

Pistas prosódicas para a segmentação da entrevista sociolinguística

Raquel Meister Ko. Freitag
Departamento de Letras Vernáculas
Universidade Federal de Sergipe, UFS
São Cristóvão, Brasil
rkofreitag@uol.com.br

Abstract— Prosodic marks in sociolinguistic interview can be correlation with text type/discourse sequence. Pauses and disfluencies (segment repetition, hesitation etc.) are quantified in two sociolinguistic interview and these results are correlated at topic theme, topic subjectivity and text type/discourse sequence. Results point that there more disfluencies occur in more formal text types. These results can be applied in segmentation of sociolinguistic interviews process, collaborating to automatic segmentation.

Style; sociolinguistic interview; pauses; segmentatiton (key words)

I. INTRODUÇÃO

Gênero do discurso acadêmico, especificamente da Sociolinguística Variacionista [3,11,13], a entrevista sociolinguística se configura como um protocolo que visa fazer emergir o vernáculo de um indivíduo representativo de uma dada comunidade de fala [14].

A realização de uma coleta de dados com entrevistas sociolinguísticas é um modo relativamente rápido e altamente eficaz para se constituir uma amostra de uma comunidade de fala, garantindo a confiabilidade e a intersubjetividade da análise. Este protocolo de coleta de dados vem sendo replicado desde a década de 1960, partindo da premissa de que o estilo como o cuidado prestado à fala: “existem mais estilos e dimensões estilísticas do que um analista pode isolar. Mas descobrimos que *os estilos podem ser dispostos ao longo de uma única dimensão, medida pelo grau de atenção prestada à fala*” [14]. Como seu foco eram fenômenos variáveis atuantes no nível fonológico, para mensurar a atenção à fala, nos estudos da década de 1960, Labov controlou estilo contrastando a fala espontânea à fala repetida (e, por isso, mais cuidadosa), como no estudo do /R/ nas lojas de departamento de Nova Iorque, ou, então, contrastando a fala espontânea à leitura de palavras, como o fez no estudo sobre a centralização dos ditongos em Martha’s Vineyard. Os resultados obtidos em suas investigações corroboraram o contínuo *leitura >> fala cuidada >> fala casual*. Quando os estudos variacionistas expandem para os domínios gramaticais mais altos, os efeitos estilísticos não deixam de existir; o seu controle, no entanto, precisa ser reelaborado, de modo a captar a dimensão estilística da variação no plano além da fonologia.

Partimos da premissa de que a dimensão estilística da variação seria muito mais apropriadamente captada se fossem utilizadas outras formas de se obter os dados sociolinguísticos, como, por exemplo, gravações de interação de um mesmo indivíduo exercendo diferentes papéis sociais, em casa, no trabalho, na escola, etc. Mas, considerando que os bancos de dados sociolinguísticos, particularmente no Brasil, foram constituídos segundo o protocolo da entrevista sociolinguística, ainda que coletas mais apropriadas para tal propósito sejam feitas, farto registro linguístico seria desperdiçado.

Apesar das críticas que vêm sendo feitas a este modelo de coleta de dados – a entrevista sociolinguística é menos natural ainda que outros tipos de entrevista, já que o entrevistado espera uma situação formal e que segue um questionário; além da intransponível relação assimétrica estabelecida entre o pesquisador de campo (da universidade) e o falante, os tipos de texto/sequência discursiva que emergem em uma entrevista sociolinguística são muito diferentes daqueles de uma conversação natural [9] –, a entrevista sociolinguística ainda pode ser considerada uma fonte produtiva para se obter diferentes estilos de fala, na medida em que seu roteiro permite que haja uma transição entre distintos tópicos discursivos e sequências discursivas, desde que sejam controladas as pistas sobre o que poderia delinear esses diferentes estilos.

