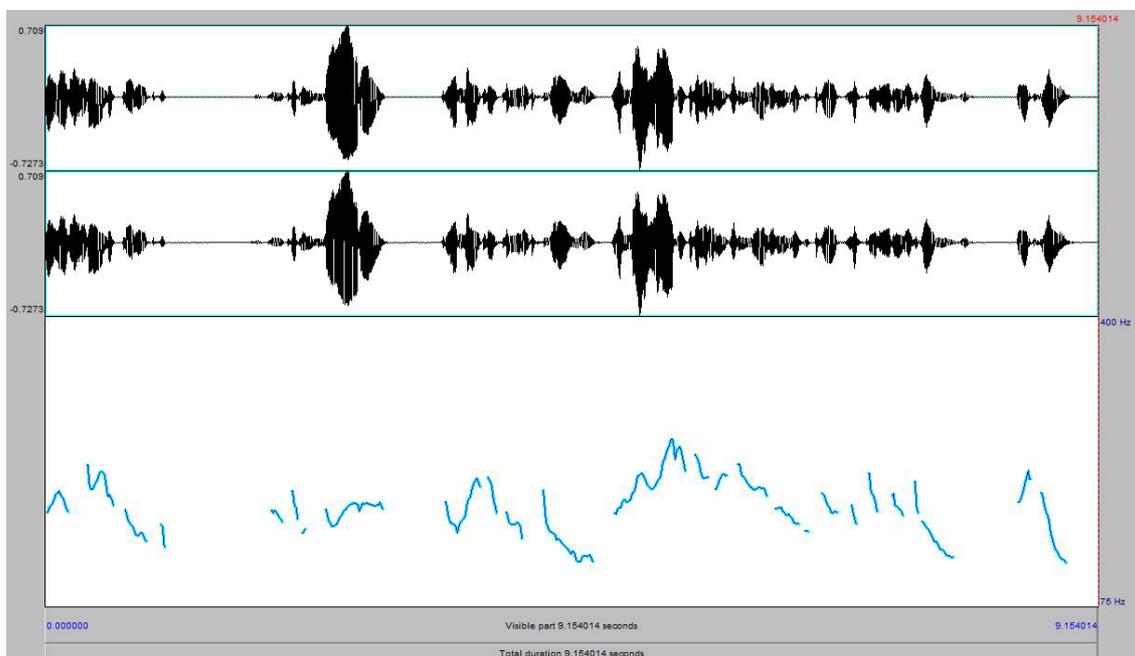


Gráfico 23 Fragmento C.3: “a relação do público de fora é muito intensa com o Brasil é maior do que a gente imagina do que a consciência da gente muito maior”

Destaca-se nesse desenho da melodia da fala uma série de contornos dinâmicos ascendentes e descendentes, sinalizando que a cantora prossegue a sua fala. Na unidade menor de análise, observamos:

