

JOÃO GONÇALVES, SARA PEREIRA & MARISA TORRES DA SILVA

id5322@alunos.uminho.pt; sarapereira@ics.uminho.pt; marisatorresilva@hotmail.com

CENTRO DE ESTUDOS DE COMUNICAÇÃO E SOCIEDADE - CECS / UNIVERSIDADE DO MINHO, BRAGA, PORTUGAL / UNIVERSIDADE NOVA DE LISBOA, LISBOA, PORTUGAL

MEDIR O AGENDA-SETTING NOS COMENTÁRIOS DOS LEITORES ÀS ELEIÇÕES LEGISLATIVAS DE 2015

RESUMO

O presente estudo propõe o uso dos comentários dos leitores como objeto para medir os fenómenos de *agenda-setting*. A análise de frequência de palavras é o método escolhido para operacionalizar a avaliação do *agenda-setting*. O método proposto foi testado numa amostra de 741 artigos e 52.064 comentários do jornal *Expresso*, que englobam um período de cinco semanas entre 4 de setembro e 10 de outubro de 2015. A análise de frequência de palavras permitiu não só avaliar quais os temas abordados pelo jornal que mais ecoaram nos comentários, mas também estudar o processo inverso, ou seja, quais os temas abordados pelos comentadores que foram sub-representados nas notícias.

PALAVRAS-CHAVE

Agenda-setting; comentários dos leitores; frequência de palavras; eleições legislativas

1. INTRODUÇÃO

O desenvolvimento das tecnologias da informação e comunicação permite explorar novos caminhos de investigação nas ciências sociais. Estes podem implicar a criação de novas técnicas de recolha e análise, a génese de áreas de investigação ou o revisitar de teorias consolidadas. No presente capítulo, propõe-se uma nova forma de estudar os fenómenos de *agenda-setting*, recorrendo à análise de frequência de palavras nas notícias e nos comentários dos leitores publicados no *website* do jornal *Expresso* durante as eleições legislativas portuguesas de 2015.

Pretendemos que este trabalho seja uma “prova de conceito” para o método proposto, demonstrando a viabilidade da sua aplicação. Nesse

sentido, levamos a cabo quatro exercícios de análise que exploram diferentes vertentes do estudo do *agenda-setting*. Estes exercícios ilustram o potencial da análise de frequência de palavras, mas destacam também algumas limitações que não devem ser descuradas. O foco da nossa reflexão é quase exclusivamente metodológico, embora esta se integre num projeto de doutoramento mais abrangente que procura estudar a articulação entre o discurso dos políticos, o tratamento e discurso jornalístico e o discurso dos públicos, aqui representados pelos comentadores. Assim, este capítulo não ambiciona apresentar conclusões sobre o contexto específico das legislativas, uma vez que esse trabalho será realizado noutros espaços, reforçando-se com métodos que complementem a frequência de palavras.

2. UMA BREVE HISTÓRIA DO AGENDA-SETTING

Há mais de 40 anos, McCombs e Shaw (1972) demonstravam o papel central assumido pelos média na definição dos temas mais importantes para os votantes nas eleições presidenciais norte-americanas de 1968. No seu célebre estudo de Chapel Hill, os autores cruzaram os resultados de 100 entrevistas a eleitores com uma análise de conteúdo aos meios de comunicação de massas de informação política a que os indivíduos tinham acesso. Os dados recolhidos por McCombs e Shaw, apresentando correlações muito elevadas, sustentam a hipótese da existência da função de *agenda-setting* dos média, ou seja, que o conteúdo da agenda mediática determina o conteúdo da agenda pública.

A preponderância dos média na influência da agenda dos públicos é desde logo preconizada por Walter Lippmann (1922) no início do século XX e mais tarde por Cohen (1963) que afirma: “a imprensa pode não ter muito sucesso na maior parte do tempo a dizer aos leitores o que pensar, mas é extraordinariamente bem sucedida a dizer-lhes sobre o que pensar”. Desde de que McCombs e Shaw cunharam o termo *agenda-setting*, a hipótese tem sido desenvolvida e expandida por vários autores¹. Funkhouser (1973), que também está na génese do estudo do fenómeno, explorou o nível macro de análise, cruzando resultados de inquéritos de opinião pública na década de 60 nos Estados Unidos com o conteúdo de três revistas semanais de informação. Iyengar e Kinder (1987) abordaram a questão a partir do nível micro, testando a teoria recorrendo a experiências com a televisão. As conclusões de ambos os estudos são favoráveis à hipótese do *agenda-setting*.

¹ Para um enquadramento mais abrangente dos estudos em torno do *agenda-setting*, consultar Dearing e Rogers (1996) e Boynton e Richardson (2016).

Desde a sua génese, centenas de estudos procuraram replicar ou desenvolver as ideias iniciais do *agenda-setting*. Para além dos estudos sobre os efeitos básicos do *agenda-setting*, têm-se desenvolvido trabalhos sobre *agenda-setting* de atributos, associados ao *framing*, sobre os efeitos psicológicos associados, sobre as fontes da agenda mediática e sobre as consequências dos efeitos do *agenda-setting* (McCombs, 2005).

Apesar da diversidade de abordagens, as investigações desenvolvidas no âmbito do *agenda-setting* recorrem com frequência à análise de conteúdo para avaliar a agenda dos média e a estudos de opinião, questionários ou entrevistas, para medir a agenda pública, seguindo os moldes dos estudos originais. Questões como “qual o problema mais importante que este país enfrenta atualmente?” da Gallup Poll norte-americana são usadas para hierarquizar as questões mais relevantes para os públicos. Funkhouser (1973) nota, por exemplo, que a formulação da questão, enquadrando-a no âmbito do país, pode fazer com que exista um enviesamento positivo a favor da influência dos média, já que estes são encarados como barómetro da nação. As respostas poderiam ser diferentes caso a questão interrogasse o inquirido diretamente sobre os problemas que este enfrenta.

Uma outra limitação dos primeiros estudos da agenda mediática prende-se com as restrições que se aplicam à análise de conteúdo das publicações e noticiários, muitas vezes confinando-se a uma contagem de artigos ou minutos dedicados a um determinado tema. “Habitualmente, os assuntos não permanecem importantes na agenda mediática por muito tempo. Localizar, contar e codificar milhares de artigos, sobretudo analisando o significado de cada peça, é uma tarefa muito aborrecida” (Dearing & Rogers, 1996). Para além de trabalhosos, estes procedimentos nem sempre espelham de uma forma precisa a saliência de um assunto, uma vez que há temas secundários ou subtilezas discursivas que ficam por registar.

