



A segmentação da fala espontânea: aspectos prosódicos, funcionais e aplicações para a tecnologia

Spontaneous Speech Segmentation: Functional and Prosodic Aspects with Applications for Automatic Segmentation

Plínio A. Barbosa

Universidade Estadual de Campinas, Campinas, São Paulo / Brazil

pabarbosa.unicampbr@gmail.com

Tommaso Raso

Universidade Federal de Minas Gerais, Belo Horizonte. Minas Gerais /Brazil

tommaso.raso@gmail.com

Este número da *Revista de Estudos da Linguagem* é dedicado a um tema enfrentado em diversas outras iniciativas promovidas pelos organizadores junto com colegas de outros países. O tema, que nos últimos anos tem adquirido um papel cada vez mais importante nas disciplinas que estudam a produção e a percepção da fala, é a segmentação da fala em unidades menores, vista sob perspectivas tanto formais quanto funcionais, tanto fundamentalmente teóricas quanto com foco mais empírico. Entre as principais iniciativas, citamos:

- Dois workshops internacionais (o IV Leel e o X Lablita International workshop *Unit of Reference for Spontaneous Speech and their Correlation Across Language*, realizado em agosto de 2015 na UFMG; e o workshop *Spoken Corpora advances: prosody as the crux of speech segmentation, annotation and multilevel linguistic studies*, organizado na Cidade do Cabo em junho de 2018, dentro do 20º International Congress of Linguists);

- O livro *In Search for a Reference Unit of Spoken Language: A Corpus Driven Approach*, a ser lançado em breve pela editora John Benjamins;
- Um número especial do *Journal of Speech Sciences* programado para sair em meados de 2019.

Todas essas iniciativas são dedicadas ao tema da segmentação prosódica da fala, um tema que tem cada vez mais se tornado central para entender a estruturação da fala em diversos níveis, bem como a relação dessa estruturação com as funções comunicativas da linguagem. As disciplinas interessadas no tema, e a linguística *in primis*, têm se desenvolvido enormemente tanto a partir da contribuição dos avanços tecnológicos e da estatística aplicados aos estudos linguísticos, quanto da contribuição dos avanços das próprias teorias linguísticas. De fato, até recentemente, o estudo da segmentação da fala considerava a segmentação, quase exclusivamente, da chamada fala de laboratório (*lab speech*), que inclui a fala lida e a fala eliciada sob várias formas (XU, 2010), a partir da manipulação de eventos externos pelo pesquisador (como pela proposta de tarefas com um ou mais participantes como *map task* e jogos eletrônicos, pela condução de entrevistas sobre temas específicos, *inter alia*). Há alguns anos, no entanto, tornou-se possível abordar a fala sem roteiro prévio (*non scripted speech*) extraída de *corpora* de fala espontânea em situações comunicativas naturais variadas e com boa qualidade acústica. Nesse artigo introdutório ao número temático apresentamos um panorama, mesmo que parcial, das questões científicas em jogo, dos resultados alcançados até aqui e dos passos que já se anunciam para o futuro.

1. Segmentação prosódica: entre forma e função

Contrariamente à escrita, que é um produto que permanece no tempo e no espaço, a fala é um processo cujo resultado desaparece logo após a sua manifestação, se não consideramos nesse exame as tecnologias que realizam seu registro. Apenas permanecem como subproduto imediato algumas consequências cognitivas do discurso, mas não a fala em si (LINELL, 2005; BLANCHE-BENVENISTE; JEANJEAN, 1987). Ausente na escrita em sua manifestação acústica, a não ser por meros indicativos pelos sinais de pontuação, a prosódia é o componente essencial para os estudos de segmentação da fala. Hoje é possível, graças

à tecnologia e a softwares dedicados, reproduzir a fala por quantas vezes se achar necessário e realizar procedimentos de anotação que permitem delimitar diferentes unidades para poder estudá-la: sílabas, grupos de sílabas ou palavras, unidades prosódicas de diferente dimensão e estatuto teórico, bem como sequências de enunciados de interesse. Isso permite a observação sistemática e a medição de muitos aspectos da fala que, sem a tecnologia, haviam sido em certa medida apenas intuídos através da sensibilidade auditiva dos precursores da prosódia contemporânea (cf. PIKE, 1945; LIEBERMAN, 1960; BOLINGER, 1965), mas que não podiam ser nem aprofundados nem demonstrados. Entre esses aspectos, um lugar de importância primária é ocupado pelos diferentes componentes em que é possível segmentar o fluxo da fala e pela reflexão sobre suas formas e funções. Enfim, se tornou possível tentar a reconstrução da complexa estrutura prosódica (e não somente) da fala humana.

Além disso, a tecnologia tornou possível compilar e investigar grandes quantidades de dados de fala, tratados e anotados de diferentes maneiras e adequados especificamente a estudos dos mais variados, numa linha de pensamento que privilegia a obtenção de conhecimento a partir de *corpora* bastante extensos, no que se convencionou chamar hoje de “*big data*” (cf. FURHT; VILLANUSTRE, 2016). O tratamento informatizado do sinal acústico nos permite segmentar o discurso em unidades menores, desde o enunciado (ou talvez desde unidades maiores como os “parágrafos”) até a sílaba e os seus constituintes; e nos permite investigar como a fala humana veicula as fronteiras (ou a ausência delas) em diferentes níveis hierárquicos.

Dependendo do interesse de estudo, a fala pode ser segmentada em unidades de diferentes tamanhos e naturezas, cada uma delas sendo capaz de mostrar algumas de suas propriedades e cada uma delas sendo delimitada por algum tipo de fronteira. Considerando apenas as unidades acima do nível da palavra, podemos dividir a fala em grupos acentuais (ou pés *n*-ários, os grupos de sílabas até uma tônica, no caso de línguas com cabeça à direita), em unidades prosódicas chamadas de entonacionais ou tonais ou de grupos prosódicos, em enunciados, ou, em uma perspectiva de natureza sintática, em sintagmas entoacionais (IP), orações e sentenças. Cada tipo de segmentação está direta ou indiretamente associado a uma visão teórica, mas em muitos casos isso não impede uma investigação empírica, cujos resultados podem ser analisados à luz de diferentes perspectivas teóricas. Nos últimos anos, vários *corpora* com a anotação

prosódica da fronteira foram compilados para diferentes línguas (AURAN *et al.*, 2004; DU BOIS *et al.*, 2000-2005; OSTENDORF *et al.*, 1996; CRESTI; MONEGLIA, 2005; SCHUURMAN *et al.*, 2003; IZRE'EL, 2002; RASO; MELLO, 2012 e em preparação; METTOUCHI *et al.*, 2010; GAROFOLO, *et al.*, 1993,).

Qualquer segmentação implica a presença de uma fronteira, seja ela de fato percebida ou proposta teoricamente. Sendo assim, a fronteira pode ser entendida como uma ruptura fisicamente percebida; pode se referir a um limite verificável para a realização de fenômenos linguísticos; e pode ainda ser considerada em uma região entre duas unidades, sendo que essa região pode ser de natureza perceptível ou não.

Este número temático busca estudar a segmentação enfocando o que pode ser considerado como a unidade de referência do processo da fala (IZRE'EL *et al.*, no prelo). A própria noção de unidade de referência pode ser entendida de diferentes maneiras, mas provisoriamente podemos defini-la como uma unidade mínima de sentido completo e autônomo comunicativamente que compõe um texto falado (CRESTI, 2000; MONEGLIA; RASO, 2014). Essa definição não é incontroversa, mas nos permite começar a busca.

Todos os tipos de unidades nomeadas acima, independentemente de como são definidos, são separados por fronteiras que são definidas dando um peso maior ou menor à percepção ou à teoria; dificilmente um dos dois critérios de individualização exclui completamente o outro. Nos trabalhos deste número temático está sempre presente uma base perceptiva, mas algumas contribuições dão um peso maior que outras aos aspectos teóricos, sendo que esses aspectos variam de uma contribuição para outra. Com essas diferenças de perspectiva muda também o conceito de fronteira.

De natureza teórica são as fronteiras de constituintes de abordagem sintática ou informacional; isso não significa que elas não possam ser associadas a fronteiras de caráter prosódico, que constituem o interesse primário desse número temático. De fato, entendemos que a prosódia guia a interpretação sintática, como em casos como o da sentença “A ovelha de raça brasileira.” A partir dessa unidade da escrita, dois enunciados podem ser emitidos segundo duas formas de agrupamento distintas, em que “/” representa uma fronteira não terminal forte:

[A ovelha de raça] / [brasileira] vs. [A ovelha] / [de raça brasileira]

No primeiro caso se trata de uma ovelha de raça não informada nascida no Brasil e, no segundo caso, de uma ovelha que é de raça desenvolvida no Brasil. São assim os constituintes prosódicos que permitem a escansão adequada da estrutura sintática de cada sentença. Isto é, a prosódia permite a desambiguação entre as duas interpretações possíveis, pois os limitados recursos da escrita não permitem resolver a distinção. Nesse caso, a prosódia guia a interpretação sintática e os constituintes sintáticos e prosódicos são consequentemente congruentes, isto é, têm os mesmos limites. Por conta disso os autores deste número temático que tratam diretamente da questão da segmentação da fala consideram unidades que são constituintes prosódicos.

Assim, quase todas as contribuições aqui estudam a organização da fala em unidades que podem ser consideradas como extensíveis a unidades entonacionais. Por conta disso, nesta apresentação utilizaremos a expressão “unidade entonacional” de forma geral, mas deixamos claro que essa unidade se estrutura com base em parâmetros que não incluem apenas a frequência fundamental (f_0), mas também parâmetros de natureza duracional, de intensidade e possivelmente de qualidade de voz. Um único trabalho (o de Ph. Martin) segmenta a fala em grupos acentuais, o que não exclui o fato de um grupo acentual ou um conjunto de grupos acentuais coincidir com a unidade entonacional. A segmentação em grupos acentuais pode, portanto, ser vista também como a oportunidade de investigar a estrutura interna da unidade entonacional, enriquecendo assim, e não contradizendo, as perspectivas que preferem se concentrar na análise da unidade entonacional.

É difícil definir a unidade entonacional sem fazer referência ou à percepção ou a um postulado de natureza teórica. Mas em geral a unidade entonacional é definida como o grupo de palavras (pode ser também uma única palavra e, em casos mais raros em que entra em jogo a ênfase em sílabas, menos de uma palavra. Nesse último caso, a fronteira é uma consequência perceptiva da proeminência da unidade) delimitado entre uma fronteira prosódica e outra, gerando um contorno entonacional coerente e separado fisicamente e perceptivamente dos contornos precedente e seguinte (DU BOIS *et al.*, 1992, p. 17; CRUTTENDEN, 1997). Essa definição mascara algumas dificuldades em capturar as propriedades de uma unidade entonacional sem fazer referência à fronteira, e por sua vez sem identificar a fronteira a partir do conceito de unidade entonacional, com evidente risco de circularidade. A própria

definição de “contorno coerente” não é completamente satisfatória, já que não sabemos com clareza quais são os parâmetros que permitem ou rompem a coerência.