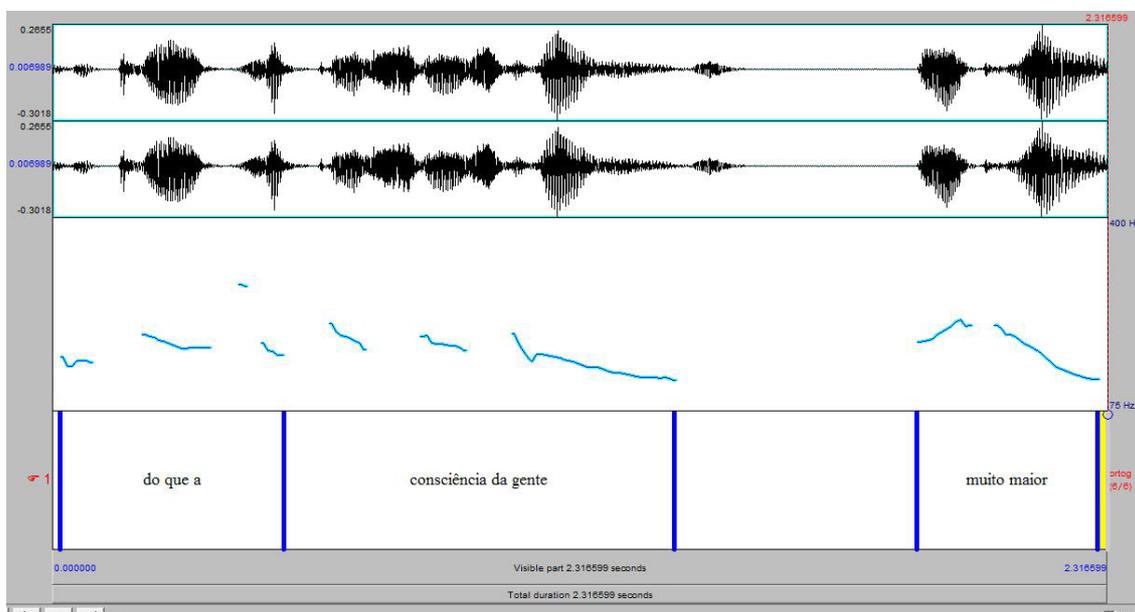


Gráfico 24 Trecho: “do que a consciência da gente muito maior”

Vemos aí que a F0 em descida, aliada ao aspecto semântico e pragmático, indica que finalmente a cantora conclui o turno, oferecendo a pista para a fala dos jornalistas. Passemos agora o último exemplo selecionado.

FRAGMENTO C.4:

“sempre estarei pronta pra desfrutar da companhia deles porque é um prazer será um prazer”

No último trecho selecionado, Ivete fala sobre a não-participação do bloco *Beijo* do carnaval 2009. A entrevistada diz sentir tristeza com relação a isso, pois se trata de um bloco muito tradicional do carnaval baiano. Ela espera que o bloco se reestruture para o próximo ano e anuncia que os participantes do *Beijo* serão convidados dela, pois fazem parte de sua história.

A mesma situação anterior ocorre nessa resposta. Ivete é muito consistente nos seus argumentos e desenvolve completamente seu raciocínio com relação à idéia principal do que está falando. Ela introduz, desenvolve e conclui o assunto, dominando o turno de fala e a vez de falar.

Quanto ao nível suprasegmental, o mesmo ocorre nos fragmentos anteriores, com a única diferença de que nesse trecho, as subidas e descidas de F0 não alcançam uma dimensão muito grande. Veremos o gráfico adiante:

