

Análise de valências, debate acadêmico e contenda política

Sentiment analysis, academic debate and political contention

A academia brasileira carece de debates internos. Essa deficiência é certamente produto de um complexo colonial que temos perante a produção acadêmica que vem do norte global, particularmente dos Estados Unidos, da Inglaterra, da Alemanha e da França. Este artigo faz parte de uma dessas raras ocasiões em que dois acadêmicos brasileiros, cientistas políticos no caso, dedicam-se a discutir um assunto de relevância científica e, ao mesmo tempo, pública. Temos muito que melhorar nesse sentido, nos quesitos da polidez, das boas práticas editoriais e da ética profissional acadêmica. Mas tenho certeza que essas coisas só são aprendidas na prática.

Este debate foi iniciado por Luis Felipe Miguel, em trabalho apresentado no encontro da Associação Brasileira de Pesquisadores em Comunicação e Política (Compólitica) de 2015, publicado em artigo alguns meses depois nesta revista (Miguel, 2015, p. 167). Logo em seguida, submeti artigo à revista, respondendo às objeções feitas por Miguel, inclusive mostrando como suas críticas à análise de valências, além de equivocadas, poderiam ser dirigidas ao seu

1 É professor de Ciência Política e vice-diretor do Instituto de Estudos Sociais e Políticos da Universidade do Estado do Rio de Janeiro (IESP-UERJ). Coordena o Grupo de Estudos Multidisciplinares da Ação Afirmativa (GEMAA), o Laboratório de Estudos da Mídia e Esfera Pública (LEMEP), que abriga os sites Manchetômetro (disponível em: <<http://www.manchetometro.com.br>>), e o Congresso em Notas (disponível em: <<http://congressoemnotas.tumblr.com>>), e o Observatório das Ciências Sociais. É autor de vários livros, estudos e artigos científicos sobre teoria política, políticas de ação afirmativa, relações raciais e de gênero, mídia e política, ciëntometria e estudos da América Latina. E-mail: <feres.joao@gmail.com>.

próprio trabalho anterior, baseado em análise de enquadramentos midiáticos (Feres Júnior, 2016).

No mesmo número em que meu artigo foi publicado, janeiro-abril de 2016, período em que Miguel era editor da revista, também aparece, em seguida à minha contribuição, um texto curto de autoria do próprio Miguel, com o título econômico *Uma Resposta* (Miguel, 2016). Esta contribuição tem por objetivo comentar as objeções e as críticas contidas neste último texto do professor da Universidade de Brasília (UnB). Ou seja, vou assumir que não há necessidade aqui de repassar todos os argumentos já publicados anteriormente, que podem ser facilmente acessados pelo leitor interessado na página da Revista Brasileira de Ciência Política (RBCP) no *site* do Scielo.

Miguel tenta desqualificar o Manchetômetro na abertura do texto, dizendo tratar-se de “um projeto de intervenção, não um projeto acadêmico”. Logo em seguida, faz uma rápida varredura do meu currículo Lattes, a fim de revelar que publiquei, até o momento da escrita de seu texto, somente um artigo com dados do Manchetômetro. De fato, devo me explicar em relação à minha produção de artigos acadêmicos que utilizam os dados do Manchetômetro. O *site* produz dados diariamente, utilizando a Metodologia de Análise de Valências (MAV). Transformar estes dados em artigos acadêmicos é trabalho que leva bem mais tempo. No momento em que escrevo este texto, há dois artigos no prelo e dois passando por processo de revisão por pares, todos em revistas qualificadas, algumas estrangeiras. Há, porém, outro detalhe no meu Lattes que talvez tenha escapado a meu crítico. Pesquisa anterior que coordenei sobre a cobertura que a grande mídia fez das políticas de ação afirmativa de recorte racial, também utilizando a análise de valências, já me renderam artigos na Revista Compolítica (Daflon e Feres Júnior, 2012), na Revista Opinião Pública (Feres Júnior e Daflon, 2015), na Revista Contemporânea (Feres Júnior, Campos e Daflon, 2011) e na própria RBCP (Campos, Feres Júnior e Daflon 2013). O ponto importante aqui não é minha biografia acadêmica, mas o fato de a MAV não ter sofrido objeção de pareceristas ou de editores quando da apreciação desses manuscritos.