Ambas as limitações podem ser ultrapassadas aproveitando as valências das tecnologias da informação e da comunicação. Por um lado, é possível recorrer a métodos não-intrusivos para medir a agenda pública, como as pesquisas em motores de busca *online* (Maurer & Holbach, 2016; Scharkow & Vogelgesang, 2011) ou as mensagens da rede Twitter (Boyn-ton & Richardson, 2016). Por outro, os programas de recolha e análise de texto permitem-nos analisar automaticamente *corpora* extensos, através de estatísticas de frequência de palavras, coocorrências e outros métodos de quantificação. Explorando estas potencialidades, pretendemos acrescentar um novo objeto aos estudos do *agenda-setting*, os comentários dos leitores nos *websites* dos jornais.

3. COMENTÁRIO: A VOZ DO LEITOR OU O ECO DO JORNALISTA

Os comentários dos leitores nas páginas da internet dos jornais têm sido um objeto de estudo popular entre os académicos (Coe, Kenski & Rains, 2014; Silva, 2013; Stroud, Scacco, Muddiman & Curry, 2015; Weber, 2014). Apesar de alguns movimentos recentes em sentido contrário (Finley, 2015), os espaços de comentário estão presentes nos *websites* de muitos dos principais meios de comunicação e os comentários dos leitores, nas redes sociais ou nos *websites* dos jornais, são lidos por uma grande parte da população (Stroud, Duyn & Peacock, 2016).

Uma fatia significativa do bolo da investigação que incide sobre os comentários tem sido guiada pela hipótese do desenvolvimento de uma esfera pública digital (Dahlberg, 2001; Papacharissi, 2002), seguindo a linha proposta por Habermas (1984a). Apesar de ter sido criticado por alguns académicos (Fraser, 1992) e reformulado pelo próprio autor (Habermas, 1992), o conceito de esfera pública continua central para os estudos da comunicação e democracia. A conceção da esfera pública, na sua base, supõe um espaço de debate público no qual pessoas privadas discutem e perante o qual o poder tem que se legitimar. A tradição deliberativa (Habermas, 1984b, 1996) determina que esta discussão deverá obedecer a um modelo crítico-racional inclusivo, em que a força do melhor argumento prevalece. Assim, muita da investigação que incide sobre os comentários dos leitores tem-se focado nos requisitos Habermasianos para a deliberação (Ruiz et al., 2011). Um destes requisitos é a autonomia, estabelecendo que o discurso deve ser orientado pelas preocupações dos cidadãos (Dahlberg, 2001).

Neste ponto, o estudo do debate público entronca com a teoria do *agenda-setting*, uma vez que esta estuda a forma como são moldadas as preocupações dos públicos. Será que é possível conceber uma esfera pública em que os temas do debate são exclusivamente determinados pelos media? A importância da discussão entre cidadãos é reconhecida pelos autores da linha do *agenda-setting*: “a variável interveniente mais estudada na explicação da relação entre a agenda dos média e a agenda pública tem sido a quantidade de discussão interpessoal gerada pela cobertura noticiosa de um assunto” (Dearing & Rogers, 1996).

Os comentários dos leitores apresentam-se como um espaço de excelência para observar esta discussão interpessoal entre cidadãos. A ligação dos comentários às notícias e a possibilidade de interação entre os utilizadores que oferecem, permitem-nos avaliar as dinâmicas de discussão de uma forma não intrusiva. Restringindo os temas das notícias abordadas, é possível focar a investigação num determinado assunto e o facto de que

grande parte das caixas de comentário estarem em funcionamento pelo menos desde os primeiros anos do milénio, possibilita a realização de estudos longitudinais com um horizonte temporal razoável. A circunstância de que muitos sistemas de comentário exigirem o registo dos seus utilizadores também permite uma análise dos hábitos e dos perfis de comentadores.

No entanto, as caixas de comentário apresentam também limitações assinaláveis enquanto objeto de estudo. Desde logo, fatores como a moderação e/ou a criação de sistemas de reputação influenciam fatores como a civilidade das discussões (Ksiazek, 2015), podendo criar um enviesamento no retrato da participação dos leitores. A diferença nas políticas de moderação e registo pode também dificultar a realização de estudos comparativos ou longitudinais, uma vez que o mesmo órgão de comunicação pode alterar o seu sistema ao longo do tempo².

A investigação que incide sobre as caixas de comentários é também limitada pelo seu potencial de generalização, uma vez que os comentadores dificilmente poderão ser considerados uma amostra representativa da população. Os comentadores representam, no entanto, um subgrupo propenso à participação e à discussão, constituindo um segmento particularmente relevante para o estudo da participação e da democracia.

Neste contexto, não será descabido estabelecer um paralelismo entre os comentadores e os líderes de opinião da teoria do *two-step flow* de Lazarsfeld, Berelson e Gaudet (2005), uma vez que os comentários podem ser vistos como um mecanismo de influência na interpretação das notícias e do trabalho jornalístico. Esta afirmação é apoiada por estudos que mostram que os comentários têm um efeito significativo na perceção da opinião pública da parte dos leitores (Lee, 2012) e que podem ter um efeito negativo na perceção da qualidade do trabalho jornalístico (Prochazka, Weber & Schweiger, 2016).

Estas conclusões relevam a importância do estudo dos comentários dos leitores no âmbito do *agenda-setting*, uma vez que estes podem ser encarados não só como um espelho da agenda pública, mas também como um dos fatores que influenciam a definição dessa mesma agenda. No entanto, o recurso aos comentários como objeto de estudo carece de um enquadramento metodológico adequado, que procuramos estabelecer na próxima secção.

² Na verdade, pelo menos três dos *websites* com as maiores comunidades de comentadores em Portugal (*Jornal de Notícias*, *Diário de Notícias* e *Público*) fizeram alterações significativas aos sistemas de comentário entre 2012 e 2015.