Do ponto de vista funcional, a unidade entonacional pode ser estudada e definida linguisticamente com base em perspectivas diferentes. As principais são a perspectiva sintática, a perspectiva informacional (CHAFE, 1994; RASO; MELLO, 2014) e a perspectiva conversacional (BARTH-WEINGARTEN, 2016). Mas a própria individualização da unidade entonacional é problemática. De fato, não é sempre óbvio reconhecer um perfil prosódico coerente ou uma fronteira prosódica. No que tange a identificação de uma fronteira, geralmente os estudos se baseiam no acordo estatístico entre segmentadores: um determinado trecho de fala é oferecido para um certo número de segmentadores e se compara o acordo que eles tiveram em segmentar, de oitiva, o trecho em unidades menores. Outras abordagens consideram a percepção associada a alguns traços formais visíveis a partir de ferramenta de software, como o que é chamado de tom de fronteira (*boundary tone*), um determinado movimento de f0 alinhado ao final da unidade, na linha de investigação da Fonologia Métrica-Autossegmental (LADD, 1996; PIERREHUMBERT, 1980).

Testes estatísticos de análise da coerência entre avaliadores (segmentadores) mostram que o acordo na identificação das fronteiras, e por consequência das unidades, é muito alto (superior a 80 %, especialmente para o caso das fronteiras terminais; MELLO *et al.*, 2012; MONEGLIA *et al.*, 2005; YOON *et al.*, 2004; BUHMANN *et al.*, 2002). É, portanto, consensual que um importante nível de organização da fala seja constituído pela unidade entonacional. As razões dessa organização, ao contrário, são controversas: segundo alguns autores (cf. COWAN, 1998) essa segmentação do fluxo da fala é devida aos limites de memória, que impõem agrupamentos de um número limitado de sílabas para o processamento linguístico. Segundo outros, as unidades seriam devidas a motivações cognitivas (CHAFE, 1994; CROFT, 1995; BYBEE, 2010). Segundo outros ainda, a segmentação corresponde a unidades de natureza sintática e, portanto, fronteira prosódica e fronteira sintática seriam correlacionadas, especialmente nas abordagens de natureza fonológica da prosódia que pressupõem um mapeamento entre constituintes sintáticos e os limites de unidades prosódicas (NESPOR; VOGEL, 1986; SELKIRK, 1995). Uma quarta proposta, dominante nesse número temático, atribui à fronteira prosódica o valor de delimitar unidades de natureza

informacional, independentemente de sua organização sintática. Outros ainda vêm uma correspondência entre a prosódia e unidades de outro domínio discursivo (COUPER-KUHLEN, 2004; SCHEGLOFF, 1998). Quem estuda a prosódia como correlacionada a domínios linguísticos de natureza não sintática tende também a considerar a prosódia como um elemento estrutural implementado antes dos elementos segmentais (cf. a teoria *Frame/Content* de MacNEILAGE, 1998). Uma visão interessante dentro dos estudos prosódicos (HIRST; Di CRISTO, 1998; BARBOSA, 2006) tenta uma conciliação entre constituintes sintáticos e prosódicos, propondo que a sintaxe se limitaria a impor algumas restrições, mas não determinaria a posição das fronteiras: estas poderiam, portanto, aparecer somente em posições compatíveis com a estruturação sintática, sem por isso marcar necessariamente constituintes dessa natureza, já que, dada uma mesma sentença, seriam várias as posições compatíveis com a estruturação sintática onde poderia ser colocada uma fronteira, com cada posição sinalizando uma interpretação cognitivo-informacional diferente. Por outro lado, muitos estudiosos da sintaxe estão percebendo como a prosódia é essencial para o funcionamento de estruturas que apresentam fortes dificuldades para as explicações sintáticas tradicionais. É o caso do fenômeno da assim chamada de insubordinação (EVANS; WATANABE, 2016; BOSSAGLIA *et al.*, no prelo). Nesses casos, a interpretabilidade da estrutura depende em maneira decisiva da sua codificação prosódica.

2. As principais questões metodológicas

As pesquisas realizadas mostram também que o estudo das fronteiras prosódicas depende da tipologia de fala e em parte da tipologia do texto falado. De fato, até recentemente, as pesquisas se concentraram no estudo da segmentação prosódica em textos lidos ou sequências limitadas executadas em laboratório, com resultados interessantes, mas que não parecem ser comparáveis com o que acontece na fala espontânea, objetivo prioritário desse número temático. É frequente em estudos de prosódia ligados à sintaxe e à fonologia que a fala de laboratório seja usada para testar relações entre prosódia e sintaxe (como no caso de desambiguação ou na investigação dos possíveis tipos de constituintes isolados por fronteiras), apresentando um número muito menor de variáveis do que a fala espontânea e uma maior previsibilidade (PRICE *et al.*, 1991). Ressalta-se ainda que, quando essa fala é lida, ela é a

realização sonora de um texto escrito, e, portanto, estruturado com base em princípios diferentes daqueles da fala espontânea.

Recentemente, alguns trabalhos sobre a fala espontânea obtiveram bons resultados na investigação dos mecanismos de segmentação, seja observando um acordo alto (maior que 80%) entre os segmentadores humanos nessa tarefa (MELLO *et al.*, 2012; MONEGLIA *et al.*, 2005; TEIXEIRA FALCÃO, 2017), seja criando softwares capazes de segmentar automaticamente textos alcançando resultados altamente comparáveis com as tarefas realizadas pelos humanos (AVANZI *et al.*, 2008; NI *et al.*, 2012; BARBOSA, 2016).

A criação de softwares capazes de automatizar a segmentação prosódica em unidades entonacionais (cf. MITTMAN; BARBOSA, 2016) só é possível porque a investigação dos parâmetros acústicos responsáveis pela percepção de fronteira tem avançado muito, graças também aos trabalhos realizados sobre a fala lida e sobre sequências realizadas em laboratório que permitiram uma primeira compreensão dos fenômenos em jogo, que são altamente complexos. Parece de fato que os parâmetros responsáveis pela nossa percepção de fronteira são diversos, que nem sempre são todos co-presentes, que seu peso pode variar dependendo das línguas e das circunstâncias da fala, levando também a questionar se é possível falar de fronteira como categoria homogênea ou se não é o caso de falar de tipos diferentes de fronteiras.

Na literatura os parâmetros que são mais mencionados são tanto de frequência fundamental (f_0), quanto de duração e de intensidade, além de parâmetros que se referem à qualidade de voz (BARTHWEINGARTEN, 2016; MO *et al.*, 2008; WAGNER; WATSON, 2010), especialmente laringalização, *creaky voice* em inglês (DILLEY *et al.*, 1996; GORDON; LADEFOGED, 2001; REDI; SHATTUCK-HUFNAGEL, 2001; HANSON *et al.*, 2001; CARLSON *et al.*, 2005). Os principais são os seguintes: a pausa silenciosa, que aqui simplesmente chamaremos de “pausa” (adiante falaremos da pausa preenchida), cuja presença parece automaticamente veicular a percepção de fronteira (MARTIN, 1973; SWERTS, 1997; SHRIBERG *et al.*, 2000; TSENG; CHANG, 2008; MO; COLE, 2010; TYLER, 2013); o alongamento das sílabas finais da unidade, ou seja uma redução da taxa de realização das últimas sílabas antes de uma fronteira (WIGHTMAN *et al.*, 1992; BARBOSA, 2008; MO *et al.*, 2008; FUCHS *et al.*, 2010; FON *et al.*, 2011; TYLER, 2013); a redução duracional das primeiras sílabas da

unidade, ou seja, a aceleração imediatamente após uma fronteira (AMIR *et al.* 2004; TYLER, 2013), correlacionada com fenômenos de *anacrusis*; o *reset* da curva de f_0 ; a mudança brusca de direção da curva de f_0 ; a mudança de intensidade no início da unidade prosódica (SWERTS *et al.*, 1994; TSENG; FU, 2005; MO, 2008); a laringalização (*creaky voice*) e talvez outras qualidades de voz não modais. A esses parâmetros, pelo menos para algumas línguas, devem ser acrescentados alguns fenômenos de natureza segmental. Por exemplo, para o inglês pode ser importante a soltura da oclusiva final ou a laringalização e golpe de glote de alguns segmentos finais como marcadores de fronteira.

Cada uma dessas pistas traz problemas para o pesquisador. Por exemplo, a pausa, que intuitivamente parece uma noção óbvia, não é identificada consensualmente: qual é o tempo mínimo de silêncio para considerar a presença de uma pausa? Como a presença dessa pausa afeta os outros parâmetros responsáveis pela percepção de fronteira? A pausa é um indicador também do tipo de fronteira ou não? Quanto à curva de f_0 , qual a contribuição relativa da diferença de nível de f_0 , de sua excursão, da direção de seu movimento e da taxa de sua variação? Já quando se considera a duração silábica, qual a extensão da região afetada pela fronteira, medida em número de sílabas? E se a mudança de duração envolve mais do que a simples sílaba fronteira, a mudança acontece na mesma proporção em cada sílaba envolvida ou não? Além disso, os trabalhos experimentais mostraram que, para avaliar de maneira confiável as medidas de natureza duracional, é necessária alguma forma de normalização que coloque à parte as propriedades intrínsecas dos segmentos, que, nesse caso, influem de maneira decisiva sobre a duração (BARBOSA, 2012). É preciso salientar ainda que a medida da duração propícia à análise prosódica deve considerar a realização e a fronteira das sílabas, o que envolve as noções de sílaba fonológica e sílaba fonética. A primeira é importante para a percepção da taxa de elocução, porque perpassa pela apreensão da sílaba pelo sistema cognitivo, enquanto a segunda fundamenta a produção da cadeia de fala e a organização de consoantes e vogais na sílaba produzida.

A pesquisa sobre os parâmetros acústicos que, no seu conjunto, veiculam a percepção de quebra (fronteira) deve considerar o peso ou contribuição relativa de cada pista acústica. Para isso, é importante considerar não somente que cada pista é perceptível somente se passar de um certo limiar, mas que esse limiar varia variando as outras pistas

(t'HART *et al.*, 1990). Isso significa, em primeiro lugar, que nós não somos capazes de perceber qualquer mudança de f_0 ou qualquer mudança de duração ou intensidade, mas somente as mudanças que ultrapassam um determinado limiar. Embora para cada parâmetro ou pista em isolado possamos conhecer a *Just Noticeable Difference* (JND), ou seja a variação mínima desse parâmetro que podemos perceber (cf. HUGGINS, 1972; KLATT; COOPER, 1975 para duração segmental; t'HART, 1981; RIETVELD; GUSSENHOVEN, 1985 para f_0 ; KOFFI, 2018 para intensidade), bem como a forma pela qual a JND varia com a partir da modificação de um outro parâmetro (por exemplo como percebemos a variação de intensidade em frequências diferentes), pouco sabemos ainda sobre como essas combinações complexas de parâmetros variam com relação à capacidade de veicular percepção de fronteira. Não é simples modelar o efeito de fronteira nas combinações de tantos parâmetros no fluxo da fala. De fato, não seria surpreendente se o peso de uma pista mudasse mudando as combinações de pistas nas quais está inserido, ou mudando os contextos de fala em que aparece: leitura ou fala espontânea, ou diferentes estilos de fala espontânea, ou ainda diferentes funções linguísticas das unidades delimitadas pelas fronteiras marcadas, sem considerar variações ligadas às características dos falantes.