Meu crítico está de fato correto em afirmar que o *Manchetômetro* é um projeto de intervenção. É óbvio e explícito nas páginas do *site* que seu objetivo é intervir no debate sobre a democratização e a qualidade da comunicação de massas em nosso país. Isso não desqualifica, contudo, a utilização de métodos acadêmicos de análise. Pelo contrário, é exatamente por utilizar métodos comprovados pela academia que o *Manchetômetro* obteve tanta repercussão no contexto eleitoral de 2014 e depois dele, no Brasil e no exterior.

Ao contrário do que sugere o professor da UnB, o uso de métodos acadêmicos em aplicações e serviços exteriores ao confinamento universitário é muito comum na sociedade contemporânea. O exemplo mais óbvio é o emprego de métodos econométricos em serviços e *sites* de análise de mercado. Se esse exemplo não servir, por tratar-se do emprego da ciência para a racionalização instrumental, tomemos o uso de metodologias de análise química para o monitoramento de emissões poluentes,² ou o uso de métodos de demografia, farmacologia e estatística para controle de adição a drogas,³ estes sim com claro intuito de intervenção no debate público. Assim, o uso de métodos acadêmicos para intervenção pública deveria ser festejado, pois somente assim podemos sair do ambiente da mera troca de opiniões impressionistas para o debate entre opções de políticas públicas e de regulação que tenham verdadeiro impacto naquilo que toca o bem comum.

Resta saber se a metodologia adotada no *Manchetômetro*, a MAV, é acadêmica, ou melhor, verdadeiramente científica. Penso já ter demonstrado no meu artigo anterior não só que a MAV é um tipo de análise de enquadramento – metodologia preferida por Miguel –, mas também que ela é mais precisa do que as metodologias concorrentes utilizadas pelos estudos de mídia, inclusive do que as modalidades de análise de enquadramento semanticamente mais

2 Ver, por exemplo, o *site* da Metro Vancouver, uma parceria de municipalidades canadenses que se dedica ao monitoramento de emissões poluentes. Disponível em: <<http://www.metrovancouver.org/services/air-quality/emissions-monitoring/emissions/emission-inventories/Pages/default.aspx>>.

3 Ver o *site* do The European Monitoring Centre for Drugs and Drug Addiction (EMCDDA). Disponível em: <<http://www.metrovancouver.org/services/air-quality/emissions-monitoring/emissions/emission-inventories/Pages/default.aspx>>.

complexas. A despeito do que pensam muitos cientistas sociais, particularmente aqueles afeitos a métodos matemáticos, não há um modelo ideal de ciência, seja ela exata, seja ela natural, seja ela social, que resista à passagem do tempo e das gerações. O estudo da história da ciência mostra que o pragmatismo é a melhor e mais sensível abordagem para entendermos a dinâmica científica, inclusive porque é uma teoria social da ciência. Como mostra Thomas Kuhn, o critério para sabermos se um método é científico é averiguar se ele é aceito como “normal” pela comunidade acadêmica de uma determinada época (Kuhn, 1962).

Ora, para provar que a MAV é ciência normal nos estudos de mídia brasileiros, bastaria citar os trabalhos pioneiros do saudoso Marcus Figueiredo (Aldé, Mendes e Figueiredo, 2007) e de seus discípulos (Aldé, 2002; Mundim, 2012; 2014; Borba, 2015), ou mesmo os artigos publicados por mim e por meus colaboradores sobre a cobertura midiática da ação afirmativa, já citados. Mas isso é pouco; constatamos um enorme número de trabalhos que utilizam a MAV quando ampliamos o escopo da busca para publicações em língua inglesa. Tal metodologia é empregada para o estudo da mídia em contextos eleitorais, em tudo similar ao Manchetômetro, desde a década de 1950 (Rowse, 1957) até os dias de hoje (Kahn e Kenney, 2002; Dalton e Beck, 1998; Hopkins e King, 2010; King, Pan e Roberts, 2013). Outros trabalhos empregam-na em conjunto com a avaliação do impacto eleitoral do viés midiático (Della Vigna e Kaplan, 2007; Kahn e Kenney, 2002; Druckman e Parkin, 2005; Gentzkow, 2006). A literatura internacional mais recente integra a MAV à análise computacional de textos (Van de Rijt et al., 2013; Irak, 2015; Hao-Chiang Koong et al., 2016), uma inovação tecnológica que o Manchetômetro vai disponibilizar em questão de meses.