4. PROPOSTA METODOLÓGICA

O estudo dos comentários dos leitores apresenta desde logo uma divergência significativa da abordagem tradicional ao *agenda-setting*. Enquanto as questões do “problema mais importante” (MIP) permitem aos inquiridos listar uma série de temas espontaneamente, o conteúdo dos comentários dos leitores encontra-se fortemente condicionado pelo tema das notícias. Esta relação do comentário com o texto de notícia é formulada com particular eloquência por Foucault (1981): “o comentário exorciza o elemento aleatório do discurso dando-lhe o que lhe é devido; permite-nos dizer algo diferente do próprio texto, mas na condição de que seja este próprio texto a ser dito e, de certo modo, completado”. Será que este constrangimento invalida o estudo dos comentários no âmbito do *agenda-setting*? No nosso entender, apesar de condicionar a interpretação dos dados, esta limitação não é impeditiva para a presente proposta.

Embora os assuntos dos comentários estejam condicionados pelos temas das notícias, este estímulo não constrange o número de comentários. Isto significa que a importância de um assunto na agenda pode também ser medida pela relação entre o número de comentários e o número de notícias de um determinado tema. Claro que este número pode ainda ser influenciado pelo destaque que as notícias têm na página do jornal, mas estudos anteriores mostram que as notícias mais comentadas não seguem, por exemplo, as dinâmicas das notícias mais lidas (Tenenboim & Cohen, 2015), sugerindo que existe uma lógica para o número de comentários que não depende apenas do destaque das peças.

No nosso entender, o comentário é o melhor indicador de saliência na agenda pública do que a leitura/clique, uma vez que pressupõe um investimento e uma atitude ativa da parte do utilizador. Enquanto a leitura pode ser motivada por uma curiosidade pontual, o comentário significa que o assunto é importante o suficiente para merecer a participação do leitor. Para apoiar esta afirmação, podemos observar que a frequência de comentário está associada a motivações maioritariamente cognitivas (Springer, Engelmann & Pfaffinger, 2015).

Devido à elevada associação entre as notícias e os comentários, o estudo do *agenda-setting* nas caixas de comentários não deve procurar confirmar ou desmentir a hipótese da influência dos média na agenda pública, mas antes explorar convergências e divergências relevantes entre as duas agendas. A presença persistente de um determinado tema nos comentários que é negligenciado nas notícias, por exemplo, é um dos fenómenos que pode ser detetado com esta abordagem. Da mesma forma, o inverso

pode ser observado, existindo casos em que a agenda dos comentadores não replica a importância atribuída pelos média a um determinado tema, acontecimento ou personalidade.

Tendo estabelecido a forma como os dados devem ser interpretados, importa explicitar como pode ser medida a agenda. Recordemos que, nos primeiros estudos, era atribuído manualmente um tema a uma notícia e era depois medida a frequência do tema e/ou o tamanho das peças dedicadas a esse mesmo tema. As tecnologias da informação permitem-nos desenvolver hoje outras abordagens ao problema da medição. A contagem de palavras parece ser o método mais prático de analisar um grande volume de dados com recursos limitados, sendo que esta abordagem pode ser aplicada tanto no tratamento de notícias como no estudo dos comentários.

Embora a frequência de um termo ou conjunto de termos nem sempre seja um indicador válido da sua importância (por exemplo, Krippendorff, 2004), parece-nos que as características dos comentários e das notícias reforçam a validade deste indicador para medir a sua saliência. Tanto as notícias como os comentários são textos relativamente curtos, o que significa que o potencial de divagação do texto é limitado, centrando-se num conjunto limitado de temas e protagonistas. As próprias regras do jornalismo reforçam a validade do indicador, uma vez que o conteúdo jornalístico deve ser claro para o leitor não podendo, em teoria, dar excessivo lastro a referências indiretas e segundos sentidos que comprometeriam a integridade do indicador. Embora os comentários sejam por vezes propensos a este tipo de liberdades estilísticas, existe uma consistência discursiva que permite ao investigador atento corrigir o dicionário³ para que estas liberdades sejam acauteladas. Por exemplo, Paulo Portas, antigo vice-primeiro-ministro de Portugal, é frequentemente mencionado como “o irrevogável” devido à sua quase-demissão em 2013. Um investigador que queira medir a presença dos dirigentes políticos nos comentários pode facilmente incorporar o termo no conjunto de palavras que se referem a Paulo Portas e assim obter um retrato mais fiel da agenda dos comentadores.

O uso da frequência de palavras como medida da saliência tem ainda algumas vantagens para além da economia de tempo e de recursos. O facto de os artigos não estarem circunscritos a uma categorização manual prévia do investigador permite detetar temas secundários e pequenas variações no contexto. A análise de coocorrência pode também dar alguns indícios

³ Na análise textual computadorizada, os dicionários são as categorizações de palavras usadas pelos programas informáticos para agrupar determinados termos. Por exemplo, as palavras “défice”, “juros” e “dívida” podem ser agrupadas numa categoria “finanças”.

sobre o enquadramento que é dado ao tema, integrando na investigação algumas considerações sobre o *framing*.

5. TESTE DE APLICAÇÃO

Para ensaiar a aplicação do método proposto e decalcar as suas fragilidades foi realizado um exercício exploratório incidindo sobre uma parte da amostra do nosso projeto de doutoramento, focado nas eleições legislativas de 2015. Assim, foram submetidas a análise todas as notícias publicadas entre 4 de setembro e 10 de outubro de 2015 no jornal *Expresso* com a *tag* “Legislativas 2015” e os respetivos comentários, englobando duas semanas de pré-campanha, o período de campanha, o dia de reflexão, o dia de eleições e uma semana de pós-eleições. No total foram recolhidos 741 artigos e 52.064 comentários, arquivados entre 1 e 9 de dezembro de 2015.

Algumas limitações devem ser mencionadas à partida neste exercício. Alguns dos artigos de opinião são de acesso pago, o que significa que apenas o título, subtítulo e primeiro parágrafo estão acessíveis ao público em geral. Não sendo possível estimar a percentagem de comentadores que têm acesso integral a estes artigos, pode-se admitir um pequeno enviesamento na análise. Enquanto estes casos serão acautelados na análise final, estes não foram sinalizados para os exercícios propostos em baixo. Uma segunda limitação prende-se com peças que dão uma grande ênfase a elementos não textuais, como as fotogalerias e os vídeos, cujo conteúdo não é refletido na contagem de palavras. Também é possível acautelar estas limitações na análise, mas uma vez que estas peças têm, por norma, um número reduzido de comentários, o efeito é quase negligenciável e não será destacado neste capítulo.