No Brasil, existem vários bancos de dados aos moldes da sociolinguística variacionista constituídos por entrevistas sociolinguísticas, aos moldes do que foi implementado pioneiramente pelo PEUL [8], e que permitem uma análise da série histórica, além da dimensão diatópica, diagenérica e diamésica da variação. Por isso, pesquisadores têm se dedicado a buscar estratégias para o controle do estilo dentro da entrevista sociolinguística, a fim de ampliar o poder explanatório desta abordagem e ampliar o poder descritivo dos bancos de dados já constituídos. Propostas metodológicas para captar nuances de mudança de estilo do falante, focando, particularmente, dois aspectos constitutivos da entrevista sociolinguística – o controle do tópico discursivo [1,2] e o tipo de texto/sequência discursiva [5,6,11,13,15] – foram desenvolvidas, sendo necessário, ainda, seu aprimoramento.

Assim, dando continuidade à discussão acerca de estratégias metodológicas para captar nuances de mudança de estilo dentro da entrevista sociolinguística [13], especificamente quanto ao controle de sequências/tipos de

texto e ao controle do tópico e complexidade do assunto [11,13], neste trabalho enfocamos as pistas prosódicas significativas produzidas pelos falantes que podem auxiliar na definição de uma trilha de segmentação da entrevista sociolinguística, articuladas às segmentações baseadas em critérios textuais-discursivos.

O controle de pistas de contextualização – que podem ser linguísticas, como a alternância de código, de dialeto ou de estilo, paralinguísticas, como pausas, tempo da fala, hesitações, e prosódicas, como entonação, acento, tom, ritmo, entre outros – auxilia na observação de pressuposições contextuais [5].

A correlação entre a regularidade de padrões prosódicos e segmentação de seqüências discursivas/tipos de texto é sugerida por Guy et alii [5] que, ao buscarem motivações para a mudança entonacional em curso no inglês australiano (elevação crescente final, conhecida como *Australian Question Intonation*), analisaram, dentre os fatores linguísticos, o controle do tipo de seqüência discursiva (*text type* no original). A motivação para a escolha, segundo os autores, foi o fato de que em investigações preliminares a ocorrência de elevação crescente final em frases declarativas foi predominante em narrativas, sugerindo que o tipo de seqüência discursiva exercia influência no fenômeno.

Em termos prosódicos, os tipo textuais/seqüências discursivas correspondem a fonostilos; a distribuição de pausas e disfluências, articulada ao controle do tipo de tópico temático, pode dar pistas para a sua segmentação em entrevistas sociolinguísticas, colaborando também para a caracterização das suas dimensões de formalidade/informalidade.

II. METODOLOGIA

Com base em estudos que avaliaram aspectos prosódicos de narrativas [4], buscamos averiguar se pistas prosódicas podem auxiliar no balizamento para a segmentação de entrevistas sociolinguísticas no português. Para tanto, procedemos ao cálculo da frequência de pausas e disfluências (hesitações, repetições de palavras ou repetições de segmentos) por segmento de tipo textual/seqüência discursiva (narrativa episódica, narrativa habitual, relato de procedimento,

explicação e opinião) [11], além do tópico temático discorrido e o seu grau de subjetividade.

Como *corpus*, foram tomadas 2 entrevistas sociolinguísticas extraídas da amostra Fala Culta – Itabaiana/SE, estratificadas quanto ao sexo do falante, gravadas em campo (Marantz PMD661), com duração entre 50-60 minutos. Inicialmente, foram identificadas as seqüências discursivas prototípicas e consensuais (casos ambíguos ou de interação foram descartados neste momento).

A amostra Fala Culta – Itabaiana/SE faz parte do Banco de Dados Falares Sergipanos [10,12], base de dados linguísticos ampla da variedade de português do Estado de Sergipe, seguindo duas linhas de coleta – a de comunidades de fala (estratificação homogeneizada) e a de comunidades de práticas (relações sociopeçoais). Seu propósito é dar subsídios à investigação de variedades linguísticas do português, em seus diferentes níveis (do morfofonológico ao discursivo) e com diferentes propósitos (dos descritivos aos moldes sociolinguísticos às aplicações educacionais, subsidiando programas de ensino de língua materna).