Uma análise superficial dessas referências mostra que a MAV foi usada em vários artigos publicados na *American Political Science Review* (APSR). Eu mesmo não gosto do perfil da APSR, mas se há *normal science* em ciência política, esta revista da Associação Norte-Americana de Ciência Política é forte candidata a representá-la. É curioso o fato de meu crítico dizer em seu texto que procurou no

Google pela abreviação MAV, utilizada por mim em meu artigo, e não a encontrou. Se tivesse procurado por expressões em inglês, como *slant analysis*, *sentiment analysis* ou *opinion mining*, todas sinônimas de “análise de valências”, teria encontrado um mundo de referências a trabalhos acadêmicos publicados em revistas internacionais com alto fator de impacto.

Por fim, queria tratar de um assunto fundamental na crítica de Miguel: a de que a MAV “promove interpretações arbitrárias e mutila de maneira irreversível a complexidade da relação comunicativa” e que, portanto, deveria ser substituída por “outras ferramentas metodológicas, menos bombásticas, mais sutis, mas também mais acuradas” (Miguel, 2016, p. 300). Aqui vou escolher uma metodologia que Miguel parece achar mais “acurada”, a análise de enquadramentos, pois a empregou algumas vezes em seus trabalhos, e compará-la à MAV. Ambas são metodologias de análise de conteúdo, ou seja, estão atentas ao significado do texto estudado. Já mostrei em meu ensaio anterior que a MAV é de fato um caso, uma espécie, do *genus* análise de enquadramento, e que ela, a MAV, requer uma interpretação mais simples do material – estabelecer se ele é contrário, favorável ou neutro em relação ao objeto em questão. Já a análise de enquadramentos demanda a identificação por parte dos codificadores de construtos semânticos bem mais complexos, e, portanto, está muito mais sujeita a erro e à discordância entre codificadores.

Imaginem o seguinte experimento: colocamos dois codificadores no acostamento de uma estrada e pedimos para cada um anotar os veículos que passarem pelo ponto em que estão. Munidos de um relógio, uma caneta, um papel e uma prancheta, eles registrarão o horário que o veículo passa e sua categoria – se carro, ônibus ou caminhão. Após uma hora de experimento, interrompemos o procedimento e colhemos os resultados. Se o movimento da estrada não for intenso e se os codificadores tiverem um mínimo de capacidade cognitiva, é bem provável que haja coincidência total nos resultados. Adicionamos, então, a variável da cor do carro à tarefa de identificação. Após uma hora de trabalho, é provável que apareçam algumas discordâncias quanto à cor dos carros, dos

ônibus e dos caminhões anotados. Em seguida, adicionamos o número de pessoas dentro do veículo ao rol de variáveis, e depois o estado dos pneus – se novos, meio usados ou carecas, e assim por diante. É óbvio que a adição de mais variáveis, ou seja, a complexificação do conteúdo a ser analisado, aumenta o potencial de discordância entre os codificadores. Da mesma forma, a identificação de unidades semânticas mais complexas, como é o caso da análise de enquadramentos, certamente redundante em maior discordância potencial entre codificadores e, portanto, na diminuição da acuidade do resultado.

Se a meta é produzir interpretações menos arbitrárias e mais acuradas, há métodos para atingi-la, ou pelo menos para avaliar o grau de falta de acuidade, prontamente disponíveis na literatura acadêmica. Um deles, talvez o mais importante, é o teste de confiabilidade entre codificadores (*intercoder reliability test*). Há várias versões, da simples conferência da concordância percentual entre codificadores ao *pi* de Scott, *kappa* de Cohen, *kappa* de Fleiss e *alfa* de Krippendorff (Neuendorf, 2002). Nenhuma revista internacional de prestígio aceita trabalhos que contenham análise de conteúdo codificado sem algum teste de confiabilidade, seja na sociologia, seja na ciência política, seja na psicologia. Junto à análise computacional das valências, o Manchetômetro em breve fará o controle da confiabilidade da codificação. Contudo, esse não é o caso das revistas brasileiras. Se há tanta preocupação com a acuidade, com a mitigação da arbitrariedade, devemos ultrapassar a ideia da autoridade do autor como intérprete do mundo e adotar ferramentas mais confiáveis para testar nossas hipóteses com dados empíricos.