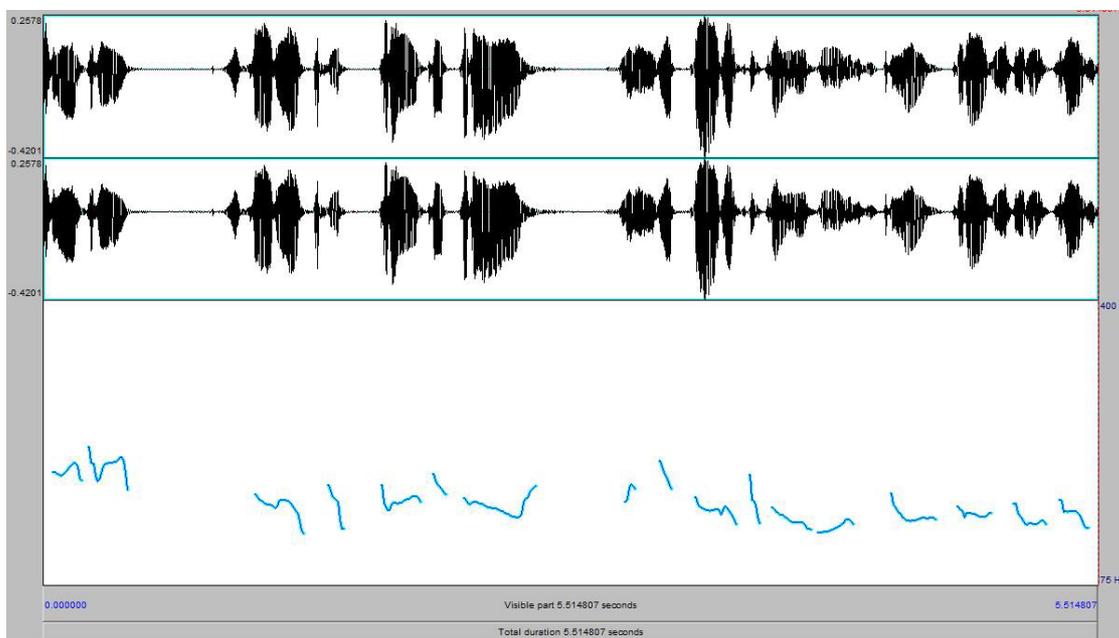


Gráfico 25 Fragmento C.4: ” sempre estarei pronta pra desfrutar da companhia deles porque é um prazer será um prazer”

E, por fim, um movimento mais alinhado, é encontrado no gráfico mais focalizado:

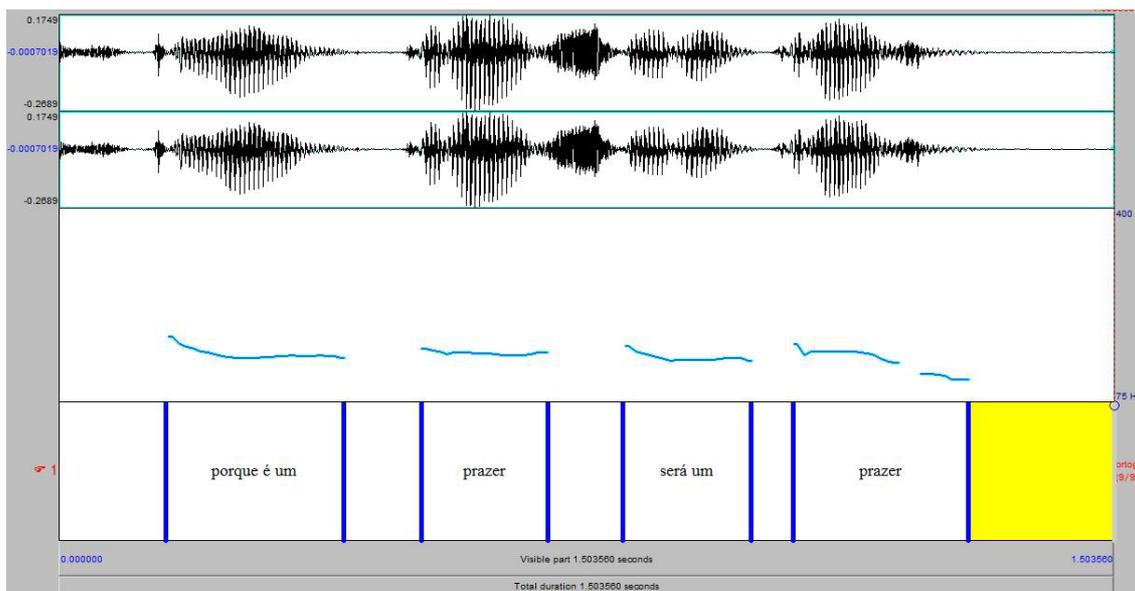


Gráfico 26 Trecho: “porque é um prazer será um prazer”

A submissão dos trechos da fala de Ivete no PRAAT geraram, assim como a fala do brigadeiro, uma sucessão de contornos ascendentes e descendentes, cada um dos quais sinalizando uma estratégia específica. O desenho descendente associa-se com a entrega de turno, assim como o ascendente, com o prosseguimento na “luta” conversacional. Com a primeira entrevista, as estratégias envolviam certa autoproteção discursiva do entrevistado, visto que estavam sendo cobradas ações dele; já na última entrevista, a da Ivete, parece haver uma autopromoção intencional. A particularidade da entrevista de Ivete é que mesmo dando um sentido completo à resposta da pergunta que lhe é feita, a cantora prossegue sua fala, fazendo com que os jornalistas tenham dificuldade em arrebanhar o piso conversacional.

