

<https://doi.org/10.11606/issn.2317-9511.v37i0p622-643>

Explorando o vestuário na literatura em português

Clothing in lusophone literature: an exploration

Diana Santos*

* Linguatca e Universidade de Oslo. E-mail: d.s.m.santos@ilos.uio.no

Resumo: O ponto de partida para o presente artigo foi a observação de que, num corpo de cerca de 800 textos literários em português, a *Literateca*, havia consideravelmente mais menções ao vestuário nos textos brasileiros do que nos portugueses. Isto levou a uma análise mais fina do campo semântico da roupa, tentando confirmar ou refutar essa tendência, e se confirmada, identificar suas possíveis causas. Algumas tentativas de explicação com base nos campos da antropologia, linguística cultural e estudos literários são alvitadas, e alguns estudos exploratórios são apresentados usando um conjunto mais comparável de obras literárias nas duas variantes do português. O artigo tenta mostrar a vantagem metodológica de efetuar leitura distante, que guia a leitura próxima e que leva a novas fases de análise, assim como a necessidade de investigar muitas diferentes vertentes quando se lida com corpos anotados.

Palavras-chave: Linguística com corpos; Roupa; Variedades do português; Métodos estatísticos.

Abstract: My starting point for this article was the realization that Brazilian authors had a significantly higher number of mentions to clothing than Portuguese ones in *Literateca*, a corpus of about 800 lusophone literary texts. In order to establish if this was the case, and why, we looked at the semantic field of clothing with a distant reading methodology. In addition to suggesting some plausible explanations from fields such as anthropological linguistics, cultural studies and literary theory, I conducted some exploratory studies, using a more comparable set of literary works. An important methodological point of the paper is the use of distant and close reading in tandem, as well as the emphasis on the variety of issues that one has to deal with in annotated corpora.

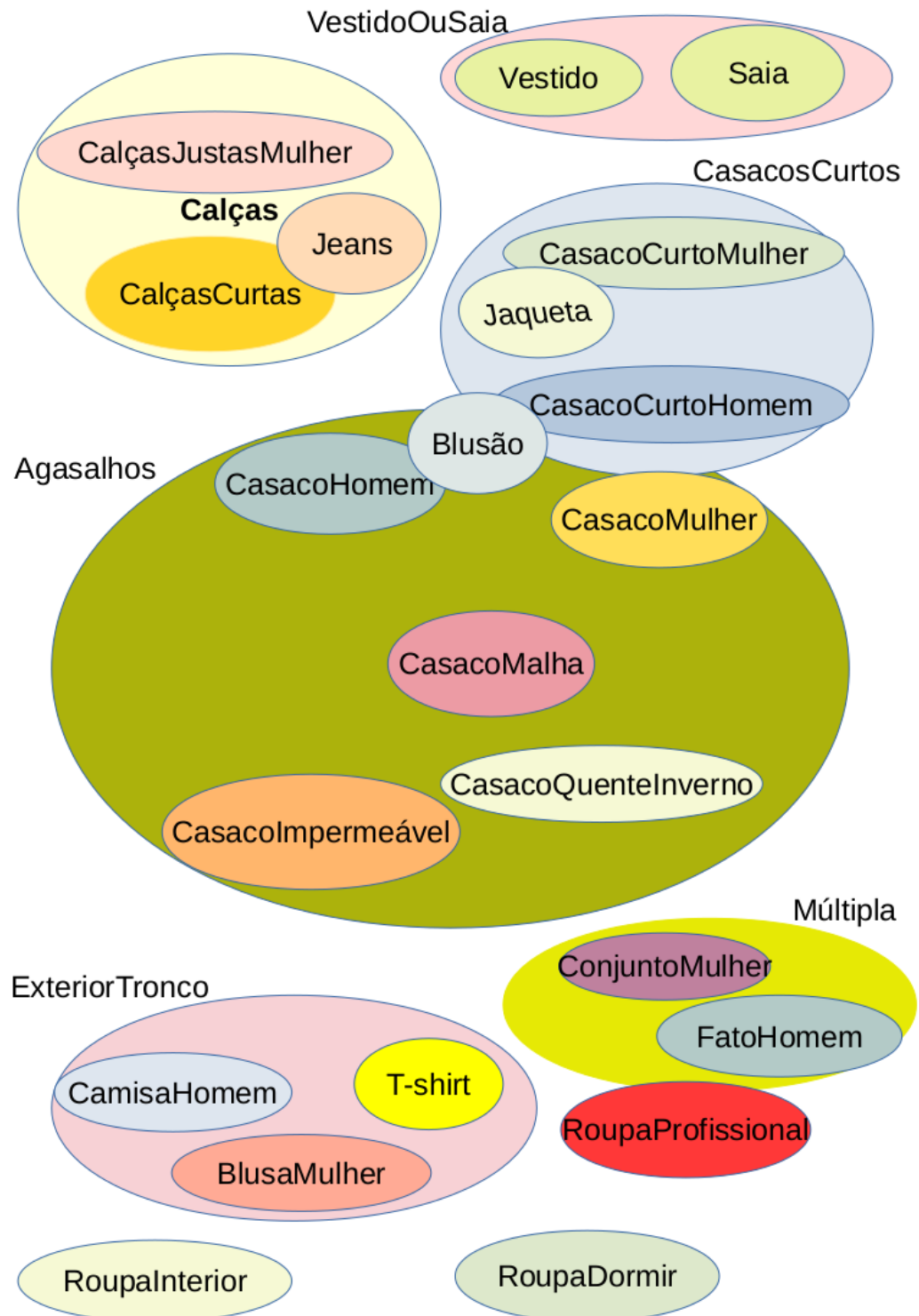
Keywords: Corpus linguistics; Clothing; Portuguese Language Varieties; Statistical methods.