De fato, os diversos estudos em diferentes línguas confirmam a importância das pistas apontadas acima para a percepção de fronteira, enquanto também revelam que cada uma dessas pistas atua com peso diferente para assinalar essa mesma fronteira (TEIXEIRA FALCÃO, 2017). Essa diferente hierarquia de pistas acústicas parece estar ligada às funções que um determinado parâmetro possui na língua. Por exemplo, em línguas tonais, a f_0 tem o papel de veicular funções linguísticas que em línguas não tonais são veiculadas por outros parâmetros. Nessas línguas, diferenças de f_0 realizam diferenças entre tons que servem para contrastar itens lexicais. Por também ter essa função, o peso da f_0 é afetado quando esse parâmetro é usado para marcar fronteira, sendo parâmetros duracionais e reset de f_0 mais relevantes para assinalar fronteiras (YANG; WANG, 2002). É provável que isso aconteça também com os outros parâmetros, que se comportariam de maneira diferente para assinalar fronteira prosódica a depender da importância que têm para sinalizar outras funções em uma determinada língua. Muito pouco sabemos também como varia o peso de um parâmetro dentro de uma combinação mais ampla ao marcar fronteiras de unidades funcionalmente diferentes.

Enquanto alguns estudos se concentram em investigar a oposição entre presença vs. ausência de fronteira (MO *et al.*; 2008; BARBOSA, 2010), outros investigam uma potencial diversidade entre as fronteiras. Nesse último caso, alguns autores propõem a existência de um número determinado de fronteiras, enquanto outros propõem um continuum entre presença e ausência de fronteira. Nesse segundo caso corre-se o risco de encontrar sempre algum grau de fronteira, por mínimo que seja, e de perder a oposição de fronteira vs. não fronteira, tornando extremamente difícil, se não impossível, qualquer consideração de natureza funcional.

Quem, por outro lado, considera que as fronteiras são um fenômeno gradiente, mas categórico, propõe uma gradação de força das diferentes fronteiras, que contudo são em número limitado; entre esses autores existe uma discordância sobre a quantidade de fronteiras de força diferente que é possível reconhecer e perceber (cf. BARBOSA, 2006, para uma discussão). Alguns trabalhos distinguem simplesmente entre fronteiras fortes e fracas, enquanto outros consideram possível individualizar mais de dois graus de força (cf. WIGHTMAN *et al.*, 1992, para o inglês; BARBOSA, 2006, para o português brasileiro; BARBOSA, 1994, para o francês), alguns chegando até a sete, o que se alinha com as já citadas teorias fonológicas da prosódia como de Nespor e Vogel (1986) e Selkirk (1995).

Outra possibilidade de inferir graus de força é pelo uso de máximos locais dos parâmetros acústicos que veiculam a fronteira prosódica como índices da força dessa fronteira (TEIXEIRA FALCÃO, 2017). Mesmo que os valores de máximos locais variem continuamente, é possível o uso de técnicas estatísticas de agrupamento (*clusterization techniques*) para inferir um número limitado de forças de fronteira que não é superior a quatro (cf. BARBOSA, 2006, para o PB e BARBOSA, 1994, para o francês). No trabalho para o PB, Barbosa (2006) utilizou máximos de duração de unidades de tamanho silábico normalizada por *z-score* para a obtenção de 3 a 4 níveis distintos, parcialmente correlacionados com fronteiras sintáticas obtidas pela projeção de uma árvore de dependência nos modelos da de Tesnière (1965). Os diferentes graus de força permitem estabelecer uma hierarquia de constituintes prosódicos que abrem a possibilidade da inferência da estrutura prosódica de um enunciado. Esse procedimento já fora proposto por Grosjean e colegas (GROSJEAN; GROSJEAN; LANE, 1979; GROSJEAN; DOMMERGUES, 1983; GEE; GROSJEAN, 1983) a partir da leitura em

taxas cada vez mais lentas e da análise de durações de vogais acrescidas das eventuais pausas silenciosas à sua direita e de índices de segmentação dos enunciados obtidos por testes com ouvintes. Esse procedimento pôde revelar assim a “*structure de performance*” (estrutura de performance), uma estrutura prosódica com as seguintes propriedades: constituintes de tamanho semelhante, organização hierárquica e estrutura simétrica (GROSJEAN; DOMMERGUES, 1983). Essas propriedades emergiram de duas restrições concorrentes: a tendência do locutor em respeitar a estrutura linguística da sentença e a tendência a equilibrar a extensão dos constituintes que produz (MONNIN; GROSJEAN, 1993, p. 28; MARTIN, 1987). A tendência de equilíbrio da extensão de constituintes prosódicos explicaria porque os sujeitos não agrupam sistematicamente o verbo com o sintagma nominal objeto ao pronunciarem frases do inglês, como seria previsto pela sintaxe, mas preferem agrupamentos do tipo (SV)O (GROSJEAN; GROSJEAN; LANE, 1979, p. 59).

A discussão sobre os tipos de fronteira, no entanto, não é somente de cunho quantitativo. Muitos autores distinguem entre fronteiras que veiculam percepção de conclusão prosódica e linguística (com interpretações diferentes sobre a natureza da unidade linguística concluída) e fronteiras que veiculam a percepção de continuidade discursiva, sinalizando que o segmento de discurso em curso não pode se considerar concluído, apesar de a fronteira marcar a conclusão de um constituinte, esse também de diferente natureza dependendo da abordagem teórica (MONEGLIA; CRESTI 1997; CRYSTAL, 1969; SWERTS, 1994; SWERTS *et al.*, 1994). Em vários autores esses dois tipos de fronteiras são chamados respectivamente de terminais e não terminais.

Mas alguns dos autores que consideram a distinção entre fronteiras terminais e não terminais defendem que existe também uma diferença interna. Para esses autores, não haveria um único tipo de fronteira terminal e um único tipo de fronteira não terminal. Segundo essa proposta, podemos observar terminais “mais terminais” que as outras. Por exemplo, as fronteiras de enunciados seriam menos terminais se comparados à fronteira entre blocos discursivos maiores, chamados de parágrafos por alguns (van DONZEL, 1999). Analogamente, existiriam diversos tipos de fronteiras não terminais, algumas mais salientes que outras, ou perceptualmente mais próximas das terminais, ou que anunciam o fato que a conclusão está próxima; essas propostas não devem ser consideradas excludentes mas podem capturar diferentes aspectos

da complexidade do fenômeno (SWERTS *et al.*, 1994; TEIXEIRA FALCÃO, 2017).

De fato, se examinarmos os parâmetros fonético-acústicos correlacionados à percepção de fronteira, em particular de fronteira não terminal, observamos combinações muito variadas dentro da mesma língua e do mesmo texto (cf. TEIXEIRA FALCÃO, 2017). Temos, por exemplo, fronteiras marcadas claramente por um movimento de f_0 ascendente, uma pista acústica de continuidade, que, frequentemente junto com outras marcas prosódicas como a duração, veicula a clara percepção de que o discurso vai continuar. Por outro lado, esse movimento ascendente de f_0 ou o alongamento final podem faltar em outras fronteiras que também são percebidas como não terminais (cf. WAGNER, 2010).

Quanto às fronteiras conclusivas, frequentemente se observa que elas são caracterizadas por um movimento descendente da curva de f_0 até o nível mais baixo, e seguidas por um reset da f_0 no começo da unidade seguinte, que iniciaria com uma f_0 em uma altura claramente distinta. Contudo é comumente reconhecido que nem todos os enunciados se concluem com uma f_0 baixa. Embora o caso mais evidente e estudado seja aquele das interrogativas polares em línguas como inglês e espanhol peninsular, há outras ilocuções, segundo a terminologia e categorização que adotamos, que são marcadas, entre outros parâmetros por uma f_0 final mais alta (CRESTI, 2000 e no prelo; MORAES; RILLIARD, 2014, *inter alia*).

A variabilidade na manifestação física das fronteiras pode estar correlacionada com diferentes valores funcionais no plano linguístico. Teríamos então não somente uma correlação entre tipos de fronteiras que veiculam conclusão e tipos de fronteiras que veiculam continuação, mas também entre tipos conclusivos diferentes, por exemplo de ilocuções diferentes, e entre tipos não conclusivos diferentes, que, por hipótese, marcariam tipos de constituintes diferentes (sintáticos ou de outra natureza). Nesse caso a manifestação específica de uma fronteira prosódica não teria apenas um valor demarcativo, mas dependeria fortemente da função linguística da unidade da qual marcam a fronteira e, portanto, pistas que assinalam também essas mesmas funções linguísticas.

Olhando por esse lado, estudar a forma como fisicamente as fronteiras são realizadas e estudar a natureza das unidades demarcadas por duas fronteiras (uma à esquerda e outra à direita) não seriam mais de âmbitos distintos, o primeiro de interesse prioritário da fonética e o segundo de níveis linguísticos superiores ou dos estudiosos dos mecanismos

cognitivos, mas se tornariam bem mais integrados. Essa perspectiva que une as funções das unidades à manifestação concreta das fronteiras que as delimitam é ainda incipiente e poderá nos dar respostas interessantes sobre a natureza das unidades que são delimitadas por fronteiras.

Antes de passar às diferentes abordagens teóricas sobre as unidades, vale a pena fazer uma observação sobre alguns tipos de fronteiras (e de unidades) que estão bem menos frequentes na fala de laboratório, ao menos quando essa se restringe à fala lida, mas que são extremamente comuns na fala espontânea: os diferentes tipos de disfluências. Na fala espontânea são muito frequentes os fenômenos de interrupção, de retratação e hesitação. Muitas unidades terminam não porque o falante planejou a sua conclusão, mas porque alguma causa imprevista de natureza interna (não recuperação de palavra apropriada, mudança de idéia, ou qualquer problema na articulação ou na elaboração do conteúdo) ou externa (interrupção por outro falante ou qualquer evento ambiental) leva à interrupção momentânea do enunciado antes que ele seja completado semântica e prosodicamente. Quanto à retratação, o enunciado não é interrompido, mas é fragmentado por causa de repetições de palavras ou partes de palavras, que depois o falante idealmente cancela e corrige, prosseguindo no enunciado como se elas não tivessem sido pronunciadas. Trata-se nesse caso do resultado de dificuldades na realização do enunciado que não levam à interrupção do mesmo e que são mais ou menos presentes em todos os falantes, mas principalmente naqueles que tem menor domínio da fala, seja porque são muito jovens, seja porque são de diastratia baixa, seja por outras razões. No caso da hesitação, as dificuldades na fala se manifestam de diferentes formas como alongamentos vocálicos ou tomadas de tempo, também chamadas de pausas preenchidas. Sempre ou quase sempre que se dão um desses três fenômenos se geram também uma ou duas fronteiras (geralmente uma no caso da interrupção e duas nos outros dois casos). Contudo, essas fronteiras em princípio não são planejadas pelo falante e não marcam unidades com função linguística. Na análise das pistas de fronteira prosódica elas constituem um elemento de ruído, não podendo ser comparadas às fronteiras que o falante faz para construir o significado do enunciado.