III. RESULTADOS E DISCUSSÃO

Os dados sob análise foram coletados segundo a metodologia da Sociolinguística Variacionista, em campo, e não em estúdio, o que traz sensíveis prejuízos à qualidade acústica, inviabilizando uma análise prosódica mais aprofundada. Tais limitações impuseram a restrição da análise aos fatores prosódicos mais salientes, como as pausas e as disfluências. Feitas estas restrições, foram tabulados em números absolutos e submetidos à análise estatística do teste ANOVA ($p < 0,05$).

Foram analisados 70 segmentos de entrevistas sociolinguísticas da amostra Fala Culta, divididos quanto ao tipo de seqüência (29 explicação, 18 opinião, 10 habitual, 7 narrativa e 6 procedimento), grau de subjetividade (34 subjetivo, 21 intersubjetivo e 15 extrasubjetivo) e tópico temático discorrido (29 trabalho, 28 estudo, 4 lazer, 3 cidade, 2 viagem, 2 família e 2 rotina), aos quais foram cotejadas as ocorrências de disfluências e pausas, conforme resultados nas figuras 1-3.

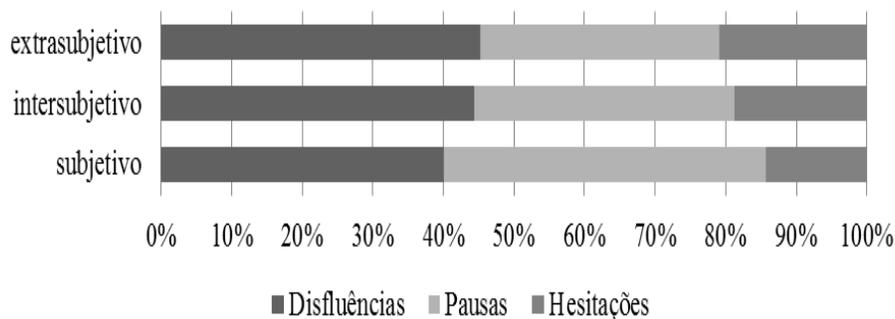


Figure 1. Grau de subjetividade

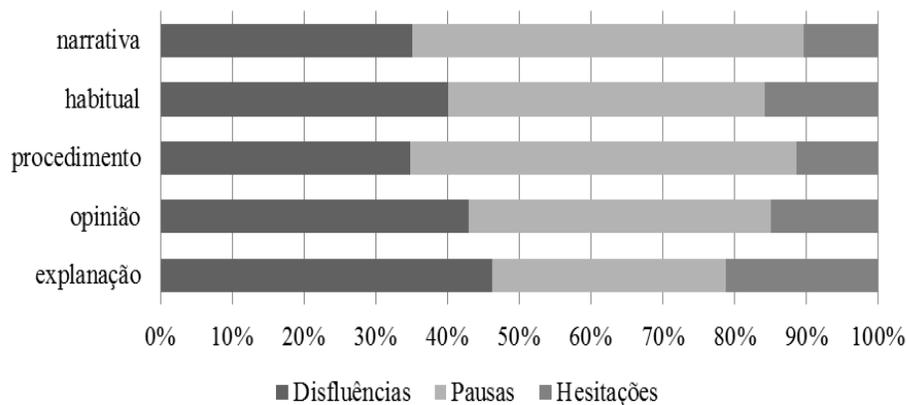


Figure 2. Tipo de texto/sequência discursiva

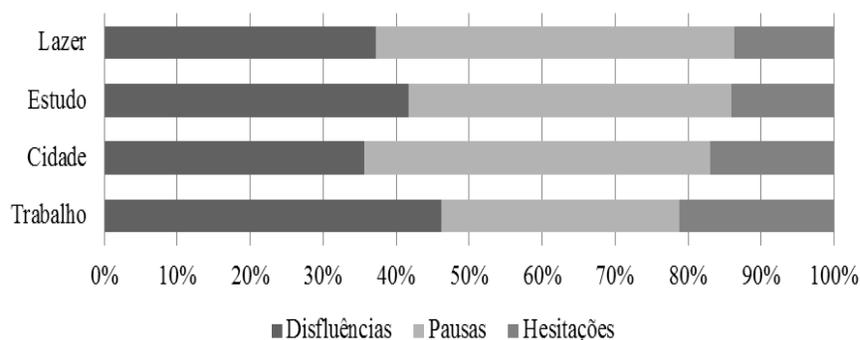


Figure 3. Tópico discursivo discorrido

Os resultados apontam que, na amostra analisada, há correlação na ocorrência de pausas e disfluências com o tópico discursivo discorrido ($p < 0,01$): tópicos mais formais, relacionados ao trabalho e ao estudo, tendem a favorecer mais a ocorrência de disfluências (fig. 3). Tipos de textos mais formais também estão correlacionados à maior ocorrência de disfluências (fig. 2).

As pausas tendem a ser associadas a contextos de maior subjetividade e informalidade (fig.1), na medida em que sua recorrência é mais significativa em tópicos mais informais, como lazer, e em sequências narrativas e relatos de procedimento, possivelmente decorrentes do relaxamento promovido pelas estratégias de minimizar os efeitos do paradoxo do observador [7] que são adotadas na realização da entrevista sociolinguística.

A necessidade de maximizar a comparação entre amostras de fala de múltiplos falantes, nas amostras de bancos de dados sociolinguísticos, tem levado à investigação da relação entre tipos textuais e a emergência de fenômenos variáveis, do envolvimento afetivo do informante com o tópico discursivo e da relação entre tipos textuais formais e informais. Pistas prosódicas também podem ser utilizadas nessa correlação.