1. Apresentação

A roupa é um fenómeno cultural por excelência (Eco 1975), e tem sido frequentemente usada pelos sociolinguistas cognitivos como indicador sociológico e dialetal, como o demonstram os estudos feitos por Geeraerts et al. (1994) e Geeraerts & Grondelaers (1999). Estes linguistas belgas criaram o corpo CONDIV neerlandês, que inspirou o corpo CONDIVport (SOARES DA SILVA 2008), dedicado ao estudo empírico da divergência entre o português de Portugal e o português do Brasil. Inspirados por sua vez por este recurso, resolvemos há cerca de dez anos também investir nesse campo semântico na Linguateca¹, iniciando o projeto Guarda-Fatos (SANTOS ET AL. 2011), no qual compilámos um conjunto de termos relativos ou associados a roupa, e os classificámos em vários grupos (ver figura 1), tentando manter a compatibilidade com o trabalho do ConDiv, mas alargando as informações para procura em todos os tipos de corpos e não só jornais e revistas sobre moda.

¹A Linguateca é uma rede de desenvolvimento de recursos para o processamento computacional da língua portuguesa, veja-se <http://www.linguateca.pt>.

Figura 1. Categorias de roupa no projeto Guarda-Fatos: de notar que os grupos não são mutuamente exclusivos. Além disso ainda existem as categorias Acessórios, Chapéu, Meias, Elementos, Roupa e Materiais



Os corpos do grupo AC/DC (SANTOS 2014) passaram então a ser automaticamente anotados com essa informação (que quase não foi revista) desde essa altura. Contudo, nunca desenvolvemos qualquer estudo diretamente relacionado com o vestuário. De facto, uma primeira tentativa, num artigo conjunto que escrevi há mais de dez anos com Stella Tagnin e Elisa Teixeira sobre anotação da cor e da roupa no CorTrad (SANTOS et al. 2012), os pareceristas foram da opinião que bastaria focarmos a cor... e nunca foi publicado o estudo sobre a roupa.

Mas, em 2019, quando estávamos a explorar a Literateca², uma das propriedades que nos chamou a atenção foi a aparente diferença entre os autores portugueses e brasileiros no que se referia ao interesse pela roupa, relatado em Santos e Simões (2019). Este artigo é o resultado de apreciar essa questão com mais atenção.

2. A roupa como fenómeno cultural

É sobejamente conhecida a importância do vestuário nas sociedades humanas. A espécie humana é a única espécie animal que cobre o seu corpo com artefactos e símbolos, quer por razões de sobrevivência (peles em regiões geladas) quer para efetuar rituais ou identificar papéis e classes numa sociedade. A moda nas sociedades modernas é uma indústria importantíssima, e muda constantemente. A forma de vestir e os acessórios que usamos são também pistas importantes de uma identidade a vários níveis: social, intelectual, política, nacional, emocional, entre outras.

A roupa está também, por razões óbvias, ligada ao corpo humano, um campo semântico essencial na cognição humana, como enfatizado por Lakoff e Johnson (1980), veja-se também Leitão de Almeida (2009) para o português.

É por isso natural supor que a menção à roupa seja frequente quando a aparência (do corpo humano) está em foco, mas que a roupa poderia ser

² A Literateca é em traços largos o subconjunto do grupo AC/DC cujo conteúdo é literário, associado a algumas ferramentas dedicadas especialmente a estudos de leitura distante, muitas delas ainda em construção. Ver <http://www.linguateca.pt/Literateca/> para informação atualizada, e Santos (2019a) para motivação e primeiros resultados.

usada complementarmente (quanto mais roupa mais corpo) ou alternativamente (ou se menciona uma ou outro). É preciso pois investigar a interação entre estes dois domínios do conhecimento presentes na língua, observando como são usados.

Witowski & Brown (1985), estudando os termos para partes do corpo em muitas línguas do mundo, propõem que o número de termos para partes do corpo seja proporcional à necessidade de se vestir, e que portanto línguas originárias de climas mais frios fazem mais distinção nos membros do corpo.³

O presente estudo só foca uma língua, que se expandiu por vários climas. Ao encontrar realidades e necessidades diferentes, além de falantes diferentes com outras línguas, será que o seu desenvolvimento foi influenciado e divergiu precisamente no que tem a ver com o corpo e o vestuário?

E será que esses desenvolvimentos diferentes são espelhados pelas obras literárias escritas em Portugal e no Brasil?

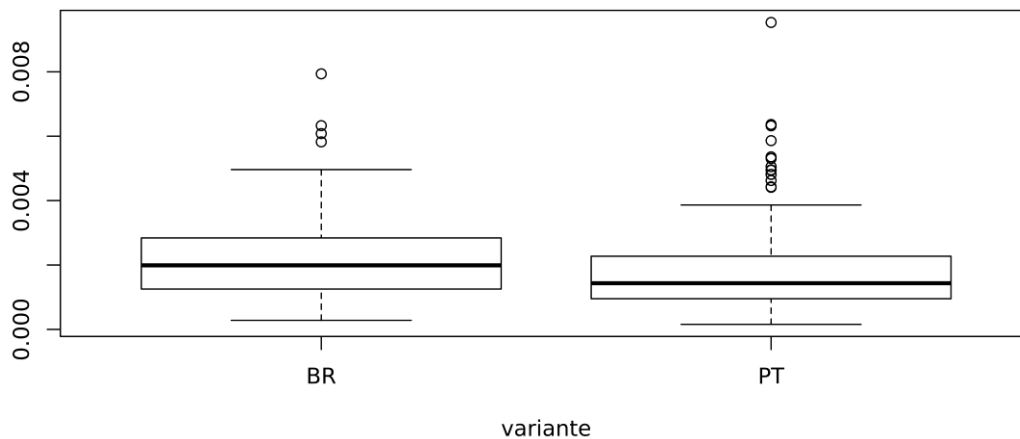
3. Primeiras explorações no material

Neste artigo, vou olhar para a questão da roupa na literatura lusófona a partir da Literateca, que contém 832 obras literárias em português.