Um último tipo de fronteira que deve ser considerado é o que delimita aquelas que, no modelo da *Language into Act Theory* (L-Act; CRESTI, 2000; MONEGLIA; RASO, 2014; MONEGLIA; CRESTI,

1997), são chamadas de *Scanning Units*. Uma *Scanning Unit*, segundo a visão da L-AcT é uma unidade informacionalmente não autônoma e que constitui uma parte de uma unidade informacional maior (por ex. um Tópico dividido em duas ou mais unidades entonacionais. Nesse caso, as unidades antes da última seriam *Scanning Units*, enquanto o perfil prosódico que marca a função da unidade informacional é sempre colocado na última unidade entonacional). Segundo a L-AcT, as fronteiras que delimitam essas unidades são devidas a diferentes possíveis causas: ênfase (para realçar as partes que compõem o texto de uma unidade informacional se segmenta seu conteúdo em mais unidades entonacionais), imperícia do falante (como se fossem pequenas hesitações ou retratações sem acréscimo de material segmental), necessidade articulatória (quando uma unidade informacional possui uma quantidade de sílabas superior àquela que cabe confortavelmente dentro de uma unidade entonacional). Essas fronteiras, que, como vimos, constituem um grupo não homogêneo, têm uma tipologização complexa em relação às outras fronteiras, pois só é possível individualizar uma *Scanning Unit* depois de uma etiquetagem informacional, o que segue a segmentação e não pode ser automatizado.

Além de todas essas questões abertas, seria interessante ainda considerar diversas outras questões de ordem não linguística: as vozes masculinas e as femininas usam os parâmetros acústicos para veicular a percepção de fronteira da mesma maneira? O que acontece nas várias patologias de fala, quando são comprometidas funções articulatórias ou cognitivas? E como evolui a capacidade de gerenciar os parâmetros para esse objetivo funcional ao longo da ontogênese?

A pesquisa das últimas décadas avançou muito na compreensão e na investigação das complexas combinações de fatores que condicionam a manifestação de fronteiras, e mais recentemente os trabalhos estão começando a investigar o fenômeno na fala espontânea. Contudo, ainda resta um longo caminho a ser percorrido. Enfim, enfrentar a questão das combinações de parâmetros não é suficiente, é necessário também olhar para o valor de cada parâmetro nas diversas combinações e para o peso relativo (hierarquia) desses parâmetros dentro de cada combinação. É evidente que isso aumenta muito as variáveis responsáveis para a marcação de fronteira prosódica, o que de fato nos impõe o uso de instrumentos computacionais e estatísticos para apreendê-las em algum grau satisfatório.

Mais recentemente, as fronteiras prosódicas têm sido objeto de investigação da psicolinguística no que tange questões de processamento (DRURY *et al.*, 2016; GLUSHKO *et al.*, 2016; NICKELS *et al.*, 2013; HWANG; STEINHAUER, 2011; PAUKER *et al.* 2011; STEINHAUSER, 2003; STEINHAUER; FRIEDERICI, 2001), de modo particular através da técnica do *Event-Related Potential* (ERP). Foi Steinhauer *et al.* (1999) e colaboradores que inicialmente usaram a técnica de ERP para mostrar que fronteiras prosódicas ouvidas estão associadas a trechos de aumento da amplitude da atividade elétrica (potencial evocado), que foi nomeado de CPS (*Closure Positive Shift*). Esse pico ocorre de 400 a 800 ms. após um momento definido, que, nos testes mais bem sucedidos, foi considerado a partir da última tônica antes de fronteira. Os experimentos foram realizados com e sem a presença de pausa e de diversos outros parâmetros considerados responsáveis por veicular percepção de fronteira, mas o pico de atividade se manteve sempre. Parece que a presença do alongamento silábico e de um tom de fronteira são suficientes para que o encéfalo do ouvinte reaja. As pesquisas atuais procuram refinar cada vez mais a observação de como reagimos a parâmetros isolados ou às suas combinações quanto à percepção de fronteira.

É especialmente interessante o fato de que a segmentação (*phrasing*) seria sensível a pistas de modalidades diferentes: não somente pistas acústicas, mas também pistas gráficas, como a vírgula na leitura, causariam um aumento de atividade elétrica em correspondência de fronteira. Além disso, o fenômeno aparece também para a segmentação musical, mas com uma latência maior (talvez devida à ausência de informações de natureza linguística como a sintaxe e o léxico). Parece também que o CPS é encontrado somente a partir de uma certa idade (cerca de três anos de idade), o que faz com que ele seja considerado dependente de uma capacidade de estruturação mínima, seja de natureza sintática seja mesmo de natureza prosódica *stricto sensu*. Este resultado parece compatível com os dados de estudos aquisicionais (THORNTON, 2016; HYAMS; ORFITELLI, 2015, *inter alia*). Por fim, o CPS parece ser maior quanto menos esperada for a fronteira, ou seja, quanto menos ela seja previsível com base em informações de outra natureza; mas também parece bastante claro que a prosódia, como veículo de fronteira, seria capaz de prevalecer em caso de conflito com as expectativas de natureza sintática (BÖGELS, TORREIRA, 2015; BÖGELS *et al.*, 2013, 2010).

Como a fronteira é marcada pelo concurso de todos os parâmetros prosódicos, especialmente, duração silábica, f0 e intensidade, é importante ainda apontar que há uma predominância, em indivíduos destros, de processamento temporal no hemisfério esquerdo enquanto o processamento espectral ativa majoritariamente áreas do hemisfério direito (ROBIN *et al.*, 1990; ZATORRE, 1997), o que também é confirmado por estudos em indivíduos lesionados seja no hemisfério esquerdo, seja no hemisfério direito, com os primeiros perdendo capacidade de processamento temporal (SHAH *et al.*, 2006). Quanto às áreas neuronais envolvidas na percepção da fala, tanto as áreas corticais temporais quanto parietais são ativadas bilateralmente (HICKOK; POEPEL, 2000).

3. Segmentação e significado linguístico

A segmentação da fala é de fundamental importância para a construção do significado linguístico (cf. FERY, 2017, para uma revisão). A prosódia é utilizada para orientar o ouvinte na reconstrução de unidades funcionais distintas e de sua hierarquia e função na decodificação da mensagem. Essa é a razão principal que motiva os pesquisadores a estudarem a natureza física das fronteiras e a sua relação com os diferentes níveis linguísticos. Olhemos apenas alguns exemplos em línguas diferentes. Em inglês, uma sequência como *People give John the book I promised him* pode ser segmentada pelo menos das quatro maneiras seguintes, gerando significados muito diferentes entre eles, tanto do ponto de vista ilocucionário quanto sintático:

- (a) *People* (Calling)! *Give John the book I promised him* (Order)!
- (b) *People give John the book I promised him* (Assertion).
- (c) *People give John the book* (Question)? *I promised him* (Assertion).
- (d) *People* (Calling)! *Give John the book* (Order)! *I promised him* (Assertion).

Em (a), (c) e (d) teríamos duas fronteiras terminais, enquanto em (b) teríamos apenas uma fronteira, essa também de natureza terminal. Quanto aos parâmetros acústicos, no entanto, as fronteiras terminais das várias segmentações são diferentes pelo menos quanto ao movimento de

f0. Se a segunda fronteira de (a), (c) e (d) é precedida de um movimento descendente, sua primeira fronteira apresenta um movimento ascendente. Esses movimentos ascendentes não são iguais, assim como não são iguais os movimentos descendentes dos outros casos. Uma diferenciação análoga poderia ser feita para os valores de intensidade e duração.

Em português uma sequência como *João vai pro Rio até amanhã* (*João will go (or go) to Rio until tomorrow (or see you tomorrow)*) pode ser segmentada pelo menos das três maneiras seguintes:

- (a) *João* (calling)! *Vai pro Rio até amanhã* (order)! (*João! Go to Rio until tomorrow*)
- (b) *João vai pro Rio até amanhã* (assertion). (*João will go to Rio until tomorrow*)
- (c) *João* (calling)! *Vai pro Rio* (order)! *Até amanhã* (greeting)! (*João! Go to Rio! See you tomorrow*)

Nessas três organizações de sentenças, é evidente que a segmentação afeta a interpretação sintática e semântico-pragmática da sequência.

Por fim, o exemplo seguinte mostra como a segmentação pode ser decisiva para a interpretação sintática e semântica em outra língua ainda, nesse caso o italiano:

- (a) *Claudia* (calling)! *Guarda* (deixis)! *Quanto è bello* (expressive)! (*Claudia! Look! How beautiful it is!*)
- (b) *Claudia* (calling)! *Guarda quanto è bello* (deixis)! (*Claudia! Look how beautiful it is!*)
- (c) *Claudia guarda quanto è bello* (assertion). (*Claudia looks how beautiful it is.*)

A exemplificação poderia ser mais complexa, levando em conta outras interpretações e diferentes tipos de unidades) e estendida a outras línguas, mas o que importa aqui evidenciar é a importância do papel da segmentação na construção do significado linguístico tanto no nível sintático quanto naquele semântico. A presença de fronteira afeta certamente também o nível morfo-fonológico, por exemplo, inibindo fenômenos de sândi.

Nos exemplos anteriores vimos somente casos de fronteiras terminais, que isolam sequências autônomas pragmática e prosodicamente e que podem ser enunciadas em isolamento. Mas o significado é afetado também em caso de fronteiras não terminais, ou seja, quando a relação sintática ou informacional entre as unidades separadas pela fronteira deve ser mantida. Por exemplo a sequência *the film I like it* é analisável como um sintagma nominal modificado por uma relativa. Se, por outro lado, inserirmos uma fronteira a análise pode mudar: *in the film, I like it* a análise pode ser aquela de Tópico e Comentário e interpretável como: *as for the film, I like it*.

Voltemos agora à noção de unidade de referência da fala, entendida como unidade com sentido comunicativo autônomo que compõe o texto. Se consideramos a dimensão prosódica, é difícil definir essa unidade com base exclusivamente em critérios sintáticos que definem as tradicionais categorias de oração e sentença. A prosódia tem uma dimensão comunicativa que conduz preferencialmente os pesquisadores a prestar atenção à produção e à percepção da fala, mesmo não faltando perspectivas mais abstratas (mas fora de um contexto comunicativo). Muitos linguistas que incorporam a prosódia como elemento primário de seus modelos consideram a percepção prosódica de conclusão de uma sequência comunicativa como a marca principal da unidade de referência (CRESTI, 2000; MONEGLIA; RASO, 2014; IZRE'EL, 2002). Outros preferem considerar como unidade de referência a unidade entonacional, independentemente de ela apresentar um contorno prosódico percebido como conclusivo ou continuativo (METTOUCHI *et al.*, 2010). Em ambas as perspectivas a marca principal da unidade de referência está na fronteira entonacional. A diferença reside na questão de se qualquer fronteira pode marcar o fim de uma unidade de referência ou se somente fronteiras com uma qualidade específica fazem isso. Essa discussão se acompanha também daquela relativa às relações linguísticas que se manifestam dentro da unidade entonacional, dentro de um conjunto de unidades entonacionais marcado por uma fronteira conclusiva, e também nas relações que atravessam a fronteira conclusiva e precisam de unidades ainda maiores (para alguns aspectos dessa discussão dentro de arcabouços teóricos diferentes embora próximos, veja-se Izre'el neste número; CRESTI, 2014; PIETRANDREA *et al.*, 2014).