O conteúdo das notícias e dos comentários foi submetido a uma análise de frequência de palavras agregada por dia. O período de tempo considerado e a unidade de agregação são mais curtos e reduzidos do que é habitual nos estudos longitudinais do *agenda-setting*. Stone e McCombs (1981) identificaram um intervalo entre dois e seis meses em que o efeito da agenda mediática na agenda pública demora a refletir-se na agenda pública. Embora o período temporal selecionado não nos permita avaliar este tipo de efeitos mais duradouros, é possível realizar outro tipo de estudo, em que são avaliadas variações mais imediatas na agenda. Este tipo de análise adequa-se à velocidade e ao ritmo das campanhas eleitorais, potenciados pela imediatez da internet. Independentemente destes fatores, nada

impede que os métodos aqui ilustrados possam ser aplicados a amostras mais alargadas temporalmente e com maiores unidades de agregação.

A recolha dos comentários foi efetuada com recurso a um programa desenvolvido propositadamente para o efeito. O arquivo da informação e parte do tratamento de dados foram efetuados no *software* Microsoft Excel. O programa WordStat foi usado para o tratamento da frequência de palavras e o programa IBM SPSS foi usado para a análise de séries temporais. Foram construídos dicionários de raiz para captar referências aos partidos políticos e principais personalidades e temas da campanha.

5.1 EXERCÍCIO 1 – MEDIR OS EFEITOS DE AGENDA-SETTING

O exercício mais simples que se pode estabelecer no enquadramento do estudo proposto é o de medir a correspondência entre a agenda mediática e a agenda dos comentadores para um determinado assunto. Para garantir a comparabilidade entre as frequências de palavras é usada a percentagem que a palavra ocupa no total do *corpus* analisado, sendo os *corpora* dos comentários e dos artigos tratados independentemente. A título de exemplo, ilustremos graficamente a presença de José Sócrates na agenda durante o período de campanha.

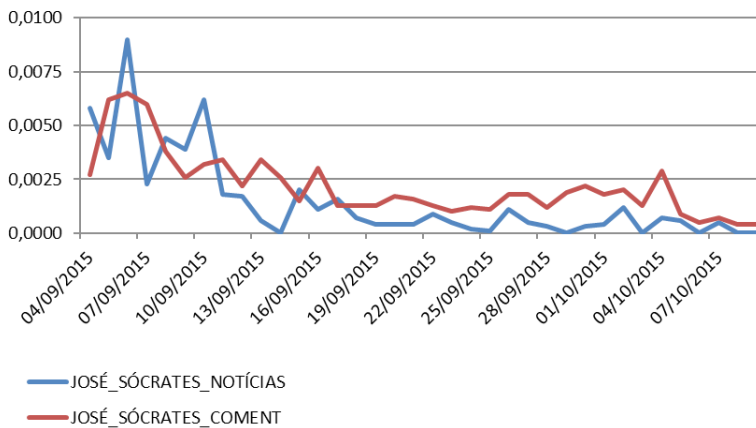


Gráfico 1: Percentagem de referências a José Sócrates no total de palavras processadas, por dia

Uma primeira interpretação do gráfico permite-nos identificar as alturas da campanha em que o antigo primeiro-ministro José Sócrates foi mais referido. Note-se, por exemplo, que no dia 6 de setembro perto de

uma em cada 100 palavras nas notícias era uma referência a José Sócrates. No entanto, mais relevante para o presente estudo, é a observação do comportamento das duas linhas, sendo que uma maior sobreposição equivale a uma maior proximidade entre a agenda dos comentadores e a agenda mediática.

De uma forma semelhante, podemos aplicar o mesmo processo aos temas abordados. Tratando o tema das pensões da mesma forma que no exemplo anterior, obtemos a seguinte representação gráfica:

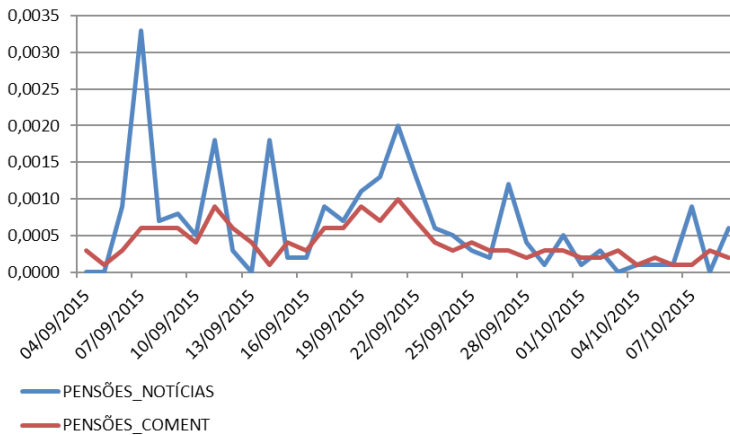


Gráfico 2: Percentagem de referências a pensões no total de palavras processadas, por dia

A representação gráfica apresentada nos exemplos acima pode ser convertida numa relação numérica, analisando a correlação cruzada entre as percentagens dos termos nas notícias e nos comentários. Assim, a título de exemplo, “José Sócrates” apresenta um valor de 0.69 e as “pensões” um valor de 0.52⁴, que significa que a agenda dos comentários coincide na globalidade com a agenda das notícias. Partindo deste processo, poderemos analisar e comparar vários termos para verificar quais os assuntos em que os comentadores mais seguem ou divergem da agenda mediática.

⁴ O coeficiente varia entre -1 e 1, sendo que os valores positivos indicam uma variação semelhante dos valores e os valores negativos uma variação inversa. O erro padrão é de 0.164.

PALAVRA-CHAVE	COEFICIENTE DE CORRELAÇÃO
BPI	1
Défi ce	0,869
Debate	0,853
.....
Europa	-0,162
Crise	-0,206
Bancarrota	-0,226

Tabela 1: Correlação cruzada entre o número de referências em notícias e em comentários, por dia

Neste caso, podemos observar uma correlação perfeita no caso do BPI e uma divergência das agendas quando os temas são a “Europa”, a “Crise” ou a “Bancarrota”. Agrupando tematicamente termos semelhantes, começamos a poder pintar um retrato mais alargado dos discursos de jornalistas e comentadores, desvelando as tendências mais abrangentes que guiam a adesão dos comentadores a uma determinada agenda.