A figura 1, que usa todos os textos brasileiros (306) e portugueses (521) da versão 4.11 da Literateca, mostra que os textos brasileiros têm mais densidade de roupa (número de palavras de roupa por número de palavras total).

³ A questão do clima e da latitude não é, evidentemente, a única questão relevante na estrutura dos termos do corpo: Burton & Kirk (1979), por exemplo, mostram como duas línguas originárias da mesma região africana apresentam conceptualizações extraordinariamente diferentes.

Figura 2. Densidade de roupa em 827 obras literárias em português



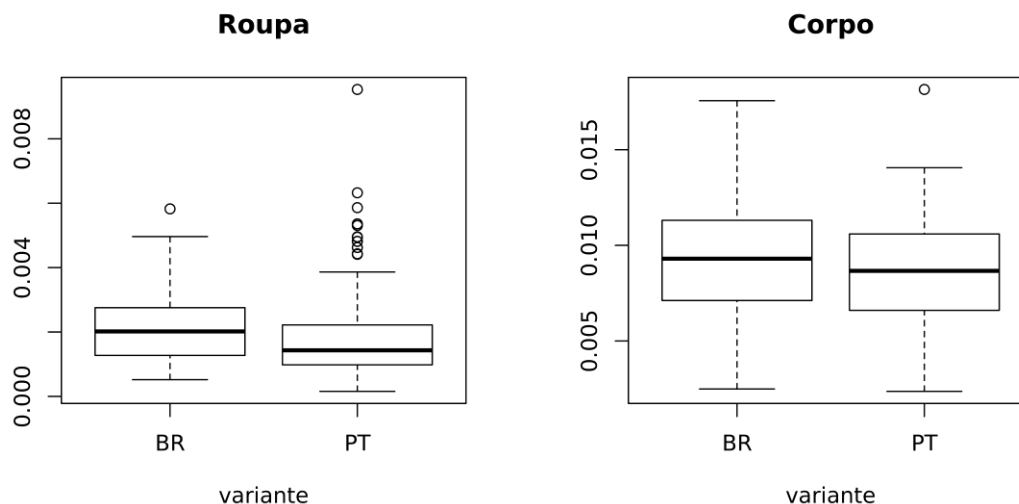
Mas: estaríamos a comparar conjuntos de textos comparáveis? Dado que o material português engloba muito mais poesia e teatro do que o material brasileiro, e que o material brasileiro é composto por muito mais contos do que o material português, é evidente que não foi apenas a variável “variedade do português” que contrastámos.⁴

Reduzimos assim o nosso universo aos romances e novelas em ambas as variantes, o que significa apenas observarmos 175 romances portugueses e 80 romances brasileiros (totalizando 255 obras).⁵ De qualquer maneira, obtivemos o mesmo resultado, visualizável no lado esquerdo da figura 3 (e confirmado por um teste estatístico, o teste t): as obras brasileiras contêm indiscutivelmente mais menções a roupa que as portuguesas.

⁴ Para a lista de obras que compõem a Literateca, ver https://www.linguateca.pt/acesso/lista_autores_literateca.html.

⁵ Por causa dos problemas de direitos de autor, quase todos os romances e novelas foram publicados antes de 1920. Dos anos 30 e 40 apenas contamos com sete obras, de cinco autores: Paulo Setúbal, Virgínia de Castro e Almeida, José Régio, Lima Barreto e Raul Brandão.

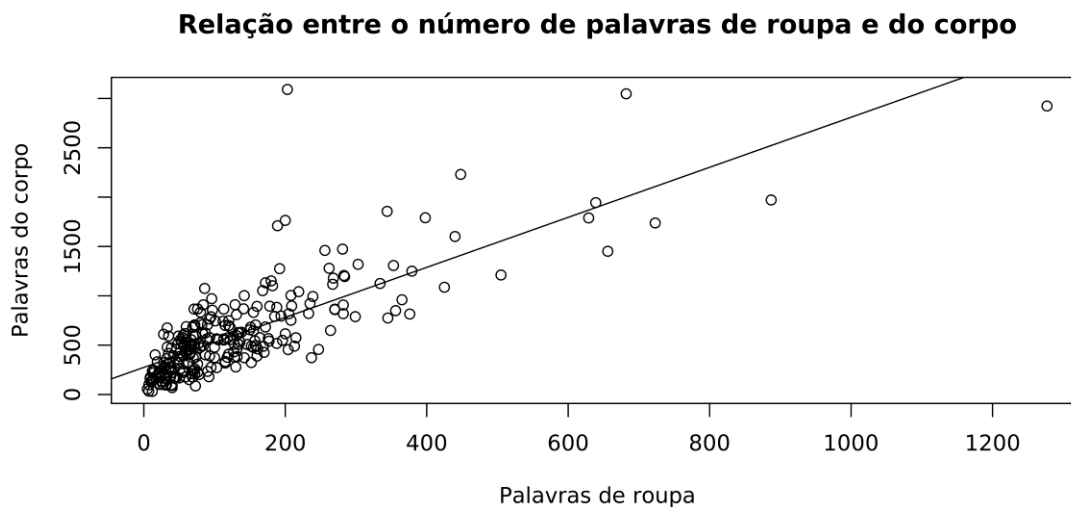
Figura 3. A distribuição da frequência relativa de roupa e de corpo por variante em 255 obras literárias em português



Para investigar se este seria um fenómeno específico da roupa ou também se verificaria em relação às palavras do corpo, a figura 3 mostra a distribuição das menções ao corpo no mesmo grupo de obras, em que, embora com diferença menor, a literatura brasileira também se destaca com maior frequência. Resta chamar a atenção para o facto de que as unidades no eixo dos YY são diferentes nos dois diagramas da figura 3: como seria de esperar, as palavras de corpo são muito mais frequentes em português.