4. As contribuições presentes no volume e as contribuições para o debate

Os nove trabalhos apresentados neste número temático tocam em aspectos diferentes da segmentação prosódica da fala espontânea. Um primeiro grupo de contribuições se concentra na elaboração de softwares que permitam a extração de dados e informações capazes de esclarecer algumas das tantas questões ligadas à segmentação. Naturalmente, também por trás desses trabalhos há sempre uma hipótese teórica, seja na função, seja na quantidade de fronteiras a serem identificadas.

O trabalho de Xu e Gao apresenta a ferramenta computacional FormantPro que usa o *software* Praat como plataforma para a extração automática de trajetórias de formantes. Embora o tema não enfoque diretamente a questão da segmentação prosódica, tanto a ferramenta quando os exemplos levantados pelos autores abrem uma discussão sobre isomorfismo entre eventos acústicos e articulatórios que marcam fronteiras de consoantes e vogais. Essas fronteiras são discutidas a partir de um alinhamento com trajetórias de f_0 que podem vir a ter implicações para a delimitação de fronteiras prosódicas. O programa gera subsidiariamente valores de duração e de intensidade e permite a apresentação das trajetórias médias aliadas a uma normalização temporal que auxilia a observar as equivalências entre instâncias de diferentes enunciados contendo palavras em contraste. Os valores de duração podem ser usados para investigar pistas de fronteiras prosódicas no caso de mudanças importantes em relação ao seu entorno.

O trabalho de Teixeira Falcão e Mittmann apresenta um interessante procedimento para extrair modelos de parâmetros acústicos para diferentes tipos de fronteira em trechos de *corpora* de fala espontânea previamente segmentados por 14 segmentadores. Dos dados de *corpora*, após deles serem tratados para que o *script* pudesse lê-los em Praat, é extraído um número muito grande de medidas em uma janela de 10 unidades V-V à esquerda e à direita de cada posição candidata a fronteira, ou seja, de cada fronteira de palavra fonológica. A segmentação em unidade V-V (BARBOSA, 2006) mostra como outros níveis de segmentação da fala interagem necessariamente com o nível da unidade entonacional. Um procedimento estatístico e o procedimento humano de refinamento revelam as combinações de parâmetros que melhor explicariam as fronteiras, e seus pesos. Todo o trabalho foi

planejado considerando que as fronteiras prosódicas podem ser divididas em dois grandes grupos: terminais e não terminais. O trabalho relativo às quebras não terminais aponta para a necessidade de considerar essas fronteiras em pelo menos três sub-grupos distintos, cada um explicado por um modelo distinto. Esse resultado alimenta as reflexões tanto sobre a existência de uma distinção entre fronteiras terminais e não terminais e também de distinções mais sutis. Seria importante investigar ao que seriam devidas essas últimas.

O trabalho de Bigi e Meunier avalia a ferramenta de *software* SPPAS que permite a segmentação automática da fala lida e da espontânea focando especialmente, no último caso, nas questões relativas às disfluências. A ferramenta apresentada pressupõe a existência de uma transcrição ortográfica e de um dicionário de pronúncia de palavras de um léxico. Além disso, contém um modelo acústico dos sons da fala do francês que permite o alinhamento de símbolos fonéticos com o sinal de fala. Os erros de alinhamento são cerca de 11% na fala lida e 15 % na fala espontânea, mas podem ser reduzidos a partir de uma transcrição ortográfica enriquecida que identifique os tipos de disfluência. A ferramenta é testada com nove *corpora* que incluem fala lida, conversa espontânea e debate político para os casos de trechos disfluents contendo risos, pausas preenchidas e ruídos. Os autores demonstram que, sendo precedido de um pré-processamento que separa a cadeia de fala em unidades entre pausas, pode-se atingir um nível de precisão na delimitação dessas unidades de apenas 20 ms.

O artigo de G. Christodoulides usa dois *corpora* de fala do francês com anotação de fronteiras de força diferente para verificar: (a) o grau de acordo entre anotações prosódicas a partir de duas abordagens teóricas diferentes, a teoria métrica autosegmental (PIERREHMBERT, 1980) e a distinção entre micro e macro-sintaxe (BLANCHE-BENVENISTE, 2002, 2003) quanto a dois níveis de anotação comparáveis; (b) quais parâmetros acústicos são mais importantes para veicular os dois tipos de fronteira e qual a sua hierarquia. O uso de *corpora* dependentes de abordagens teóricas tão diferentes é um teste importante para o tema das fronteiras, ainda mais considerando que um *corpus* é segmentado com base em critérios teóricos e o outro com base em critérios perceptuais. Os parâmetros investigados são a presença e a duração de pausa, o alongamento pré-fronteiriço e duas medidas de f_0 associadas com a fronteira. A análise mostra um acordo muito alto entre os dois *corpora*

quanto aos parâmetros prosódicos envolvidos nas posições onde foi marcada a fronteira e na distinção dos dois tipos de fronteiras comparados. A conclusão é que o parâmetro mais importante associado com a presença de fronteira e com a sua maior força seria a pausa, seguida pelo alongamento silábico. A f_0 seria importante apenas para marcar a diferença entre presença ou ausência de fronteira, mas não para distinguir as forças dos dois tipos de fronteira.

O trabalho de Ph. Martin se distingue dos outros por analisar uma unidade diferente: o grupo acentual. O objeto de estudo é, portanto, uma unidade menor que a unidade entonacional, mesmo se às vezes pode coincidir com ela. A observação dessa unidade é importante para os nossos objetivos, já que Martin individualiza um número restrito de possíveis movimentos de f_0 dentro de um grupo acentual e um número restrito de sequências de movimentos dentro da unidade entonacional, com um critério de dependência entre eles. Isso nos permite investigar a estrutura interna de uma unidade entonacional com base em unidades marcadas pelo acento. Entre outras consequências, os resultados dessa análise poderão trazer mais luz sobre as características das diversas estruturas internas das unidades entonacionais e sobre quanto e como essas estruturas se correlacionam com a função linguística veiculada pela unidade. Diferentes aspectos da unidade, com a presença de certas proeminências em certas posições, já são discutidos na literatura, mesmo se ainda não suficientemente na perspectiva adotada aqui. Propostas como a de Martin nos levam a considerar o papel de um outro nível prosódico e a sua específica função linguística, além de outras características (proeminências, tipo de fronteira) que possam nos fazer entender melhor como construímos uma sequência que possui uma função linguística gerenciando diferentes níveis da estrutura prosódica.

Um terceiro grupo de artigos investiga as fronteiras com objetivos diretamente ligados a algum nível linguístico, seja sintático, seja informacional.

O trabalho de A. Mettouchi, realizado sobre o Kabyle, língua afro-asiática da Argélia, mostra como a presença/ausência de fronteira pode constituir uma marca de uma função de natureza sintática, nesse caso o objeto direto. A fronteira se revela como o traço formal decisivo para distinguir essa estrutura de outras estruturas que possuem funções diferentes, talvez de natureza especificamente informacional, mas que aparecem no enunciado com os mesmos traços formais do objeto, a não

ser pela presença da fronteira que é ao contrário ausente na realização com função de objeto direto. Esse trabalho levanta uma questão importante: a relação entre presença de fronteira e ruptura da composicionalidade sintática. Outros estudos (CRESTI, 2014; RASO; VIEIRA, 2016) tratam dessa importante questão, que ainda é controversa. Se por um lado é fácil encontrar casos em que parece que a composicionalidade sintática se interrompe em coincidência da fronteira (e seria, portanto, possível hipotetizar que a fronteira ou algum tipo de fronteira tenha um papel em marcar essa interrupção), por outro lado outros casos são interpretáveis também salvando a composicionalidade em presença de fronteira.

O trabalho de da Silva e Fonseca apresenta também diversos motivos de interesse. Um primeiro motivo, como no caso do artigo anterior e do seguinte, é a importância que uma marca prosódica adquire para a identificação de uma unidade linguística, no caso em questão o Tópico. Um segundo motivo é a natureza experimental do trabalho, sobre a qual diremos algo mais para frente. Um terceiro motivo é que mostra como os resultados apresentados dentro de uma teoria formalista podem ser úteis também para diferentes visões da categoria de Tópico, evidenciando como o estudo empírico dos dados beneficia o debate. Os experimentos idealizados e realizados por da Silva e Fonseca podem ser muito interessantes para o debate entre os estudiosos da estruturação informacional da fala. Seus resultados podem ser utilizados para comparar a definição sintática de Tópico com as definições mais pragmáticas, em particular com a definição proposta por L-AcT, que dá um peso muito grande aos aspectos prosódicos, e que apresenta vários resultados de pesquisas em diversas línguas, entre as quais o PB (cf. CRESTI, 2000; SIGNORINI, 2004; FIRENZUOLI; SIGNORINI, 2003; MONEGLIA; RASO 2014; ROCHA; RASO, 2013; CAVALCANTE, 2016; MITTMANN, 2012; RASO; CAVALCANTE; MITTMANN, no prelo). De fato, os resultados não esperados do terceiro experimento poderiam ser explicados assumindo que o Tópico seja uma categoria pragmática que não depende da estrutura argumental e, portanto, pode estar em posição de sujeito, mas ser marcada por fronteira e por um foco prosódico funcional que o distingue claramente do sujeito, que por sua vez não apresenta fronteira prosódica com o resto do enunciado e não possui foco prosódico funcional. Nesse caso, a diferença entre sujeito e Tópico não seria de natureza sintática, mas seria devida à diferença entre um item sintático interno ao comentário (o sujeito) e um item pragmático

externo ao comentário (o Tópico). Um diálogo mais aprofundado entre essas diferentes abordagens teóricas poderia levar a uma maior clareza sobre o conceito de Tópico e estimular ambas as teorias a refinar a própria análise e as próprias argumentações, utilizando tanto abordagens experimentais como aquelas propostas por Da Silva e Fonseca quanto os dados de *corpora* de fala espontânea elaborados dentro de L-AcT.