A relação entrevistado “*entrevista*” entrevistador ilustra como o gênero coletiva de imprensa pode mostrar flexibilidade quanto a sua estrutura. Vimos que cada uma das entrevistas em A, B e C indicou uma configuração prosódica que, no emparelhamento com a cena interacional, atuou na construção do sentido. Um ponto crucial para se estabelecer essas distinções entre as entrevistas foi o assunto envolvido em cada caso. Em todas as coletivas, os jornalistas exerciam sua função de buscar informações, fontes de notícias etc. Contudo, em A, o acidente ocorrido criou um clima de muita tensão e angústia entre todos. O entrevistado, o brigadeiro, estava naquela posição cumprindo o dever de esclarecer as informações, já que essa era a tarefa que seu cargo exigia. Nas coletivas de Daniel e Álvaro, os participantes da cena estavam lá, pois seguiam um

protocolo que era a apresentação dos novos jogadores do clube. As perguntas dos jornalistas, em suma, eram esperadas para aquela situação.

Com Ivete Sangalo, era do interesse da cantora a divulgação do seu bloco, pois com a iminência do carnaval, as pessoas buscavam informações que condicionavam a compra dos abadás. A cantora, então, aproveita para se promover e usa de estratégias e de toda sua simpatia com os jornalistas. O clima da entrevista é de muita descontração. Nos trechos das respostas da cantora, encontramos oscilações de F0 no sentido de ela assegurar seu turno quando o contorno era ascendente. Voltando às analogias entre gestos corporais e gestos do aparelho fonador, nossa hipótese, de que a F0 associa-se a movimentos dos braços em uma luta, continua valendo. Nessa luta, todavia, Ivete é quem se movimenta mais, por isso permanece mais tempo em combate. Quando seus movimentos se tornam menos frequentes, isto é, a F0 da sua fala cai, os jornalistas conseguem direcionar questões a ela.

A luta “sonora”, nesse caso, não tem a carga de exigência de A. O ataque de Ivete envolve brincadeira, diversão, e não preocupação e tragédia como o do brigadeiro. Assim, há a possibilidade de diferentes mapeamentos com os mesmos domínios da coletiva e da guerra. Nesse novo contexto, o adversário que parte para o ataque é o entrevistado e não o entrevistador.

Para fechar nossa discussão acerca dos aspectos cognitivos envolvidos nesse jogo simbólico que inclui informações prosódicas, reafirmamos a importância da frequência das nossas ações no resultado final que elas alcançam. Além disso, ressaltamos o limite que o nosso corpo impõe durante todo o processo de produção e percepção da fala. Fechamos a análise estabelecendo a seguinte projeção: MAIS FREQUÊNCIA É MAIS CONTROLE DA SITUAÇÃO. Isso porque quem detém a F0 alta está possivelmente mais propenso a dominar a situação e continuar ou cessar seu turno de fala à medida que lhe convir.

Após todo o exposto, passaremos às considerações finais do presente trabalho.

5. CONSIDERAÇÕES FINAIS

Ao longo deste trabalho, realizamos primeiramente um apanhado dos pressupostos teóricos sobre os quais iríamos nos assentar para cumprirmos nossa proposta. Apresentamos as idéias principais da FC e sugerimos uma abordagem construcionista para o fonema. Além disso, chamamos a atenção para o funcionamento do nosso corpo licenciando e limitando o processo de produção e percepção da fala. Defendemos que há uma ação conjunta e coordenada quando damos melodia à nossa fala. Mencionamos teorias como a Fonologia Natural que investiga e explica a naturalidade dos processos fonológicos. Isso porque nos amparamos na idéia de que falantes produzem naturalmente ou inconscientemente padrões particulares de *pitch*.