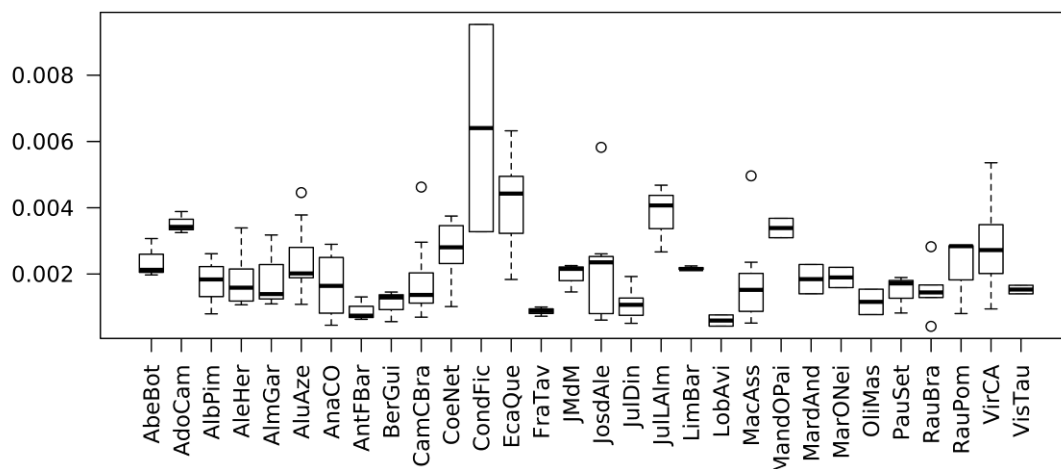
No que respeita à correlação entre os campos semânticos da roupa e do corpo, a figura 4 mostra que a correlação é positiva (0,79), mas que existe variação considerável.

Figura 4: Menções à roupa e ao corpo, em números absolutos, por obra



Tentando ver se a ocorrência de roupa depende do autor, ou se, pelo contrário, há autores que variam a referência à roupa de obra para obra, mostramos na figura 5 a variabilidade dos (30) autores que têm mais de um romance e/ou novela no nosso material. O autor com mais variabilidade, o Conde de Ficalho, conta apenas com duas obras no material, e deve essa “honra” a uma novela de apenas 6200 palavras, intitulada “Mais Uma”, sobre a escolha de uma rapariga para fugir à miséria. Seguem-se Júlia Lopes de Almeida, brasileira, e Eça de Queirós, português, como os autores com médias mais elevadas.

Figura 5. Variação de menções à roupa em vários autores



Como já mencionado em Santos & Simões (2019), não há, por outro lado, diferença significativa entre autores homens e mulheres. A nossa amostra é, contudo, reduzida a esse respeito, já que apenas contamos com 23 romances ou novelas escritas por 15 autoras no material.

4. O léxico da roupa no material

Dado que o léxico de roupa contém centenas de lemas que designam peças de roupa, acessórios e materiais, é preciso recorrer às categorias que os agrupam para ter uma ideia mais precisa do seu uso. A tabela 1 mostra, por ordem de frequência, as categorias mais mencionadas neste campo semântico, juntamente com as três palavras mais frequentes nessa categoria.

Tabela 1. Grupos de roupa mais frequentes

Categoria	Quantidade	Lemas mais frequentes
RoupaNãoEspecificada	8219	Vestir, roupa, despir
Acessório	6074	Lenço, coroa, luva
Casaco	4249	Casaca, manto, colete
Chapéu	2816	Chapéu, boné, carapuça
Materiais	2775	Seda, couro, lã
CasacoQuenteInverno	2093	Manto, paletó, capote
Vestido	2086	Vestido, vestidinho
Sapatos	2065	Sapato, bota, chinela
ElementosRoupa	1630	Botão, aba, colarinho
CasacoCerimónia	1530	Casaca, paletó, sobrecasaca

Camisola (no sentido PT) ⁶	1125	Camisa, blusa, camisola
Saia	1102	Saia, saio, saião
CasacoCurto	1071	Colete, jaqueta, gibão
Jaqueta	911	Casaca, jaqueta, jaquetão
Calças	851	Calça, calção, pantalone
RoupaProfissional	775	Farda, uniforme, libré
Meias	538	Meia, ceroula, peúga
RoupaInterior	459	Cinta, espartilho, tanga

É importante ressaltar aqui de novo que as categorias foram originalmente pensadas para o vestuário do século XX em texto jornalístico, enquanto que a maioria dos romances que analisamos neste artigo foram escritos no século XIX ou nas duas primeiras décadas do século XX. Além disso, um grande número são romances históricos (retratando, necessariamente, modas muito diferentes). Daí que muitas das categorias, como Blusão ou Jeans, nem aparecem... e provavelmente várias peças de vestuário antigo não foram reconhecidas pela anotação automática, por não constarem do léxico.