Também o artigo de Panunzi e Saccone é fortemente orientado teoricamente. De fato, seu objetivo é observar se, quanto e como as fronteiras entre pares de unidades informacionais se realizam com características diferentes entre elas. Os dois pares (ou raramente sequências de mais de duas unidades) que são explorados no artigo são diferentes combinações de unidades ilocucionárias. Um tipo de par é caracterizado por compor uma interpretação ilocucionária complexa mas única, prosódica e pragmaticamente padronizada. O outro tipo de par, ao contrário, é constituído por duas ilocuções independentes, apesar de serem separadas por fronteira não terminal. Portanto, para analisar as fronteiras, é necessário que o texto já esteja anotado informacionalmente segundo um determinado arcabouço teórico, neste caso o da L-AcT (CRESTI, 2000; MONEGLIA; RASO, 2014). Os primeiros resultados parecem apontar para diferenças fortes na forma das fronteiras entre os dois pares de unidades. Este é um exemplo instigante de como as características das fronteiras podem correlacionar com a função das unidades que elas separam. Esse tipo de trabalho, que tenta correlacionar função da unidade com as características das fronteiras pode ser aplicado a diferentes tipos de unidades e com base em diferentes quadros teóricos.

O trabalho de Izre'el foi colocado no final deste número porque, a partir da consideração dos aspectos prosódicos e em particular das fronteiras, propõe uma revisão geral das categorias tradicionais de frase, oração, sentença e predicação, mostrando como a incorporação dos traços prosódicos pode levar a uma reformulação geral das categorias canônicas no estudo da fala espontânea. Izre'el retoma a discussão linguística sobre essas categorias desde a linguística clássica até Chomsky, para mostrar como certas categorias, assim como definidas na tradição de imposição sintaticista, não funcionam na análise da fala e principalmente da fala espontânea que, em princípio, deve ser o domínio natural para análise da linguagem. Levando em conta a prosódia e dados de *corpora* de fala espontânea, emerge claramente a importância da categoria de ilocução (que Izre'el chama de *modalidade*) enquanto categoria decisiva para a

individualização de uma unidade comunicativa e enquanto categoria marcada diretamente pela prosódia. E ainda emerge claramente a importância das fronteiras prosódicas para definir o domínio em que se dão as relações linguísticas em sua realização comunicativa. Como outros trabalhos deste número, mas com um escopo maior, a contribuição de Izre'el traz mais argumentos (cf. também BIBER *et al.*, 1999; as contribuições em RASO; MELLO, 2014; CRESTI, 2005; RASO; MITTMANN, 2012, *inter alia*) que apontam para a necessidade de definir a unidade comunicativa da fala revendo a noção de predicação (e de proposição), ou as noções de oração e de sentença, e sustentando a necessidade de incorporar a prosódia como elemento central na marcação dessa unidade de referência comunicativa. Como diversas contribuições deste número, o artigo de Izre'el não deixa dúvida sobre a necessidade de incorporar a prosódia entre os níveis de análise da linguística, e, mais do que isso, sobre o peso hierarquicamente decisivo da prosódia na individualização dos constituintes linguísticos da fala.

Referências

- AMIR, N.; SILBER-VAROD, V.; IZRE'EL, S. Characteristics of Intonation Unit Boundaries in Spontaneous Spoken Hebrew: Perception and Acoustic Correlates. In: SPEECH PROSODY INTERNATIONAL CONFERENCE, 2004. Nara. *Proceedings...* Nara: ISCA, 2004. p. 677-680.
- AURAN, C.; BOUZON, C.; HIRST, D. The Aix-MARSEC Project: an Evolutive Database of Spoken British English. In: SPEECH PROSODY INTERNATIONAL CONFERENCE, 2004. Nara, Japan. *Proceedings...* Nara: ISCA, 2004.
- AVANZI, M.; LACHERET-DUJOUR, A.; VICTORRI, B. ANALOR. Tool for Semi-Automatic Annotation of French Prosodic Structure. In: ANALOR. A Tool for Semi-Automatic Annotation of French Prosodic Structure. Campinas, Brazil, May 2008. p. 119-122.
- BARBOSA, P. A. *Caractérisation et génération automatique de la structuration rythmique du français*. 1994. Tese (Doutorado) – Institut National Polytechnique de Grenoble, França, 1994.
- BARBOSA, P. A. *Incursões em torno do ritmo da fala*. Campinas: Pontes, 2006.

BARBOSA, P. A. Prominence-and Boundary-Related Acoustic Correlations in Brazilian Portuguese Read and Spontaneous Speech. In: SPEECH PROSODY INTERNATIONAL CONFERENCE, 4., 2008, Campinas. *Proceedings...* Campinas: ISCA, 2008. p. 257-260.

BARBOSA, P. A. Automatic Duration-Related Saliency Detection in Brazilian Portuguese Read and Spontaneous Speech. In: SPEECH PROSODY INTERNATIONAL CONFERENCE, 5., 2010, Chicago. *Proceedings...* Chicago: ISCA, 2010.

BARBOSA, P. A. Panorama of Experimental Prosody Research. In: GSCP INTERNATIONAL CONFERENCE – SPEECH AND CORPORA, VIIth., Belo Horizonte. *Proceedings...* Florence: Firenze University Press, 2012. p. 33-42.

BARTH-WEINGARTEN, D. *Intonation Units Revisited*. Cesura in Talk-In-Interaction. Amsterdam: John Benjamins, 2016.

BIBER, D.; JOHANSSON, S.; LEECH, G.; CONRAD, S.; FINEGAN, E. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited, 1999.

BLANCHE-BENVENISTE, C. Macro-syntaxe et micro-syntaxe: les dispositifs de la rection verbale. In: ANDERSEN, H. L.; NØLKE, H. (Éd.). *Macro-Syntaxe e Macro-Sémantique*. Bern: Peter Lang, 2002. p. 95-115.

BLANCHE-BENVENISTE, C. Le recouvrement de la syntaxe et de la macro-syntaxe. In: SCARANO, A. (Ed.). *Macro-syntaxe et pragmatique*. L'analyse linguistique de l'oral. Roma: Bulzoni, 2003. p. 53-75.

BLANCHE-BENVENISTE, C.; JEANJEAN, C. *Le français parlé*. Transcription et édition. Paris: Didier Érudition; Institut National de la Langue Française, 1987.

BLANCHE-BENVENISTE C. Macro-syntaxe et micro-syntaxe: les dispositifs de la rection verbale. In: ANDERSEN, H. L.; NØLKE, H. (Éd.). *Macro-syntaxe et macro-sémantique*. Actes du Colloque International d'Århus [17-19 mai 2001]. Bern: Peter Lang, 2002. p. 95-118.

BÖGELS, S.; SCHRIEFERS, H.; VONK, W.; CHWILLA, D. J.; KERKHOFS, R. The Interplay Between Prosody and Syntax in Sentence Processing: The Case of Subject- and Object-Control Verbs. *Journal of Cognitive Neuroscience*, Cambridge, v. 22, n. 5, p. 1036-1053, 2010.

- BÖGELS, S.; SCHRIEFERS, H.; VONK, W.; CHWILLA, D.; KERKHOFS, R. Processing Consequences of Superfluous and Missing Prosodic Breaks in Auditory Sentence Comprehension. *Neuropsychologia*, Oxford, v. 51, p. 2715-2728, 2013.
- BÖGELS, S.; TORREIRA, F. Listeners Use Intonational Phrase Boundaries to Project Turn Ends in Spoken Interaction. *Journal of Phonetics*, [s.l.], v. 52, p. 46-57, 2015.
- BOLINGER, D. Pitch Accent and Sentence Rhythm. In: ABE, I.; KANEKIYO, T. (Ed.). *Forms of English: Accent, Morpheme*, Order. Cambridge, Mass: Harvard University Press, 1965. p. 139-180.
- BOSSAGLIA, G.; MELLO, H.; RASO, T. Insubordination and the Syntax/Prosody Interface in Spoken Brazilian Portuguese: Data on Adverbial Clauses. In: IZRE'EL, S.; MELLO, H.; PANUNZI, A.; RASO, T. *In Search for a Reference Unit of Spoken Language: A Corpus Driven Approach*. Amsterdam: John Benjamins. (Forthcoming).
- BUHMANN, J.; CASPERS, J.; HEUVEN, V. J. van; HOEKSTRA, H.; MARTENS, J-P.; SWERTS, M. Annotation of Prominent Words, Prosodic Boundaries and Segmental Lengthening by Non-Expert Transcribers in the Spoken Dutch Corpus. In: LREC, 3rd, 2002, Las Palmas. *Proceedings...* Las Palmas: ELRA, 2002. p. 779-785.
- BYBEE, J. *Language, Usage and Cognition*. Cambridge: CUP, 2010.
- CARLSON, R.; HIRSCHBERG, J.; SWERTS, M. Cues to Upcoming Swedish Prosodic Boundaries: Subjective Judgment Studies and Acoustic Correlates. *Speech Communication*, [s.l.], v. 46, p. 326-333, 2005. Doi: <https://doi.org/10.1016/j.specom.2005.02.013>
- CAVALCANTE, F. *The Topic Unit in Spontaneous American English: a Corpus-Based Study*. 2016. Master (Thesis) – Universidade Federal de Minas Gerais, Belo Horizonte, 2016.
- CHAFE, W. *Discourse, Consciousness and Time. The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: Chicago University Press, 1994.
- COUPER-KUHLEN, E. Prosody and sequence organization in English conversation. In: COUPER-KUHLEN, E.; FORD, C. E. (Ed.). *Sound Patterns in Interaction: Cross-Linguistic Studies from Conversation*. Amsterdam: John Benjamins, 2004. p. 335-376.

COWAN, N. *Attention and Memory: An Integrated Framework*. Oxford: Oxford University Press, 1998. Doi: <https://doi.org/10.1093/acprof:oso/9780195119107.001.0001>

CRESTI, E. *Corpus di italiano parlato*. Firenze: Accademia della Crusca, 2000. 2 v.

CRESTI, E. Notes on Lexical Strategy, Structural Strategies and Surface Clause Indexes in the C-ORAL-ROM Spoken Corpora. In: CRESTI, E.; MONEGLIA, M. (Ed.). *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam; Philadelphia: John Benjamins, 2005. p. 209-256.

CRESTI, E. Syntactic Properties of Spontaneous Speech in the Language into Act Theory: Data on Italian Complements and Relative Clauses. In: RASO, T.; MELLO, H. (Ed.). *Spoken Corpora and Linguistic Studies*. Amsterdam: John Benjamins, 2014.

CRESTI, E. The Pragmatic Analysis of Speech and Its Illocutionary Classification According to Language into Act Theory. In: IZRE'EL, S.; MELLO, H.; PANUNZI, A.; RASO, T. *In Search for a Reference Unit of Spoken Language: A Corpus Driven Approach*. Amsterdam: John Benjamins. (Forthcoming).

CRESTI, E.; MONEGLIA, M. (Ed.). *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam: John Benjamins, 2005.

CROFT, W. Intonational Units and Grammatical Structure. *Linguistics*, [s.l.], v. 33, n. 5, p. 839-882, 1995.

CRUTTENDEN, Alan. *Intonation*. 2nd edition. New York: Cambridge University Press, 1997.

CRYSTAL, D. *Prosodic Systems and Intonation in English*. Cambridge: CUP, 1969.

DILLEY, L.; SHATTUCK-HUFNAGEL, S.; OSTENDORF, M. Glottalization of Word-Initial Vowel as a Function of Prosodic Structure. *Journal of Phonetics*, [s.l.], v. 24, p. 423- 444, 1996.