Nesse momento, após a análise dos dados, podemos dizer se a hipótese de trabalho, de que o lugar relevante para transição do turno no gênero coletiva de imprensa configura-se por um contorno melódico descendente, confirma-se. Essa tendência de declínio das curvas entoacionais se deve, entre outras razões, ao desgaste de pressão subglotal que ocorre no fluxo da fala. Em linhas gerais, vimos que a pista que leva o jornalista a perguntar é a queda da F0 interagindo com um enunciado semanticamente completo. No caso do brigadeiro, por exemplo, as hesitações na sua fala contribuíam para a sobreposição de vozes e para a tentativa dos jornalistas de tomada de turno. Os movimentos de F0 alta revelaram situações em que os falantes queriam assegurar seus turnos e continuar a falar (o brigadeiro para se defender; e Ivete para se promover). O encerramento das ideias dos entrevistados coincidiam com o movimento descendente da F0, de maneira que, em muitos casos, categorizamos esse jogo interacional como uma entrega consentida do turno. Essa transição da vez de falar se daria após a apresentação completa dos argumentos do entrevistado. Vimos que o aspecto prosódico não é o único a atuar nesse processo, porquanto houve casos em que, mesmo com uma queda de F0 seguida de pausa, o turno se manteve com o entrevistado. Isso se deve ao fato de que estão em construção, de forma dinâmica, informações semânticas, sintáticas, pragmáticas e prosódicas. A passagem do turno pode, pois, se dar tanto de maneira suave quanto de maneira competitiva, e uma ou outra opção é refletida na configuração melódica, tendo a última contornos ascendentes e descendentes mais marcados e a primeira menos.

E o que dizer a respeito do que esses desenhos prosódicos simbolizam? A partir da metáfora DISCUSSÃO É GUERRA, apostamos na analogia entre os gestos

produzidos pelas pregas vocais e os gestos dos braços em um combate. Quanto mais frequentes os movimentos, maior o ataque. Assim sendo, cessando os movimentos, a tendência é o combatente se entregar ao adversário. Com isso, detectamos a seguinte projeção metafórica: MAIS FREQUENTE É MAIS FORTE. A persistência das ações das pregas vocais faz com que o falante se mantenha na conversa e detenha o poder de resposta. Podemos, então, inferir que, além de a pressão subglotal ser uma das causas do declínio do contorno melódico (embora não possamos provar isso com experimentos), os gestos envolvendo as pregas vocais apresentam-se também como fator atuante. Não fazemos aqui associações do tipo F0 alta indica este ou aquele afeto e sentimento, uma vez que não estamos estudando emoções. Todavia, em nosso contexto de pesquisa, F0 alta, de alguma forma, relaciona-se à resistência.

Por último, mas não menos importante, retomamos nossas ponderações acerca do alcance de nossos resultados. Nosso estudo não pretende impor asserções conclusivas sobre o tema. Assumimos que há ainda muito que se fazer e conhecer a respeito. Para tanto, lançamos mão do presente trabalho como uma primeira incursão por esses caminhos.

Referências

BARBOSA, P. A. **Incursões em torno do ritmo da fala**. Pontes: Campinas, 2006.

BAUDOUIN DE COURTENAY, J. An attempt at a theory of phonetic alternations. In: STANKIEWICZ, E (ed). **A Baudouin de Courtenay Anthology**: The beginnings of structural linguistics. London: Indiana University Press, 1972.

BOLINGER, D. L. Intonation across languages. In: GREEMBERG, J. H; FERGUSON, C. A; MORAVESIK, E. A. **Universals of Human Language**, v. 2, Stanford: University Press, 1978.

BYBEE, J. **A view of phonology from a cognitive and functional perspective**. *Cognitive Linguistics*, v. 5, 1994. p. 285-306.

CLARK, H. **Using language**. Cambridge: Cambridge University Press, 2000.

COUPER-KUHLEN, E. **An introduction to english prosody**. Tübingen: Niemeyer, 1986.

CRYSTAL, D. **Your turn**. *English today*, n. 32, 1992.

CROFT, W. **Radical Construction Grammar**: syntactic theory in typological perspective. New York: Oxford University Press, 2001.