Uma revisão atenta terá de ficar para outra ocasião, mas não resisto a ilustrar uma classificação completamente inapropriada: a palavra *tanga* (que ocorre sobretudo em romances passados em África, e que identifica sempre uma peça de roupa bem visível, por única) está marcada nesta taxonomia como... RoupaInterior.

Seja como for, a tabela 1 permite-nos identificar que o grupo mais frequente é aquele que não especifica que roupa— o que é natural se pensarmos que em grande número de vezes que a roupa é mencionada com o verbo suporte *vestir*. O segundo grupo, dos Acessórios, provavelmente vem em segundo lugar devido à grande variedade de acessórios arrolados (82).

Outra forma de ter uma ideia do conteúdo dos textos é listar as palavras mais frequentes marcadas como roupa, e atentarmos na sua proporção relativa nas duas variantes (12.709.997 unidades em PT e 5.329.176 em BR). A tabela 2 apresenta os números de ocorrências, absolutos e relativos, nas duas variantes, e a probabilidade p de a diferença ser devida a puro acaso (através de um teste de proporções).

⁶ Para os leitores não conscientes desta diferença, cumpre dizer que *camisola* (PT) é referida como suéter ou casaco de lã em português do Brasil, enquanto que *camisola* (BR) é referida como camisa de noite em português de Portugal. Assim sendo, foi evidentemente uma má escolha o nome de Camisola para grupo de roupa, que exige sempre esta precisão adicional.

Tabela 2. As 20 palavras de roupa mais frequentes e sua ocorrência em BR e PT

	Casos PT	Casos BR	Relativo PT	Relativo BR	Valor de p
vestir	2230	782	17,54	14,67	1,8175 e-5
chapéu	1265	841	9,95	15,78	1,8273e-25
vestido	1319	720	10,37	13,51	1,3000e-8
roupa	914	872	7,19	16,36	3,7175e-71
seda	1089	411	8,56	7,71	0,07339
lenço	914	455	7,19	8,54	0,00301
saia	631	404	4,96	7,58	2,7541e-11
camisa	573	369	4,51	6,92	1,1740e-10
coroa	558	132	4,39	2,48	2,6318e-9
calça(s)	373	277	2,93	5,19	3,7956e-13
luva	460	138	3,62	2,59	6,2472e-4
casaca	414	156	3,25	2,93	0,2749
sapato	368	198	2,89	3,72	0,005258
botão	320	235	2,52	4,41	5,2679e-11
couro	234	319	1,84	5,98	2,1791e-47
anel	353	192	2,77	3,60	0,004194
bota	376	130	2,96	2,43	0,06431
despir	316	134	2,49	2,51	0,95386
gravata	298	147	2,34	2,76	0,11818
manto	354	90	2,78	1,69	2,3333e-05

Só no caso das palavras *bota*, *casaca*, *despir* e *gravata* é que a diferença entre os dois conjuntos de obras não é estatisticamente significativa. Assim, as obras brasileiras utilizam muito mais frequentemente do que as portuguesas as palavras *chapéu*, *vestido*, *saia* e *camisa*, enquanto as obras portuguesas só ultrapassam as brasileiras no caso de *coroa*, *manto*, *anel*

e *luvas*. Mas, se *luvas* é mais usado nas obras portuguesas, curiosamente, o verbo *enluvar* (sempre no particípio passado) é mais frequente nas brasileiras.

A palavra *coroa* fez-nos desconfiar, e consultando as concordâncias verifica-se que esta é usada sobretudo como sinónimo de realeza, e não como adereço físico, apontando para que uma futura revisão da anotação deva apenas manter como roupa os (poucos) casos usados na cabeça. Seja como for, a questão da destriça entre usos metafóricos e literais não é simples, podendo algumas pessoas argumentar que continuamos a estar em presença do campo da roupa.

Observando (agora não sistematicamente) os casos das palavras de roupa menos frequentes, mais algumas observações que me parecem de interesse:

- Ao contrário do esperado, tanto *agasalhar* como *agasalho* são mais usados nas obras brasileiras, assim como *sobrecasaca*. Pela literatura, não poderíamos concluir que o clima do Brasil é muito mais quente que o de Portugal...
- Apenas *xale* (ou *xaile*, outra forma, só usada em Portugal, de designar essa peça de vestuário feminina) é (muito) mais frequente nas obras portuguesas, assim como *corpete*. Continuando com as peças com que as mulheres se cobrem, *mantilha* é mais frequente no Brasil, enquanto *estola* volta a sê-lo em Portugal.
- Por outro lado, se *robe-de-chambre* é mais frequente na literatura portuguesa, isso deve-se a ser um galicismo muito ao gosto de Eça de Queirós, responsável pelo seu emprego em 50 num total de 66 casos.
- Como muitas obras de Portugal são romances históricos (e quase nenhuma das do Brasil), é também fácil explicar porque é que a literatura portuguesa usa *turbante* e *albornoz*, roupas claramente mouriscas, muito mais frequentemente.
- No que respeita a materiais, o *couro*, a *camurça* e o *feltro* são claramente preferidos pelos escritores brasileiros, enquanto que a *seda*, o *cetim* e o *veludo* encontram-se igualmente bem distribuídos

entre as duas literaturas.