- DRURY, J. E.; BAUM, Sh. R.; VALERIOTE, H.; STEINHAUSER, K. Punctuation and Implicit Prosody in Silent Reading: An ERP Study Investigating English Garden-Path Sentences. *Frontiers in Psychology*, [s.l.], Sept. 2016. Doi: <https://doi.org/10.3389/fpsyg.2016.01375>
- DU BOIS, J. W.; CHAFE, W. L.; MEYER, Ch.; THOMPSON, S. *Santa Barbara Corpus of Spoken American English*. Washington DC: Linguistic Data Consortium, 2000-2005.
- DU BOIS, J. W.; CUMMING, S.; SCHUETZE-COBURN, S.; PAOLINO, D. Discourse transcription. *Santa Barbara Papers in Linguistics*, Santa Barbara, v. 4, n. 1, p. 225, 1992.
- EVANS, N.; WATANABE, H. The dynamics of insubordination: An overview. In: _____. (Ed.). *Insubordination*. Amsterdam: John Benjamins, 2016.
- FRY, C. *Intonation and Prosodic Structure*. Cambridge: CUP, 2017.
- FIRENZUOLI, V.; SIGNORINI, S. L'unità informativa di topic: correlati intonativi. In: MAROTTA, G. (Ed.). *La coarticolazione: atti delle XIII Giornate di Studio del Gruppo di Fonetica Sperimentale*, [28-30 nov. 2002]. Pisa: ETS, 2003. p. 177-184.
- FON, J.; JOHNSON, K.; CHEN, S. Durational Patterning at Syntactic and Discourse Boundaries in Mandarin Spontaneous Speech. *Language and Speech*, [s.l.], v. 54, n. 1, p. 5-32, 2011.
- FUCHS, S.; KRIVOKAPIC, J.; JANNEDY, S. Prosodic Boundaries in German: Final Lengthening in Spontaneous Speech. *The Journal of the Acoustical Society of America*, [s.l.], v. 127, n. 3, p. 1851, 2010. Doi: 10.1121/1.3384378
- FURHT, B.; VILLANUSTRE, F. *Big Data Technologies and Applications*. [s.l.]: Springer, 2016.
- GAROFALO, J. S.; LAMEL, L. F.; FISHER, W. M.; FISCUS, J. G.; PALLETT, D. S.; DAHLGREN, N. L.; ZUE, V. TIMIT Acoustic-Phonetic Continuous Speech Corpus LDC93S1. Web Download. Philadelphia: Linguistic Data Consortium, 1993.
- GEE, J. P.; GROSJEAN, F. Performance Structures: A Psycholinguistic and Linguistic Appraisal. *Cognitive Psychology*, [s.l.], v. 15, n. 4, p. 411-458, 1983.

- GLUSHKO, A.; STEINHAEUER, K.; De PRIEST, J.; KOELSCH, S. Neurophysiological Correlates of Musical and Prosodic Phrasing: Shared Processing Mechanisms and Effects of Musical Expertise. *PLoS ONE*, San Francisco, v. 11, n. 5, 2016.
- GORDON, M.; LADEFOGED, P. Phonation Types: a Cross-Linguistic Overview. *Journal of Phonetics*, [s.l.], v. 29, p. 383-406, 2001.
- GROSJEAN, F.; DOMMERGUES, J. Y. Les structures de performance en psycholinguistique. *L'Année Psychologique*, [s.l.], v. 83, n. 2, p. 513-536, 1983.
- GROSJEAN, F.; GROSJEAN, L.; LANE, H. The Patterns of Silence: Performance Structures in Sentence Production. *Cognitive Psychology*, [s.l.], v. 11, n. 1, p. 58-81, 1979.
- HANSON, H. M.; CHUANG, E. S. Glottal Characteristics of Male Speakers: Acoustic Correlates and Comparison with Female Data. *Journal of the Acoustical Society of America*, [s.l.], v. 106, n. 2, p. 1064-1077, 2001.
- t'HART, J. Differential Sensitivity to Pitch Distance, Particularly in Speech. *The Journal of the Acoustical Society of America*, [s.l.], v. 69, n. 3, p. 811-821, 1981.
- t'HART, J.; COLLIER, R.; COHEN, A. *A Perceptual Study on Intonation: An Experimental Approach to Speech Melody*. Cambridge: CUP, 1990. Doi: <https://doi.org/10.1017/CBO9780511627743>
- HICKOK, G.; POEPEL, D. Towards a Functional Neuroanatomy of Speech Perception. *Trends in Cognitive Sciences*, [s.l.], v. 4, n. 4, p. 131-138, 2000.
- HIRST, D.; Di CRISTO, A. A Survey of Intonation Systems. In: _____. (Ed.). *Intonation Systems: A Survey of Twenty Languages*. Cambridge: Cambridge University Press, 1998.
- HUGGINS, A. W. F. Just Noticeable Differences for Segment Duration in Natural Speech. *The Journal of the Acoustical Society of America*, [s.l.], v. 51, n. 4B, p. 1270-1278, 1972.

HWANG, H.; STEINHAEUER, K. Phrase Length Matters: The Interplay Between Implicit Prosody and Syntax in Korean ‘Garden Path’ Sentences. *Journal of Cognitive Neuroscience*, Cambridge, v. 23, n. 11, p. 3555-3575, 2011. Doi: https://doi.org/10.1162/jocn_a_00001

HYAMS, N.; ORFITELLI, R. The Acquisition of Syntax. In: CAIRNS, H.; FERNANDEZ, E. (Ed.). *Handbook of Psycholinguistics*. [s.l.]: Wiley; Blackwell Publishers, 2015.

IZRE’EL, S. The Corpus of Spoken Israeli Hebrew: Textual Samples. *Lesson Enu*, v. 64, p. 289-314, 2002.

IZRE’EL, S.; MELLO, H.; PANUNZI, A.; RASO, T. (Ed.). *In Search for a Reference Unit of Spoken Language: A Corpus Driven Approach*. Amsterdam: John Benjamins. (Forthcoming)

KLATT, D. H.; COOPER, W. E. Perception of Segment Duration in Sentence Contexts. In: COHEN, A.; NOOTEBOOM, S. G. (Ed.). *Structure and Process in Speech Perception*. Berlin; Heidelberg: Springer, 1975. p. 69-89. Doi: https://doi.org/10.1007/978-3-642-81000-8_5

KOFFI, E. A Just Noticeable Difference (JND) Reanalysis of Fry’s Original Acoustic Correlates of Stress in American English. *Linguistic Portfolios*, St. Cloud, v. 7, 2018.

LADD, D. R. *Intonational phonology*. Cambridge: Cambridge University Press, 1996.

LIEBERMAN, P. Some Acoustic Correlates of Word Stress in American English. *Journal of the Acoustical Society of America*, [s.l.], v. 32, p. 451-454, 1960.

LINELL, P. *The Written Language Bias in Linguistics: Its Nature, Origins and Transformations*. London; New York: Routledge, 2005. Doi: <https://doi.org/10.4324/9780203342763>

MacNEILAGE, P. F. The Frame/Content Theory of Evolution of Speech Production. *Behavioral and Brain Sciences*, Cambridge, v. 21, n. 4, p. 499-511, 1998.

MARTIN, P. Les problèmes de l’intonation: recherches et applications. *Langue Française*, [s.l.], v. 19, p. 4-32, 1973.

MARTIN, P. Prosodic and Rhythmic Structures in French. *Linguistics*, [s.l.], v. 25, n. 5, p. 925-950, 1987.

MELLO, H.; RASO, T.; MITTMANN, M. VALE, H.; CÔRTEZ, P. Transcrição e segmentação prosódica do *corpus* c-oral-brasil: critérios de implementação e validação. In: RASO, T.; MELLO, H. (Ed.). C-ORAL-BRASIL I. *Corpus* de referência do português brasileiro falado informal. Belo Horizonte: UFMG, 2012. p. 125-174.

METTOUCHI, A.; CAUBET, D.; VANHOVE, M.; TOSCO, M.; COMRIE, B.; IZRE'EL, S. CORPAFROAS, A Corpus for Spoken Afroasiatic Languages: Morphosyntactic and Prosodic Analysis. In: ITALIAN MEETING OF AFRO-ASIATIC LINGUISTICS, 13th, 2010, Padova. *Proceedings...* Padova: Sargon, 2010. p.177-180.

MITTMANN, M. M. *O C-ORAL-BRASIL e o estudo da fala informal: um novo olhar sobre o Tópico no Português Brasileiro*. 2012. Dissertation (Ph.D.) – Universidade Federal de Minas Gerais, Belo Horizonte, 2012.

MITTMANN, M. M.; BARBOSA, P. A. An Automatic Speech Segmentation Tool Based on Multiple Acoustic Parameters. *CHIMERA. Romance Corpora and Linguistic Studies*, Madrid, v. 3, p.133-147, 2016.

MO, Y. Duration and Intensity as Perceptual Cues for Naïve Listeners' Prominence and Boundary Perception. In: SPEECH PROSODY INTERNATIONAL CONFERENCE, 4th., 2008, Campinas. *Proceedings...* Campinas: ISCA, 2008. p. 739-742.

MO, Y.; COLE, J.; LEE, E. K. Naïve Listeners' Prominence and Boundary Perception. In: SPEECH PROSODY INTERNATIONAL CONFERENCE, 4th., 2008, Campinas. *Proceedings...* Campinas: ISCA, 2008. p. 735-738.

MO, Y.; COLE, J. Perception of Prosodic Boundaries in Spontaneous Speech with and Without Silent Pauses. *The Journal of the Acoustical Society of America* [s.l.], v. 127, n. 3, p. 1956, 2010.

MONEGLIA, M.; CRESTI, E. L'intonazione e i criteri di trascrizione del parlato adulto e infantile. In: BORTOLINI, U.; PIZZUTO, E. (Ed.). *Il Progetto CHILDES Italia*. Pisa: Del Cerro, 1997. p. 57-90.

MONEGLIA, M.; FABBRI, M.; QUAZZA, S.; ANDREA; PANIZZA, A.; DANIELI, M.; GARRIDO, J. M.; SWERTS, M. Evaluation of Consensus on the Annotation of Terminal and Non-Terminal Prosodic Breaks in the C-ORAL-ROM *corpus*. In: CRESTI, E.; MONEGLIA, M. (Ed.). *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance*

Languages. Amsterdam: John Benjamins, 2005. p. 257-276. Doi: <https://doi.org/10.1075/scl.15.09mon>

MONEGLIA, M.; RASO, T. Notes Language into Act Theory (L-Act). In: RASO, T.; MELLO, H. (Ed.). *Spoken Corpora and Linguistic Studies*. Amsterdam: John Benjamins, 2014. p. 468-495. Doi: <https://doi.org/10.1075/scl.61.15mon>

MONNIN, P.; GROSJEAN, F. Les structures de performance en français: caractérisation et prédiction. *L'Année Psychologique*, [s.l.], v. 93, p. 9-30, 1993.