_____; CRUSE, D. A. **Cognitive Linguistics**. Cambridge: Cambridge University Press, 2004.

CRUTTENDEN, A. **Intonation**. 2. ed. Cambridge: Cambridge University Press, 1997.

FAUCONNIER, G. **Mappings in language and thought**. Cambridge: Cambridge University Press, 1997.

FELTES, H. P. M. **Semântica cognitiva**: Ilhas, pontes e teorias. Porto Alegre: Edipusrs, 2007.

FILLMORE, C. J. **Lectures on deixis**. Stanford, CA: CSLI publications, 1997.

_____; KAY, P; O'CONNOR, M. C. **Regularity and idiomaticity in grammatical constructions**: the case of let alone. *Language*, 1988.

GOLDBERG, A. **A construction grammar approach to argument structure**. Chicago: The University of Chicago Press, 1995.

GONÇALVES, C. A. V. **Focalização no Português do Brasil**. Tese (doutorado). Rio de Janeiro: UFRJ, 1997.

GUSSENHOVEN, C. **The phonology of tone and intonation**. Cambridge: Cambridge University Press, 2004.

HIRST, D. **Form and function in the representation of speech prosody**. Speech communication, v. 46, 2006. p. 334-347.

_____; DI CRISTO, A. A survey of intonation systems. In: _____ (eds). **Intonation systems: a survey of twenty languages**. Cambridge: Cambridge University Press, 1998. p. 1-44.

LAKOFF, G. **Women, fire and dangerous things**. Chicago: University of Chicago Press, 1987.

_____. Cognitive Phonology. In: GOLDSMITH, J. **The last phonological rule**. Chicago: Chicago University Press, 1993.

LAKOFF, G; JOHSON, M. **Metaphors we live by**. Chicago: Chicago University Press, 1980.

_____; **Philosophy in the flash: the embodied mind and its challenge to western thought**. New York: Basic books, 1999.

LOCAL, J. **Phonetic detail and the organization of talk-in-interaction**. XVI ICPs. Saarbrücken, 2007.

LUCENTE, L. **DaTo: um sistema de notação entoacional do português brasileiro baseado em princípios dinâmicos**. Dissertação (Mestrado em Linguística). Campinas: Unicamp, 2008.

MARCUSCHI, L. A. **Análise da conversação**. 4. ed. São Paulo: Editora ática, 1998.

MARTINS, H. **Sobre linguagem e pensamento no paradigma experiencialista**. Revista Veredas, v. 6, n. 2, Juiz de Fora, 2002. p. 85-100.

MATEUS, M. H. M. **Estudando a melodia da fala: traços prosódicos e constituintes prosódicos**. Setúbal: Encontro sobre o ensino das línguas e a Linguística, 2004. Disponível em: <http://www.iltec.pt/pdf/wpapers/2004-mhmateus-prosodia.pdf> . Acesso em: 26 jul. 2008.

MEIRELES, A. R. **Reestruturações rítmicas da fala no português brasileiro**. Tese (Doutorado em Linguística). Campinas: Unicamp, 2007.

MORAES, J. A. **Os fenômenos supra-segmentais no português do Brasil**. 1998. Mimeo.

MORI, A. C. Fonologia. In: MUSSALIN, F; BENTES, A.C (orgs). **Introdução à Linguística: domínios e fronteiras**. São Paulo: Cortez, 2001.

NATHAN, G. **Phonemes as Mental Categories**. In: Proceedings of the 12 th Annual Meeting of the Berkeley Linguistics Society, v. 12, 1986. p. 212-224.

_____. **Is the phoneme Usage-Based?**. International Journal of English Studies. 6 v. Murcia: Universidad de Murcia, 2006.

_____. Phonology. In: GEERAERTS, D; CUYCKENS, H. **The Oxford Handbook of Cognitive Linguistics**. New York: Oxford University Press, 2007.

_____. **Phonology: a cognitive grammar introduction**. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2008.

_____. **Where is the Natural Phonology Phoneme in 2009?**. In: Poznan studies in contemporary linguistics, v. 45, Warsaw: Versita, 2009.