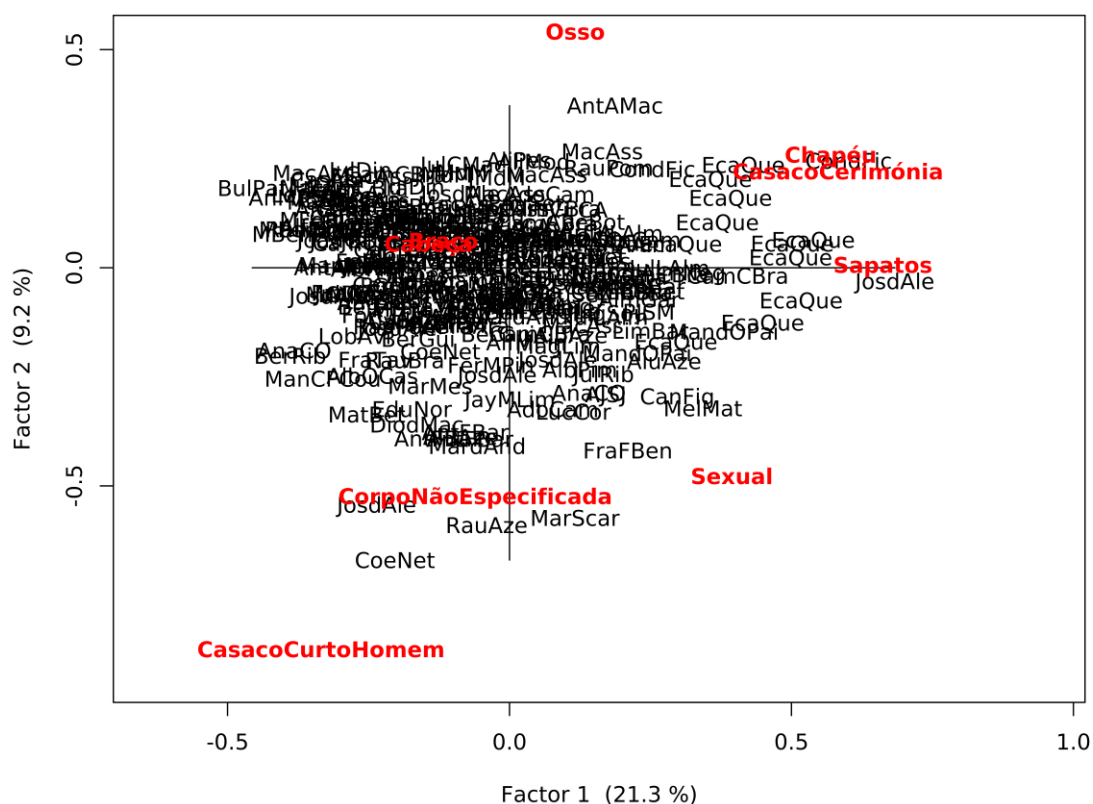
- É também interessante notar que *paletó*, um termo ainda em uso no Brasil, considerado pelos portugueses da minha geração como tipicamente brasileiro, era também usado em Portugal há mais de cem anos. De facto, *paletó* até é mais usado em Portugal na nossa amostra literária.
- Finalmente, gostaria de salientar a grande diferença entre as frequências de *vestir* e *despir*, assim como de *calçar* e *descalçar*, que indicam claramente que a primeira palavra do par também descreve um estado (o que se tem vestido ou calçado) e não apenas nem sobretudo o ato de colocar ou retirar roupa do corpo.

4. Caracterização de roupa e corpo: leituras distantes

Tendo informação de quantos casos de cada grupo de vestuário, e também de quantos casos de parte do corpo (FREITAS et al. 2015), aparecem em cada obra, será possível descobrir dimensões invisíveis ao olho humano, usando técnicas estatísticas de redução de dimensionalidade? Veja-se Baayen (2008) para uma introdução a estas técnicas, como a chamada análise fatorial (celebrizada por Biber (1985) ao descobrir dimensões comunicativas nos géneros do inglês), a análise de correspondências ou a análise de discriminantes (LDA).