MORAES, J. A.; RILLIARD, A. Illocution, Attitude and Prosody: A Multimodal Analysis. In: RASO, T.; MELLO, H. (Ed.). *Spoken Corpora and Linguistic Studies*. Amsterdam: John Benjamins, 2014. p. 233-270. Doi: <https://doi.org/10.1075/scl.61.09mor>

NESPOR, Marina; VOGEL, Irene. *Prosodic Phonology*. Dordrecht: Foris Publications, 1986.

NI, C. J.; ZHANG, A. Y.; LIU, W. J.; XU, B. Automatic Prosodic Break Detection and Feature Analysis. *Journal of Computer Science and Technology*, [s.l.], v. 27, n. 6, p. 1184-1196, 2012.

NICKELS, S.; OPITZ, B.; STEINHAUER, K. ERPs Show that Classroom-Instructed Late Second Language Learners Rely on the Same Prosodic Cues in Syntactic Parsing as Native Speakers. *Neuroscience Letters*, [s.l.], v. 557, p. 107-111, 2013.

OSTENDORF, Mari; PRICE, Patti; SHATTUCK-HUFNAGEL, Stefanie. Boston University Radio Speech Corpus LDC96S36. DVD. Philadelphia: Linguistic Data Consortium, 1996.

PAUKER, E.; ITZHAK, I.; BAUM, S. R.; STEINHAUER, K. Effects of Cooperating and Conflicting Prosody in Spoken English Garden Path Sentences: ERP Evidence for the Boundary Deletion Hypothesis. *Journal of Cognitive Neuroscience*, Cambridge, v. 23, n. 10, p. 2731-2751, 2011. Doi: <https://doi.org/10.1162/jocn.2011.21610>

PIERREHUMBERT, J. B. *The Phonology and Phonetics of English Intonation*. 1980. Dissertation (Doctoral) – Massachusetts Institute of Technology, 1980.

PIETRANDREA, P.; KAHANE, S.; LACHERET, A.; SABIO, F. In: RASO, T.; MELLO, H. (Ed.). *Spoken Corpora and Linguistic Studies*. Amsterdam: John Benjamins, 2014.

PIKE, K. *The Intonation of American English*. Ann Arbor: University of Michigan Press, 1945.

PRICE, P. J.; OSTENDORF, M.; SHATTUCK-HUFNAGEL, S.; FONG, C. The use of prosody in syntactic disambiguation. *The Journal of the Acoustical Society of America*, [s.l.], v. 90, n. 6, p. 2956-2970, 1991.

RASO, T.; CAVALCANTE, F.; MITTMANN, M. M. Prosodic Forms of the Topic Information Unit in a Cross-Linguistic Perspective: a First Survey. In: GSCP INTERNATIONAL CONFERENCE, 2016, Napole. *Proceedings...* Napole. (Forthcoming).

RASO, T.; MELLO, H. (Ed.). C-ORAL-BRASIL I. *Corpus de referência do português brasileiro falado informal*. Belo Horizonte: UFMG, 2012.

RASO, T.; MELLO, H. (Ed.). *Spoken Corpora and Linguistic Studies*. Amsterdam: John Benjamins, 2014.

RASO, T.; MELLO, H. (Ed.). The C-ORAL-BRASIL II. *Corpus de referência do português falado (formal em contexto natural, mídia e telefone)*. (Em preparação).

RASO, T.; MITTMANN, M. As principais medidas da fala. In: RASO, T.; MELLO, H. (Ed.). C-ORAL-BRASIL I. *Corpus de referência do português brasileiro falado informal*. Belo Horizonte: UFMG, 2012. p. 177-220.

RASO, T.; VIEIRA, M. A Description of Dialogic Units/Discourse Markers in Spontaneous Speech Corpora Based on Phonetic Parameters. *CHIMERA. Romance Corpora and Linguistic Studies*, Madrid, v. 3, p. 221-249, 2016.

REDI, L.; SHATTUCK-HUFNAGEL, S. Variation in the Realization of Glottalization in Normal Speakers. *Journal of Phonetics*, [s.l.], v. 29, p. 407-29, 2001.

RIETVELD, A.; GUSSENHOVEN, C. On the Relation Between Pitch Excursion Size and Prominence. *Journal of Phonetics*, [s.l.], v. 13, p. 299-308, 1985.

ROBIN, D. A. *et al.* Auditory Perception of Temporal and Spectral Events in Patients with Focal Left and Right Cerebral Lesions. *Brain and Language*, [s.l.], v. 39, p. 539-555, 1990.

ROCHA, B.; RASO, T. O pronome lembrete e a Teoria da Língua em Ato: uma análise baseada em corpora. *Veredas*, Juiz de Fora, v. 17, n. 2, p. 39-59, 2013.

SCHEGLOFF, E. A. Reflections on Studying Prosody in Talk-in-Interaction. *Language and Speech*, [s.l.], v. 41, n. 3-4, p. 235-263, 1998.

SCHUURMAN, I.; SCHOUPPE, M.; HOEKSTRA, H.; Van der WOUDE, T. CGN, an Annotated Corpus of Spoken Dutch. In: INTERNATIONAL WORKSHOP ON LINGUISTICALLY INTERPRETED CORPORA (LINC-03), 4TH, 2003, Budapest. *Proceedings...* Budapest: EACL, 2003.

SELKIRK, E. Sentence prosody: Intonation, Stress and Phrasing. In: GOLDSMITH, J. A. (Ed.). *The Handbook of Phonological Theory*. Oxford: Blackwell, 1995. p. 550-569.

SHAH, A, P.; BAUM, S. R.; DWIVEDI, V. D. Neural Substrates of Linguistic Prosody: Evidence from Syntactic Disambiguation in the Productions of Brain-Damaged Patients. *Brain and Language*, [s.l.], v. 96, p. 78-89, 2006.

SHRIBERG, E.; STOLCKE, A.; HAKKANI-TÜR, D.; TÜR, G. Prosody-Based Automatic Segmentation of Speech into Sentences and Topics. *Speech Communication*, [s.l.], v. 32, n. 1-2, p. 127-154, 2000.

SIGNORINI, S. *Topic e soggetto in corpora di italiano parlato spontaneo*. 2004. Dissertation (Ph.D.) – Università Firenze, Firenze, 2004.

STEINHAUER, K.; ALTER, K.; FRIEDERICI, A. D. Brain Potentials Indicate Immediate Use of Prosodic Cues in Natural Speech Processing. *Nature Neuroscience*, [s.l.], v. 2, n. 2, p. 191-196, 1999.

STEINHAUER, K. Electrophysiological Correlates of Prosody and Punctuation. *Brain and Language*, [s.l.], v. 86, n. 1, Special issue on the Neuronal Basis of Language, p. 142-164, 2003.

STEINHAUER, K.; FRIEDERICI, A. D. Prosodic Boundaries, Comma Rules, and Brain Responses: The Closure Positive Shift in ERPs as a Universal Marker for Prosodic Phrasing in Listeners and Readers. *Journal of Psycholinguistic Research*, [s.l.], v. 30, n. 3, p. 267-295, 2001.

SWERTS, M. Prosodic Features of Discourse Units. 1994. Thesis (PhD) – Technische Universiteit Eindhoven, 1994.

SWERTS, M. Prosodic Features at Discourse Boundaries of Different Strength. *The Journal of the Acoustical Society of America*, [s.l.], v. 101, n. 1, p. 514-521, 1997.

SWERTS, M.; COLLIER, R.; TERKEN, J. Prosodic Predictors of Discourse Finality in Spontaneous Monologues. *Speech Communication*, [s.l.], v. 15, n. 1-2, p. 79-90, 1994.

TEIXEIRA FALCÃO, B. H. *Correlatos fonético-acústicos de fronteiras prosódicas na fala espontânea*, 2017. Tese (Doutorado) – Universidade Federal de Minas Gerais, Belo Horizonte, 2017.

TESNIÈRE, L. *Éléments de Syntaxe Structurale*. Paris: C. Klincksieck, 1965.

THORNTON, R. Children's Acquisition of Syntactic Knowledge. In: ARONOFF, M. (Ed.). *Oxford Research Encyclopedia of Linguistics*, 2016. Doi: <https://doi.org/10.1093/acrefore/9780199384655.013.72>

TSENG, C. Y.; CHANG, C. H. Pause or No Pause? Prosodic Phrase Boundaries Revisited. *Tsinghua Science and Technology*, [s.l.], v. 13, n. 4, p. 500-509, 2008.

TSENG, C. Y.; FU, B. Duration, Intensity and Pause Predictions in Relation to Prosody Organization. In: EUROPEAN CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOGY, INTERSPEECH, 9th, 2005. Lisboa. *Proceedings...* Lisboa: SISCA, 2005. p. 1405-1408. Available at: <<http://www.ling.sinica.edu.tw/eip/FILES/publish/2007.4.12.99500673.0143164.pdf>>. Retrieved on: Dec. 1, 2016.

TYLER, J. Prosodic Correlates of Discourse Boundaries and Hierarchy in Discourse Production. *Lingua*, [s.l.], v. 133, p. 101-126, 2013.

Van DONZEL, M. E. Prosodic Aspects of Information Structure in Discourse. Den Haag: Holland Academic Graphics/IFOTT, 1999.

WAGNER, A. Acoustic Cues for Automatic Determination of Phrasing. In: SPEECH PROSODY 2010 – INTERNATIONAL CONFERENCE, 5th., 2010, Chicago. *Proceedings...* Chicago: ISCA, 2010. Available at: <https://www.isca-speech.org/archive/sp2010/papers/sp10_196.pdf>. Retrieved on: Sept. 2018.

WAGNER, M.; WATSON, D. G. Experimental and Theoretical Advances in Prosody: A Review. *Language and Cognitive Processes*, [s.l.], v. 25, n. 7-9, p. 905-945, 2010.

WIGHTMAN, C. W.; SHATTUCK-HUFNAGEL, S.; OSTENDORF, M.; PRICE, P. J. Segmental Durations in the Vicinity of Prosodic Phrase Boundaries. *The Journal of the Acoustical Society of America* [s.l.], v. 91, n. 3, p. 1707-1717, 1992.

XU, Y. In Defense of Lab Speech. *Journal of Phonetics*, [s.l.], v. 38, n. 3, p. 329-336, 2010.

YANG, Y.; WANG, B. Acoustic Correlates of Hierarchical Prosodic Boundary in Mandarin. In: SPEECH PROSODY, 2002, Aix-en-Provence. *Proceedings...* Aix-en-Provence: Laboratoire Parole et Langage, 2002.

YOON, T-J.; CHAVARRÍA, S.; COLE, J.; HASEGAWA-JOHNSON, M. Intertranscriber Reliability of Prosodic Labeling on Telephone Conversation Using ToBI. In: SPEECH PROSODY INTERNATIONAL CONFERENCE, 2004. Nara, Japan. *Proceedings...* Nara: ISCA, 2004. p. 2722-2732.

ZATORRE, R. J. Cerebral Correlates of Human Auditory Processing: Perception of Speech and Musical Sounds. In: SYKA, J. (Ed.). *Acoustical Signal Processing in the Central Auditory System*. Prague: Plenum Press, 1997. p. 453-468.