OHALA, J.J. The frequency code underlies the 'sound-symbolic use of voice pitch. In: HILTON, L; NICHOLS, J; OHALA, J (orgs). **Sound symbolism**. Cambridge: Cambridge University Press, 1994. p. 325-347.

PATEE, H. H. **The physics of symbols and the evolution of semiotic controls**. In: Proceedings from the workshop on control mechanisms for complex systems: issues of measurement and semiotic analysis. Santa fé Institute Studies in the Sciences of complexity. Redwood CITY, CA: Addison Wesley, 1997.

RACZASZEK-LEONARDI, J; KELSO, J. A. S. **Reconciling symbolic and dynamic aspects of language: toward a dynamic psycholinguistics**. New Ideas in Psychology, 2007.

REICH, U. **Batidas e contornos**. Minicurso Pragmática e Entoação apresentado no VI Congresso Internacional da ABRALIN. João pessoa, 2009.

ROCHA, L. F. M. **A construção da mimesis no reality show: uma abordagem sociocognitivista para o discurso reportado**. Tese (Doutorado em Linguística). Rio de janeiro: UFRJ, 2004.

_____. Proposta de dissertação [mensagem pessoal]. Mensagem recebida por patriciarvcoutinho@yahoo.com.br em 18 março 2008.

SACKS, H; SCHEGLOFF, E; JEFFERSON, G. **A simplest systematics for the organization of turn taking for conversation**. Language, v. 50, n. 4. In: Veredas, v. 7, n. 1 e 2. Tradução para o português: Maria Clara Castellões de Oliveira e Paulo Cortes Gago. Juiz de Fora, 2003.

SALOMÃO, M. M. M. **A questão da construção do sentido e a revisão da agenda dos estudos da linguagem**. Revista Veredas, v. 3, n. 1, Juiz de Fora, 1999. p. 61-79.

_____. **Teorias da linguagem: a perspectiva sociocognitiva**. II Fórum de Linguagem: Linguagem natureza e cultura. Rio de Janeiro, 2006.

SAPIR, E. The psychological reality of phonemes. In: MAKKAI, V. B. **Phonological theory: evolution and current practice**. New York: Holt, Rinehart and Winston, 1972.

SAUSSURE, F. **Curso de Linguística Geral**. São Paulo: Cultrix, 1974.

STAMPE, D. **The acquisition of phonetic representation**. Chicago Linguistic Society, v. 5, 1969.

SILVA, A. S. **A linguística cognitiva**: uma breve introdução a um novo paradigma em linguística. Revista Portuguesa de Humanidades. Braga: Faculdade de Filosofia da UCP, 1997.

_____. **Os estudos de Linguística Cognitiva do Português**. Revista Portuguesa de Humanidades. Braga: Faculdade de Filosofia da UCP, 2007.

SOUZA, J. C. P; CAGLIARI, L.C. **Pequena abordagem sobre fonologia**. Revista universitária Vale do Rio Verde, v. 4, n. 2, Três Corações, 2001.

TAYLOR, J. **Where do phonemes come from?** A view from the bottom. International Journal of English Studies, v. 6, 2006.

TOMASELLO, M. **Origens culturais da aquisição do conhecimento humano**. Trad. Claudia Berliner. São Paulo: Martins Fontes, 2003.

VÄLIMAA-BLUM, R. **Cognitive Phonology in Construction Grammar**. Analytic Tools for Students of English. Berlin: Mouton de Gruyter, 2005.

WERMKE, K. Bebês choram no idioma materno, diz estudos. **Uol**, nov. 2009.

Disponível em:

<<http://noticias.uol.com.br/ultnot/cienciaesaude/ultnot/efe/2009/11/05/bebes-choram-no-idioma-materno-diz-estudo.jhtm>>. Acesso em: 07 nov. 2009.

YOUTUBE. Disponível em: <http://www.youtube.com> . Acesso em: 26 março. 2008.