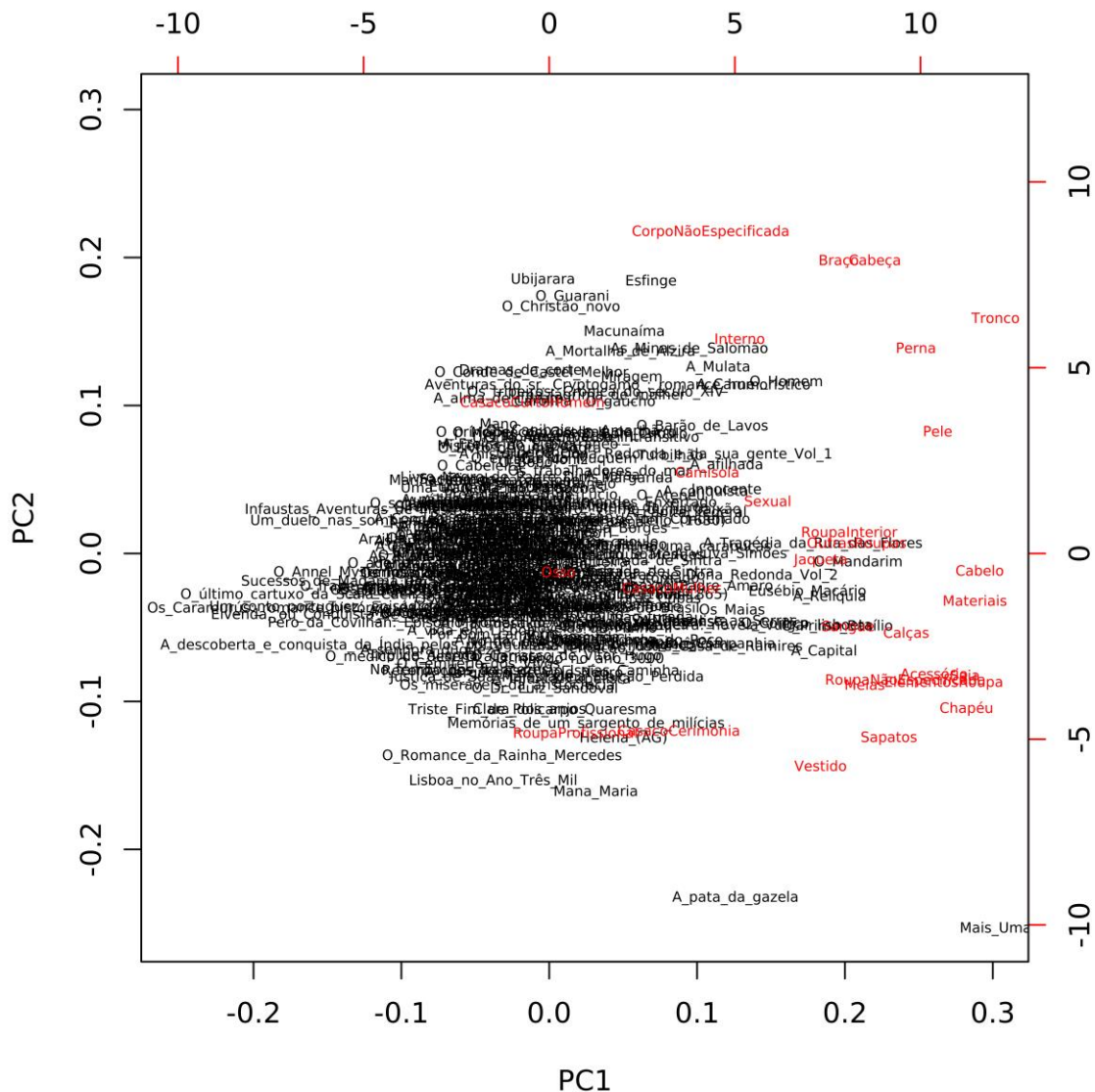
Alguns estudos exploratórios com essas técnicas são apresentados nas figuras seguintes. A figura 6, representando uma análise de correspondências, mostra os casos que são mais discriminativos (a vermelho), e que colocam as obras em quatro quadrantes diferentes.

Figura 6. Como os fatores colocam as diferentes obras no plano (análise de correspondências)



Da figura 6 vemos que as palavras do grupo CasacoCurtoHomem identificam algumas obras de Coelho Neto, Raul de Azevedo e José de Alencar, enquanto que as obras de Eça de Queirós estão próximas do grupo Chapéu e CasacoCerimônia, não muito longe do grupo Sapatos, que, não inesperadamente, coincide quase com o romance *A Pata da Gazela*, de José de Alencar.

Figura 7. Análise de componentes principais.



Na figura 7 está representada uma análise de componentes principais, mostrando os dois primeiros componentes. O segundo componente parece distinguir entre obras mais preocupadas com o corpo e obras mais interessadas no vestuário. O primeiro é mais difícil de interpretar, mas talvez distinga entre roupa por necessidade e roupa como ornamento.

Convém realçar que estes estudos exploratórios têm como missão visualizar de maneiras diferentes o material, e não fornecer uma resposta a perguntas concretas. Eles permitiram-me dar conta de algumas especificidades de algumas obras, por estas terem aparecido em locais

inesperados⁷: Foi o caso da palavra *Batas* na *Peregrinação* de Fernão Mendes Pinto, referente a um povo e não a uma peça de roupa. Ao observarmos esta obra também nos apercebemos de que várias obras portuguesas, por terem grafia antiga, não foram marcadas com todos os termos de roupa que continham (por exemplo, *velludo*).

Além disso, verificámos também que a palavra *hábito* como vestimenta de religiosos, muitíssimo usada nos romances portugueses e muito menos nos romances brasileiros, não tinha sido anotada devido à homonímia. Estes casos poderão fazer com que haja menos diferença entre as variantes, afinal. Contudo, quase metade das ocorrências de *hábito* como roupa são metafóricas (*vestir o hábito de Cristo* significa pertencer a uma ordem religiosa, *despir o hábito* significa abandonar a carreira eclesiástica, etc.).

Menciono estes pormenores aqui para mostrar a necessidade de refinar e melhorar o material muitas vezes, e assim como esclarecer e até eventualmente alterar a filosofia de anotação, antes de avançar para conclusões definitivas.

Finalmente, outro assunto que seria interessante estudar seria averiguar da predominância da roupa feminina ou masculina e mesmo calcular a percentagem de casos em que uma mulher e um homem são descritos através da forma como se vestem, na esteira do trabalho de Cláudia Freitas e colegas (FREITAS, 2019; FREITAS et al., 2020). Muito preliminarmente, podemos comparar, na Tabela 3, as menções a palavras dos Grupos Vestidos e Saias, femininos, e Calças e Jaqueta, masculinos (no período em questão). Verifica-se imediatamente que o vestuário feminino é duplamente frequente, em ambas as variantes.

Tabela 3. Termos de vestuário feminino e masculino, em grupos não ambíguos

Grupo	PT	BR
Calças	530	321
Jaqueta_CasacoCurto	691	220
Vestido	1343	743
Saia	689	413

⁷ Nem todas as visualizações são mostradas aqui.

5. Considerações finais

Este artigo apenas apresentou algumas explorações iniciais que se podem fazer sobre o vestuário em português – e que estão ao alcance de todos, visto que os corpos e anotação estão públicos.⁸ Gostaria também de investigar – talvez com a própria Stella Tagnin, fica aqui o repto - a tradução dos termos de roupa e do seu uso, por exemplo no CorTrad.