Os bancos de dados sociolinguísticos brasileiros adotam transcrição ortográfica, ortográfica adaptada, e,

excepcionalmente, transcrição fonética, baseados numa abordagem impressionista, que depende da sensibilidade do transcritor. No entanto, pausas e disfluências costumam ser controladas, mesmo em bancos que adotam transcrição ortográfica. Considerando as premissas de comparabilidade e de compatibilidade, seria possível automatizar a identificação de sequências discursivas dentro das entrevistas sociolinguísticas a partir do controle de pistas prosódicas como as pausas e as disfluências [12].

IV. CONSIDERAÇÕES FINAIS

Segmentar uma entrevista sociolinguística não é tarefa simples e objetiva; faz-se necessário ainda refinar os critérios e aprimorar a matriz de traços constituintes de estilo que permitam a segmentação.

A análise da correlação entre disfluências e pausas na entrevista sociolinguística, cotejadas ao grau de subjetividade do tópico, aponta para resultados que respaldam a hipótese assumida, de que pistas prosódicas podem também auxiliar na segmentação de tipos de texto/sequências discursivas dentro da entrevista sociolinguística.

No entanto, é preciso considerar que a mudança de estilo, dentro da entrevista sociolinguística, se dá não só por conta de fatores externos ao falante, como a formalidade da situação, a

audiência e o tópico, mas fatores mais internos, como a relação entre os interlocutores (entrevistado/entrevistador) e a sua própria posição dentro do grupo social ou valores sociais/normas, configurando o *speaker design* [9]. Por isso, a segmentação da entrevista sociolinguística é uma tarefa que deixa margem de dúvidas no que tange à intersubjetividade: é preciso refinar a classificação das sequências discursivas considerando o perfil sociopessoal dos participantes e contexto. O controle das pistas prosódicas, ainda que em pequena escala, mostrou-se promissor para auxiliar no propósito de segmentar a entrevista sociolinguística.

Destacamos, ainda, que a sistematização do controle da distribuição de frequências de pausas e disfluências pode contribuir também para automatização do processo de segmentação, ainda que realizada de modo manual e impressionístico pelos pesquisadores, ou, ao menos, auxiliando no refinamento do controle, na medida que permite o desenvolvimento de um algoritmo que correlacione a frequência de traços prosódicos a determinado tipo de sequência discursiva.