5.2 EXERCÍCIO 2 – COMPETIÇÃO OU COEXISTÊNCIA NA AGENDA

Uma das relações que também pode ser observada nos dados é a simbiose ou a competição entre determinados assuntos na agenda mediática e na agenda pública. Zhu (1992) é um dos autores que explora este problema, conceptualizando a agenda como sendo um jogo de soma-nula, em que a saliência de um assunto implica necessariamente a diminuição da importância de outros assuntos. Apesar de concluir que esta relação se verifica na prática, Zhu realça que determinados temas podem assumir trajetórias concordantes, citando o exemplo de um estudo sobre a SIDA, o cancro e as doenças sexualmente transmissíveis na agenda (Hertog, Finnegan & Kahn, 1994). Neste caso, a hipótese da soma-zero não se verifica, indiciando que as relações entre os temas na agenda são complexas e que as generalizações devem ser evitadas.

Este tipo de relações temáticas também pode ser observado na nossa proposta, com o comportamento das percentagens a revelar os temas que variam em conjunto e os que se comportam como uma soma-zero.

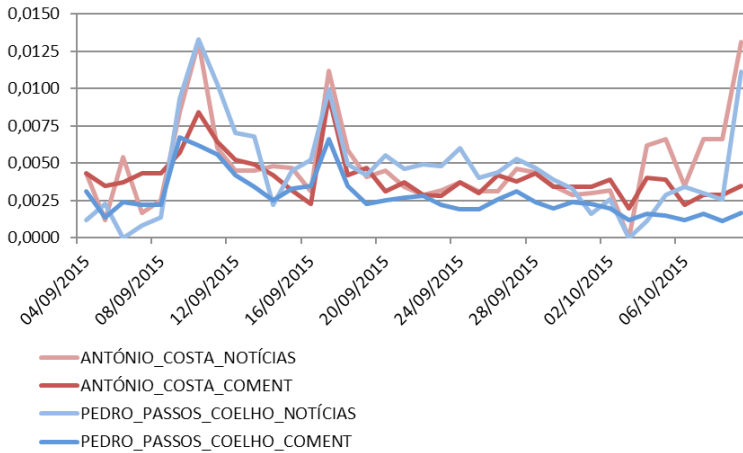


Gráfico 3: Percentagem de referências a António Costa e Pedro Passos Coelho no total de palavras processadas, por dia

No exemplo acima, podemos observar que Pedro Passos Coelho e António Costa, apesar de serem os principais adversários para ocupar o cargo de primeiro-ministro, acabam por subir e descer na saliência nas notícias (0,72) e nos comentários (0,84) em conjunto. Por outro lado, existem personalidades que competem pela agenda, como António Costa e Jerónimo de Sousa, com uma correlação negativa que, curiosamente, se verifica na agenda dos comentadores (-0,34) com maior intensidade do que na agenda mediática (-0,07).

Este tipo de análise permite iluminar relações entre temas que nem sempre são evidentes na coocorrência de palavras. Por vezes as causas para a associação de dois assuntos na agenda são mais abrangentes e não se manifestam explicitamente na mesma peça ou comentário. Aplicado a amostras e unidades de agregação maiores, este exercício pode desvelar tendências da mesma forma que os estudos longitudinais tradicionais, com a capacidade de incorporar temas não contemplados inicialmente.

5.3 EXERCÍCIO 3 –CAUSALIDADE

O intervalo temporal entre a composição da agenda dos média e os efeitos na agenda pública sugere que os segundos são causados pela primeira. No entanto, com a emergência de novos canais de participação e de comunicação entre os públicos, é mais fácil admitir hoje a possibilidade de uma inversão do sentido causal, ou seja, de os públicos colocarem

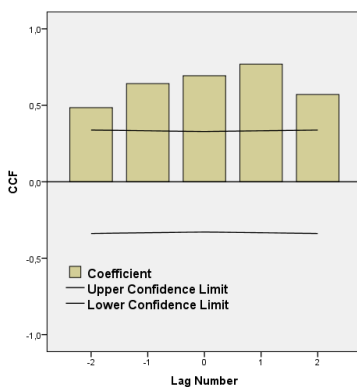
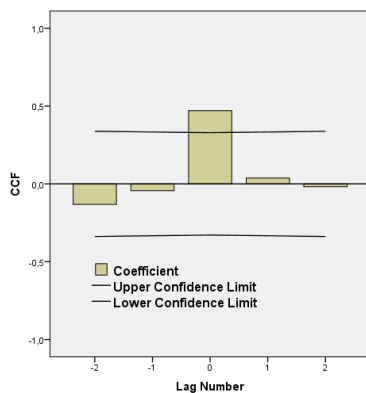
assuntos na agenda mediática. Uma das conclusões do estudo de Boynton e Richardson (2016) sobre o Twitter aponta neste sentido, sugerindo que as redes sociais introduzem uma nova dimensão no *agenda-setting*, reduzindo a influência dos média tradicionais e das elites políticas na agenda pública.

Será que o mesmo princípio se pode aplicar nos comentários? Apesar da relação estímulo-resposta que existe entre os comentários e a notícia, não será descabido conceber uma resposta desproporcional a uma notícia, que conduz posteriormente à publicação de mais peças pelo jornal sobre aquele assunto. É possível avaliar este tipo de fenómenos com uma análise temporal que mede a correlação cruzada ajustada para intervalos (*lags*). Vejamos três exemplos:

PALAVRA	LAG				
	-2	-1	0	1	2
CDU	0,533	0,738	0,617	0,543	0,215
Salários	-0,132	-0,043	0,471	0,037	-0,018
José Sócrates	0,485	0,642	0,693	0,769	0,571

Tabela 2: Correlação cruzada entre percentagem nos comentários e nas notícias dos termos CDU (Coligação Democrática Unitária), Salários e Segurança Social

A tabela lista as correlações cruzadas para cada *lag* em cada uma das palavras. O *lag* -1, por exemplo, antecipa em um dia os valores para as percentagens do termo nas notícias. Isto significa que se o valor de correlação no *lag* -1 é superior à do *lag* 0, os comentadores se anteciparam aos jornalistas a atribuir a importância a um evento, invertendo a relação causal esperada. Analisando a tabela, vemos que isto acontece no caso da CDU e observamos o fenómeno inverso no caso de José Sócrates, ou seja, os comentadores demoraram um dia a convergir com a saliência dada pelos media ao tema. Estas relações tornam-se mais claras quando representadas graficamente:



Gráficos 4, 5 e 6: Correlação cruzada entre percentagem nos comentários e nas notícias dos termos CDU, salários e segurança social, respetivamente

Observando onde se situa o valor máximo do gráfico é possível perceber até que ponto uma das partes, comentadores e média, se antecipa à outra⁵. A dispersão do gráfico também ajuda a avaliar se se tratou de um efeito pontual, como no caso da palavra salário, ou mais abrangente, como nos dois outros casos. Os gráficos mostram-nos que a palavra salário só foi evocada em peças muito específicas enquanto Sócrates e a CDU são transversais a todo o período de campanha.