O emprego da MAV pode fazer uma imensa contribuição ao debate sobre a democratização da mídia em nosso país. Nos Estados Unidos, essa metodologia é aplicada a uma gama de estudos em ciências sociais, inclusive para o controle regulatório da propriedade de canais de comunicação, a fim de garantir a diversidade de opiniões (Ho e Quinn, 2009), algo que infelizmente nunca fizemos em nosso país. A democracia brasileira, tão promissora até há pouco, adoeceu gravemente, e a grande mídia, ainda no controle do

quase monopólio da produção de conteúdo jornalístico em nosso país, é uma das principais causas dessa doença. Precisamos de mais análises de valência, enquadramento, agendamento, estudos de economia política, teoria política normativa etc., que tratem desse assunto tão importante para nossa vida coletiva. O trabalho acadêmico sólido é aquele que não foge ao debate metodológico ou substantivo, e, acima de tudo, que não se encolhe frente ao calor da ágora, no coração da pólis.

Referências

- ALDÉ, Alessandra. As eleições presidenciais de 2002 nos jornais. In: RUBIM, Antonio; CANELAS, Albino (Orgs.). *Eleições presidenciais em 2002: ensaios sobre mídia, cultura e política*. São Paulo: Hacker Editores, 2002. p. 106-128.
- ALDÉ, Alessandra; MENDES, Gabriel; FIGUEIREDO, Marcus. Tomando partido: imprensa e eleições presidenciais em 2006. *Política e Sociedade*, n. 10, p. 153-172, 2007.
- BORBA, Felipe de Moraes. Propaganda negativa nas eleições presidenciais brasileiras. *Opinião Pública*, v. 21, p. 268-295, 2015.
- CAMPOS, Luiz Augusto; FERES JÚNIOR, João; DAFLON, Verônica Toste. Administrando o debate público: O Globo e a controvérsia em torno das cotas raciais. *Revista Brasileira de Ciência Política*, n. 11, 2013.
- DAFLON, Verônica Toste; FERES JÚNIOR, João. Ação afirmativa na revista Veja: estratégias editoriais e o enquadramento do debate público. *Compólitica*, v. 2, n. 2, 2012.
- DALTON, Russell; BECK, Paul. Partisan cues and the media: information flows in the 1992 presidential election. *The American Political Science Review*, v. 92, n. 1, 1998.
- DELLA VIGNA, Stefano; KAPLAN, Ethan. The Fox News effect: media bias and voting. *Quarterly Journal of Economics*, v. 122, 2007.
- DRUCKMAN, James; PARKIN, Michael. The impact of media bias: how editorial slant affects voters. *The Journal of Politics*, v. 67, n. 4, 2005.

- FERES JÚNIOR, João. Em defesa das valências: uma réplica. *Revista Brasileira de Ciência Política*, n. 19, p. 277-298, 2016.
- FERES JÚNIOR, João; CAMPOS, Luiz Augusto; DAFLON, Verônica Toste. Fora de quadro: a ação afirmativa nas páginas d'O Globo. *Contemporânea*, n. 2, p. 61-83, 2011.
- FERES JÚNIOR, João; DAFLON, Verônica Toste. A nata e as cotas raciais: genealogia de um argumento público. *Opinião Pública*, v. 21, p. 238-267, 2015.
- GENTZKOW, Matthew. Television and voter turnout. *Quarterly Journal of Economics*, v. CXXI, n. 3, 2006.
- HAO-CHIANG KOONG, Lin et al. Construction of multi-mode affective learning system: taking affective design as an example. *Journal of Educational Technology & Society*, v. 19, n. 2, p. 132-147, 2016.
- HO, Daniel; QUINN, Kevin. Viewpoint diversity and media consolidation: an empirical study. *Stanford Law Review*, v. 61, n. 4, p. 781-868, 2009.
- HOPKINS, Daniel; KING, Gary. A method of automated nonparametric content analysis for social science. *American Journal of Political Science*, v. 54, n. 1, p. 229-247, 2010.
- IRAK, Da. Istanbul United Football fans entering the political field. In: DAVID, Isabel; TOKTAMIS, Kumru (Eds.). *Everywhere taksim*. Amsterdam: Amsterdam University Press, 2015. p. 137-152.
- KAHN, Kim Fridkin; KENNEY, Patrick. The slant of the news: how editorial endorsements influence campaign coverage and citizens' views of candidates. *American Political Science Review*, v. 96, p. 381-394, 2002.
- KING, Gary; PAN, Jennifer; ROBERTS, Margaret. How censorship in China allows government criticism but silences collective expression. *American Political Science Review*, v. 107, n. 2, p. 1-18, 2013.
- KUHN, Thomas. *The structure of scientific revolutions*. Chicago: University of Chicago Press, 1962.