REFERÊNCIAS

- [1] A. Bell (1984), "Language style as audience design" *Language in Society*, 13 (2):145-201.
- [2] A. Bell, "Back in style: reworking audience design" in: P. Eckert and J. Rickford (eds.). *Style and sociolinguistic variation*. Cambridge: Cambridge Press, 2001, pp. 139-169.
- [3] A. Bell, "Style in dialogue: Bakhtin and sociolinguistic theory" in R. Bailey, and C. Lucas (eds.). *Sociolinguistic variation: theories, methods, and applications*. Cambridge: Cambridge University Press, 2007, pp. 90-109.
- [4] A. C. F. Matte (2006). "Taxa de elocução, grupo acentual, pausas e fonostilística: temporalidade na prosa e na poesia com interpretação livre". *Estudos Linguísticos* 35(1): 276-285. Available online at: <http://www.gel.org.br/estudoslinguisticos/edicoesanteriores/4publica-estudos-2006/sistema06/28.pdf>
- [5] G. Guy et alii (1986), "An intonational change in progress in Australian English". *Language in Society* 15(1): 23-51.
- [6] G. Oliveira E Silva, and A. T. Macedo "Análise sociolinguística de alguns marcadores conversacionais" in A. T. Macedo, C. Roncarati, and M. C. Mollica (eds.) *Variação e discurso*. Rio de Janeiro: Tempo Brasileiro 1996, pp. 11-49.
- [7] J. J. Gunperz, "Convenções de contextualização" In: B. T. Ribeiro and P. M. Garcez. (eds.). *Sociolinguística Interacional*. São Paulo: Loyola, 2005, pp. 89-107.
- [8] M. M. P. Scherre, C. Roncarati, "Programa de Estudos sobre o uso da língua (PEUL): origens e trajetórias" in S. Votre, and C. Roncarati (eds.). *Anthony Julius Naro e a Linguística no Brasil – uma homenagem acadêmica*. Rio de Janeiro: 7 Letras/FAPERJ, 2008, pp. 37-49.
- [9] N. Schilling-Estes (2007) "Variation and the Sociolinguistic Interview: a reconsideration". *25th International AESLA Conference* (Spanish Society for Applied Linguistics), University of Murcia. Available online at <http://www.um.es/lacell/aesla/contenido/pdf/9/schilling.pdf>
- [10] R. M. K. Freitag, (2013) "Banco de dados Falaes Sergipanos" *Working Papers em Linguística*, 14 (2): 156-164. Available online at: <http://dx.doi.org/10.5007/1984-8420.2013v14n2p156>
- [11] R. M. K. Freitag, "Dissecando a entrevista sociolinguística: estilo, sequência discursiva e tópico" in E. M. Görski, I. L. Coelho and C. M. N. Souza (orgs.). *Variação estilística: reflexões teórico-metodológicas e propostas de análise*. Florianópolis: Insular, 2014, pp. 125-141.
- [12] R. M. K. Freitag, M. A. Martins, and M. A. Tavares (2012), "Bancos de dados sociolinguísticos do português brasileiro e os estudos de terceira onda: potencialidades e limitações" *Alfa*, 56 (6): 917-944. Available online at: <http://dx.doi.org/10.1590/S1981-57942012000300009>
- [13] R. M. K. Freitag, M. S. Reis, A. C. P. Back, C. A. R. Snichelotto and D. Dal Mago (2009) "O controle do gênero textual/sequências discursivas na motivação da variação linguística: apontamentos metodológicos". *Odisseia*, 3 (1): 1-15. Available online at: <http://www.periodicos.ufrn.br/index.php/odisseia/article/view/2051/1485>
- [14] W. Labov 2008 (1972). Padrões sociolinguísticos. São Paulo: Parábola Editorial.
- [15] W. Labov, "The anatomy of style-shifting" in: P. Eckert and J. Rickford (eds.). *Style and sociolinguistic variation*. Cambridge: Cambridge Press, 2001, pp. 85-108.