Após identificar os casos em que os comentadores se antecipam aos jornalistas, é possível explorar a situação com mais pormenor para determinar as causas específicas deste fenómeno.

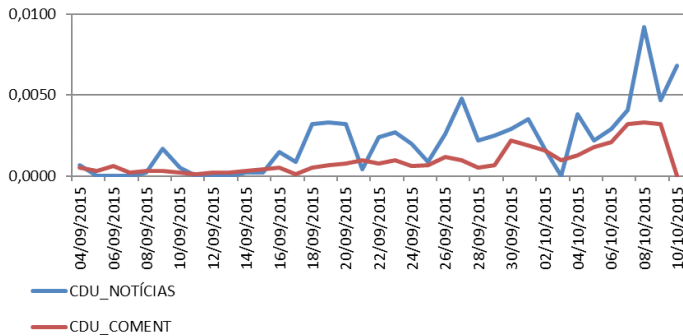


Gráfico 7: Percentagem de referências à CDU no total de palavras processadas, por dia

Observando o gráfico é possível detetar a partir do dia de eleições (4 de outubro de 2015) uma subida da CDU nos comentários que só é realmente espelhada pelas notícias no dia 8 de outubro de 2015. Analisando o contexto destes dias, os dados sugerem que ambas as subidas se devem à hipótese de a CDU viabilizar um governo do Partido Socialista (PS). Enquanto os comentadores consideraram esta possibilidade a partir do momento que foram conhecidos os resultados eleitorais, esta só foi reconhecida pela média a partir do momento em que a CDU declarou explicitamente a viabilidade da hipótese.

Este caso ilustra a forma como os dados da análise de séries temporais podem ser associados a acontecimentos e tendências reais. A análise

⁵ É necessário notar que a unidade de agregação pode ter alguma influência nos resultados, já que, por exemplo, uma notícia publicada às 23h55 terá forçosamente uma correlação maior no dia seguinte, já que a maior parte dos comentários é publicada nessa altura.

de conteúdo proposta é enriquecida quando complementada com outros métodos, que permitem explicar a variação observada nos dados e solidificar as conclusões.

5.4 EXERCÍCIO 4 – DIFERENÇAS NA SALIÊNCIA

Devido à relação direta entre os comentários dos leitores e o conteúdo das notícias, a análise proposta pode ser particularmente útil para detetar quais os assuntos em que o nível de saliência da agenda mediática mais difere da agenda dos comentadores. Estas diferenças podem ser lidas nos dois sentidos, ora buscando os assuntos destacados pelos média que não são ecoados pelos comentadores, ora olhando para as prioridades dos comentadores que não são espelhadas pelos média.

Estas diferenças de saliência podem ser calculadas, na sua forma mais básica, subtraindo à percentagem de texto que um determinado termo ocupa nos comentários dos leitores a percentagem equivalente do termo nas notícias. Os resultados serão semelhantes à Tabela 3, em que os valores positivos ilustram maior saliência nos comentários e os valores negativos refletem uma maior saliência nas notícias.

PALAVRAS MAIS SALIENTES NOS COMENTÁRIOS				PALAVRAS MAIS SALIENTES NAS NOTÍCIAS			
Palavra	Frequência	% Processada	Diferença	Palavra	Frequência	% Processada	Diferença
Dívida	3531	0,32%	0,22%	Campanha	1438	0,13%	-0,49%
Sócrates	4396	0,40%	0,22%	Líder	604	0,05%	-0,43%
Povo	2781	0,25%	0,18%	António	2956	0,27%	-0,30%
Dinheiro	2535	0,23%	0,17%	Coligação	4245	0,38%	-0,28%
Nao	1747	0,16%	0,16%	Passos	6523	0,59%	-0,24%
Troika	2244	0,20%	0,14%	Bloco	544	0,05%	-0,23%
Milhões	3662	0,33%	0,13%	Eleições	3654	0,33%	-0,19%
Pagar	1906	0,17%	0,13%	Pedro	578	0,05%	-0,19%
PS	13597	1,23%	0,11%	Socialista	1237	0,11%	-0,19%
Melhor	2482	0,22%	0,11%	Presidente	1034	0,09%	-0,19%
Gente	2263	0,20%	0,11%	Lisboa	673	0,06%	-0,18%
Votar	2359	0,21%	0,10%	Ministro	1837	0,17%	-0,18%
PEC	1213	0,11%	0,10%	Catarina	716	0,06%	-0,18%
Verdade	1926	0,17%	0,10%	Martins	302	0,03%	-0,17%

Bancarrota	1056	0,10%	0,10%	Disse	1607	0,15%	-0,17%
Será	2496	0,23%	0,09%	Jerónimo	437	0,04%	-0,17%
Pior	1294	0,12%	0,08%	Portas	2029	0,18%	-0,16%
Aumento	1477	0,13%	0,08%	CDS	2672	0,24%	-0,15%

Tabela 3: Palavras com maior diferença de
saliência entre comentários e notícias

Os dados apresentados acima permitem não só perceber quais os temas e os protagonistas sub-representados nos comentários e nas notícias, mas também nos oferece um vislumbre para as diferenças que existem na construção do discurso dos jornalistas e dos comentadores. O uso do termos como “melhor” e “pior” pelos comentadores distingue-se, por exemplo, das referências explícitas à campanha por parte dos jornalistas e de verbos de introdução do discurso como “disse”. Aplicado a um contexto mais alargado, este tipo de exercício pode contribuir para desvelar diferenças relevantes sobre a forma como determinado tema é abordado por jornalistas e comentadores.