- MIGUEL, Luis Felipe. Quanto vale uma valência? *Revista Brasileira de Ciência Política*, n. 17, p. 165-178, 2015.
- _____. Uma resposta. *Revista Brasileira de Ciência Política*, n. 19, p. 299-301, 2016.
- MUNDIM, Pedro Santos. Imprensa e voto nas eleições presidenciais brasileiras de 2002 e 2006. *Revista de Sociologia e Política*, v. 20, p. 123-147, 2012.
- _____. Cobertura da imprensa e eleições presidenciais de 2006 efeitos realmente limitados? *Revista Brasileira de Ciências Sociais*, v. 29, p. 91-107, 2014.
- NEUENDORF, Kimberly. *The content analysis guidebook*. Los Angeles: Sage, 2002.
- ROWSE, Arthur. *Slanted news: a case study of the Nixon and Stevenson fund stories*. Boston: Beacon Press, 1957.
- VAN DE RIJT, Arnout et al. Only 15 minutes? The social stratification of fame in printed media. *American Sociological Review*, v. 78, n. 2, p. 266-289, 2013.

Resumo

Este é o quarto artigo de um debate com Luis Felipe Miguel sobre a MAV e, como tal, responde às objeções feitas pelo autor em texto anterior, intitulado *Uma Resposta*. Miguel argumenta que a MAV é utilizada pelo *site* Manchetômetro com propósitos de intervenção pública, e não com fins acadêmicos. Mostro que, na sociedade de hoje, é comum metodologias acadêmicas serem usadas para subsidiar o debate público. Ele também critica a cientificidade e a acurácia da MAV. Em resposta, mostro que a MAV foi empregada muitas vezes no Brasil em estudos de mídia e é fartamente aplicada na academia de língua inglesa, inclusive nos jornais mais prestigiosos da área de ciência política. Por fim, comparo a MAV à análise de enquadramentos, metodologia fartamente utilizada na academia nacional, para mostrar que a determinação de valências de textos é operação mais simples, e, portanto, muito menos sujeita à variação, do que a interpretação de complexas construções semânticas que são os enquadramentos.

Palavras-chave: análise de valências; mídia e política; análise de enquadramento; metodologia.

Abstract

This is the fourth article in a series of exchanges with Luis Felipe Miguel about the methodology of Sentiment Analysis (SA). In his last piece, simply entitled *An Answer*, Miguel argues that SA is employed in the website Manchetômetro as a political intervention and not as academic work. I respond to that objection showing that academic methodologies are currently employed in several initiatives aimed at intervening in public debates. Miguel also claims that SA is not scientific or accurate. I then show that SA has been abundantly employed by Brazilian and English speaking scholars working on media and politics and that the most prominent political science periodicals frequently publish articles that use that methodology. Finally, through a hypothetical experiment I compare SA with frame analysis, a commonly used methodology in Brazilian media studies, to show that the latter must necessarily produce less accurate results than the former.

Keywords: sentiment analysis; news media and politics; frame analysis; methodology.

Recebido em 2 de julho de 2015.

Aprovado em 17 de novembro de 2015.

Obs.: A resposta do autor João Feres ao debate com Luis Felipe Miguel foi avaliada apenas pelo Comitê Editorial da RBCP.