6. DISCUSSÃO

Os exercícios apresentados acima ilustram a viabilidade do estudo dos comentários dos leitores no âmbito do *agenda-setting*. A comparação de agendas e a análise de séries temporais assumem-se como ferramentas valiosas para a interpretação dos dados. Há, no entanto, alguns cuidados importantes a ter ao recorrer a este tipo de métodos que se tornaram mais evidentes no decorrer da aplicação prática.

O método proposto é assumidamente pobre em contexto. A contagem de palavras, apesar das suas vantagens em termos de operacionalidade e de visualização, pode conduzir a interpretações enviesadas do conteúdo pelo reducionismo que implica. Em primeiro lugar, há questões operacionais e técnicas que devem ser acauteladas, já que muitas vezes os programas que permitem a recolha automática do texto das páginas de internet podem captar excertos que não estão associados ao *corpus* definido pelo investigador. Legendas de imagem ocultas, *tags* ou menus de navegação são alguns dos elementos que podem ser acidentalmente incluídos na amostra.

O contexto é também importante para a construção de dicionários e para a interpretação dos termos. Nos exercícios que realizámos, por exemplo, verificou-se que os erros ortográficos eram frequentes no texto dos

comentadores. Assim, os dicionários foram adaptados para contemplar as variações das palavras mais frequentes (por exemplo, Socrates) e assim espelhar melhor a saliência dos temas e das personalidades.

Há também opções lexicais que, não constituindo erros ortográficos, só podem ser apreendidas no seu verdadeiro sentido com o devido contexto. Ao supracitado ‘irrevogável’ Paulo Portas poderemos acrescentar, por exemplo, o “preso 44” José Sócrates ou o “lapsista” Passos Coelho, entre outras alcunhas a que comentadores e cronistas recorrem para referir os principais atores da política nacional. As mudanças podem ser, contudo, mais profundas do que a simples atribuição de alcunhas, podendo palavras de uso corrente sofrer uma metamorfose de significado em alturas de crise ou mudança social (Volosinov, 1973)⁶.

Por estes motivos, é indispensável que o investigador se familiarize com o *corpus* e com o objeto de estudo antes de levar a cabo este tipo de análise. A leitura de alguns excertos do texto e a funcionalidade *keyword-in-context* que muitos programas de análise textual integram são algumas formas de lidar com estes problemas. No entanto, estas práticas não escusam o investigador de atender às circunstâncias em que o texto é produzido e ao enquadramento temático em que este se insere. No nosso caso, seria irresponsável retirar qualquer tipo de conclusões dos dados sem considerar as características do discurso jornalístico e dos comentários dos leitores e sem considerar o contexto específico (eleições legislativas) e mais alargado (política portuguesa) em que as intervenções de ambas as partes se inserem.

Com as devidas salvaguardas sobre a importância do contexto, a análise de conteúdo dos comentários permite tratar uma grande quantidade de informação com o mínimo de recursos exigidos ao investigador. Neste momento, qualquer computador comercial é capaz de processar um grande volume de palavras para realizar análises de frequências e coocorrências. Adicionalmente, este tipo de análise também favorece os fenómenos de *serendipity* de Merton Merton (1968), uma vez que o investigador se confronta com a emergência de tendências e termos inesperados na construção de dicionários e no processo de análise.

7. CONCLUSÕES

A análise de frequência de palavras nos comentários dos leitores nos *websites* dos jornais e nas peças jornalísticas provou ser um método viável

⁶ Também atribuído frequentemente a Mikhail Bakhtin.

de estudar fenómenos de *agenda-setting*. Este caminho de investigação insere-se na tendência mais abrangente que avança no sentido de uma ciência social computacional (Chang, Kauffman & Kwon, 2014), incorporando a *big data* no estudo dos media e do comportamento social. No entanto, este tipo de tratamento de dados é mais eficaz quando complementado por outros métodos manuais ou tendencialmente mais qualitativos (Lewis, Zamith & Hermida, 2013). A medição da saliência pode ser usada não só para estudar a influência na agenda dos públicos, mas também a construção do discurso jornalístico face às reações e pressões dos públicos.

A aplicação dos exercícios apresentados acima a outros contextos e horizontes temporais contribuirá para refinar o método e avaliar melhor a sua pertinência. Uma caracterização mais detalhada de quem são os comentadores e dos fatores que determinam a sua agenda também poderá contribuir para uma melhor interpretação dos resultados.

REFERÊNCIAS BIBLIOGRÁFICAS

- Boynton, G. & Richardson, G. W. (2016). Agenda-setting in the twenty-first century. *New Media & Society*, 18(9), 1916-1934. doi: 10.1177/1461444815616226
- Chang, R. M., Kauffman, R. J. & Kwon, Y. (2014). Understanding the paradigm shift to computational social science in the presence of big data. *Decision Support Systems*, 63, 67-80. doi: 10.1016/j.dss.2013.08.008
- Coe, K., Kenski, K. & Rains, S. A. (2014). Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *Journal of Communication*, 64(4), 658-679. doi: 10.1111/jcom.12104
- Cohen, B. C. (1963). *The press and foreign policy*. Princeton: Princeton University Press.
- Dahlberg, L. (2001). The internet and democratic discourse: Exploring the prospects of online deliberative forums extending the public sphere. *Information, Communication & Society*, 4(4), 615-633. doi: 10.1080/13691180110097030
- Dearing, J. W. & Rogers, E. M. (1996). *Agenda-setting*. Thousand Oaks, California: SAGE.
- Finley, K. (2015, 10 de agosto). A brief history of the end of the comments. *Wired*. Retirado de <https://www.wired.com/2015/10/brief-history-of-the-demise-of-the-comments-timeline/>

- Foucault, M. (1981). The order of discourse. In R. Young (Ed.), *Untying the Text: A Post-Structuralist Reader* (pp. 48-78). Londres: Routledge & Kegan Paul.
- Fraser, N. (1992). Rethinking the public sphere: a contribution to the critique of actually existing democracy. In C. J. Calhoun (Ed.), *Habermas and the Public Sphere* (pp. 109-142). Cambridge, Massachussets: The MIT Press.
- Funkhouser, G. R. (1973). The issues of the sixties: an exploratory study in the dynamics of public opinion. *Public Opinion Quarterly*, 37(1), 62-75. doi: 10.1086/268060
- Habermas, J. (1984a). *Mudança estrutural da esfera pública*. Rio de Janeiro: Tempo Brasileiro.
- Habermas, J. (1984b). *The theory of communicative action*. Boston: Beacon Press.
- Habermas, J. (1992). Further reflections on the public sphere. In C. J. Calhoun (Ed.), *Habermas and the Public Sphere* (pp. 421-461). Cambridge: MIT Press.
- Habermas, J. (1996). *Between facts and norms: contributions to discourse theory of law and democracy*. Cambridge: The MIT Press.
- Hertog, J. K., Finnegan, J. R. & Kahn, E. (1994). Media coverage of AIDS, cancer, and sexually transmitted diseases: A test of the public arenas model. *Journalism & Mass Communication Quarterly*, 71(2), 291-304. doi:10.1177/107769909407100204
- Iyengar, S. & Kinder, D. R. (1987). *News that matters: Television and American opinion*. Chicago: University of Chicago Press.
- Krippendorff, K. (2004). *Content analysis: an introduction to its methodology*. Thousand Oaks: Sage Publications.
- Ksiazek, T. B. (2015). Civil interactivity: How news organizations' commenting policies explain civility and hostility in user comments. *Journal of Broadcasting & Electronic Media*, 59(4), 556-573. doi: 10.1080/08838151.2015.1093487
- Lazarsfeld, P. F., Berelson, B. & Gaudet, H. (2005). *The people's choice: How the voter makes up his mind in a presidential campaign*. Nova Iorque: Columbia University Press.
- Lee, E. J. (2012). That's not the way it is: How user-generated comments on the news affect perceived media bias. *Journal of Computer-Mediated Communication*, 18(1), 32-45. doi: 10.1111/j.1083-6101.2012.01597.x

- Lewis, S. C., Zamith, R. & Hermida, A. (2013). Content analysis in an era of big data: A hybrid approach to computational and manual methods. *Journal of Broadcasting & Electronic Media*, 57(1), 34-52. doi: 10.1080/08838151.2012.761702
- Lippmann, W. (1922). *Public opinion*. Nova Iorque: Harcourt, Brace and Company.
- Maurer, M. & Holbach, T. (2016). Taking online search queries as an indicator of the public agenda: The role of public uncertainty. *Journalism & Mass Communication Quarterly*, 93(3), 572-586. doi: 10.1177/1077699015610072
- McCombs, M. (2005). A look at agenda-setting: past, present and future. *Journalism Studies*, 6(4), 543-557. doi: 10.1080/14616700500250438
- McCombs, M. E. & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, 36(2), 176-187. doi: 10.1086/267990
- Merton, R. K. (1968). *Social theory and social structure*. Nova Iorque: The Free Press.
- Papacharissi, Z. (2002). The virtual sphere: The internet as a public sphere. *New Media & Society*, 4(9), 9-27. doi: 10.1177/14614440222226244
- Prochazka, F., Weber, P. & Schweiger, W. (2016). Effects of civility and reasoning in user comments on perceived journalistic quality. *Journalism Studies*, 1-17. doi: 10.1080/1461670X.2016.1161497
- Ruiz, C., Domingo, D., Micó, J. L., Díaz-Noci, J., Meso, K. & Masip, P. (2011). Public sphere 2.0? The democratic qualities of citizen debates in online newspapers. *The International Journal of Press/Politics*, 16(4), 463-487. doi: 10.1177/1940161211415849
- Scharkow, M. & Vogelgesang, J. (2011). Measuring the public agenda using search engine queries. *International Journal of Public Opinion Research*, 23(1), 104-113. doi: 10.1093/ijpor/edq048
- Silva, M. T. d. (2013). Participação e deliberação: um estudo de caso dos comentários às notícias sobre as eleições presidenciais brasileiras. *Comunicação e Sociedade*, 23, 82-95. doi: 10.17231/comsoc.23(2013).1615
- Springer, N., Engelmann, I. & Pfaffinger, C. (2015). User comments: motives and inhibitors to write and read. *Information, Communication & Society*, 18(7), 798-815. doi: 10.1080/1369118X.2014.997268
- Stone, G. C. & McCombs, M. E. (1981). Tracing the time lag in agenda-setting. *Journalism & Mass Communication Quarterly*, 58(1), 51-55. doi: 10.1177/107769908105800108

- Stroud, N. J., Duyn, E. V. & Peacock, C. (2016). *News commenters and news comment readers*. Engaging news project, The University of Texas at Austin. Retirado de https://engagingnewsproject.org/enp_prod/wp-content/uploads/2016/03/ENP-News-Commenters-and-Comment-Readers1.pdf
- Stroud, N. J., Scacco, J. M., Muddiman, A. & Curry, A. L. (2015). Changing deliberative norms on news organizations' Facebook Sites. *Journal of Computer-Mediated Communication*, 20(2), 188-203. doi: 10.1111/jcc4.12104
- Tenenboim, O. & Cohen, A. A. (2015). What prompts users to click and comment: A longitudinal study of online news. *Journalism*, 16(2), 198-217. doi: 10.1177/1464884913513996
- Volosinov, V. N. (1973). *Marxism and the philosophy of language*. Nova Iorque: Seminar Press.
- Weber, P. (2014). Discussions in the comments section: Factors influencing participation and interactivity in online newspapers' reader comments. *New Media & Society*, 16(6), 941-957. doi: 10.1177/1461444813495165
- Zhu, J.-H. (1992). Issue competition and attention distraction: A zero-sum theory of agenda-setting. *Journalism & Mass Communication Quarterly*, 69(4), 825-836. doi: 10.1177/107769909206900403

Financiamento e agradecimentos

Este trabalho foi financiado pela Fundação para a Ciência e Tecnologia (FCT) através de bolsa individual de doutoramento, referência SFRH/BD/103335/2014, no âmbito do POCH – Programa Operacional Capital Humano, comparticipado pelo Fundo Social Europeu e por fundos nacionais do Ministério da Ciência, Tecnologia e Ensino Superior.

Os autores gostariam de agradecer a Daniel Malhadas pela criação do programa de recolha e arquivo dos comentários.

Citação:

Gonçalves, J., Pereira, S. & Silva, M. T. (2017). Medir o *agenda-setting* nos comentários dos leitores às eleições legislativas de 2015. In Z. Pinto-Coelho, T. Ruão & N. Zagalo (Eds.), *Arte, Políticas e Práticas*. V *Jornadas Doutorais Comunicação e Estudos Culturais* (pp. 205-226). Braga: CECS.