Muitas perguntas, quer do foro linguístico quer do literário, ficaram por responder, mas gostaria de colocá-las aqui para estimular trabalhos futuros:

Será que, quando a língua é comum, e não originária de diferentes climas e geografias, não se dão os fenómenos mencionados por Witowski & Brown (1980), que realçam a influência do meio-ambiente? E, nesse caso, poderíamos concluir que não haveria grandes diferenças em relação aos termos de roupa nas duas variantes do português no século XIX.

Será que as duas comunidades (pelo menos as descritas na literatura) se vestiam de forma semelhante, visto que a influência da moda vinha principalmente de França, de modo igual para os dois países? Mas, então, porquê mencionar mais roupa no Brasil?

Será que a forma de vestir era mais relevante como discriminador social no Brasil, devido a maiores diferenças sociais? Ou que, devido à forma como a sociedade estava organizada, ou às vicissitudes do clima e da geografia, seria preciso mudar mais vezes o vestuário ao longo do dia ou devido a viagens?

Ou estas diferenças (não esquecer também a menção ao corpo) devem-se às escolas literárias presentes no material⁹? Será que por haver uma diferente distribuição das obras por escola literária nas duas variantes a distribuição por roupa aparece significativamente diferente? Ou, pelo menos, a distribuição de peças de roupa específicas, associadas a grupos socio-

⁸ Os números concretos referem-se à versão 4.11 da Literateca, e foram obtidos no dia 27 de maio de 2020. Como este corpo está constantemente em aumento e melhoria, também colocamos os valores exatos em www.linguateca.pt/Gramateca/Literateca/artigoRoupa/ para que os leitores possam repetir e melhorar os estudos se o desejarem.

⁹ A Literateca não está totalmente classificada em relação à vertente “escola literária”, veja-se Santos et al. (2020) para uma problematização dessa tarefa. Seja como for, essa noção é relevante no que respeita à menção do campo semântico da saúde (SANTOS, 2019b).

económicos diferentes?

Repare-se, também, que por vezes a roupa é mencionada por fazer parte do enredo: por um protagonista trabalhar numa loja de tecidos, ou por ter havido um roubo de um anel, etc. Nesses casos seria provavelmente mais correto retirar essas menções do cômputo geral antes de comparar as obras, algo que não fizemos, mas que seria interessante aprofundar.

Tudo isto terá de ficar para futuros estudos, mas espero ter pelo menos interessado alguns leitores nestas questões, e dado a conhecer as possibilidades que a Linguateca – e em particular a Literateca – oferecem a este respeito. Tendo este artigo sido escrito para homenagear Stella Tagnin, que é uma pedagoga extraordinária, tenho esperança de que ela também possa encontrar assunto e métodos para a sala de aula inspirados por este trabalho.

Referências

- BAAYEN, H. *Analyzing Linguistic Data: A practical introduction to Statistics using R*. Cambridge University Press, 2008.
- BIBER, D. Investigating macroscopic textual variation through multifeature/multidimensional analyses. *Linguistics*, n. 23, v. 2, 1985, Mouton publishers, pp. 337-360.
- BURTON, M. L.; KIRK, L. Ethnoclassification of body parts: A three culture study. *Anthropological Linguistics*, n. 21, 1979, pp. 377-399.
- ECO, U. O hábito fala pelo monge. In: Assírio & Alvim, *Psicologia do vestir*, 3.a edição, 1989 [1975]: 7-20.
- FREITAS, C.; SANTOS, D.; CARRIÇO, B.; MOTA, C.; JANSEN, H. O léxico do corpo e anotação de sentidos em grandes corpora – o projeto Esqueleto. *Revista de Estudos da Linguagem*, n. 23.3, 2015, pp. 641-680. <http://www.periodicos.letras.ufmg.br/index.php/relin/article/view/8895/8794>
- FREITAS, Cl. Leitura distante: O que podemos com corpora anotados? Apresentação no *Primeiro Encontro de Leitura Distante em Português* (Oslo, 27-28 outubro 2019). <http://www.linguateca.pt/ELD/aprClaudiaFreitasELD.pdf>
- FREITAS, C.; BIAR, L.; MARTINS, F. Construções do feminino na literatura brasileira: discurso, grandes corpora e leitura não-linear. *Revista Brasileira de Linguística Aplicada*, 2020, no prelo.
- GEERAERTS, D.; GRONDELAERS, S. Purism and fashion. French influence on Belgian and Netherlandic Dutch. *Belgian Journal of Linguistics*, n. 13, 1999, pp. 53-68.
- GEERAERTS, D.; GRONDELAERS, S.; BAKEMA, P. *The structure of lexical variation. Meaning, naming, and context*. Berlin/New York: Mouton de Gruyter, 1994.
- LAKOFF, G.; JOHNSON, M. *Metaphors We Live By*. Chicago & London: University of Chicago Press, 1980.
- LEITÃO DE ALMEIDA, M. L.; FERREIRA, R. G.; PINHEIRO, D. O. R.; SILVA, N. H. da; MARQUES, M. O. A hipótese de corporificação da categorização e do léxico. In: ALMEIDA, M. L. L. ET AL. (orgs.). *Linguística Cognitiva em foco: morfologia e semântica do português*. Rio de Janeiro: Publit, 2009: 187-204.

- SANTOS, D. Corpora at Linguateca: Vision and Roads Taken. In: BERBER SARDINHA, T.; FERREIRA, T. L. S. B. (eds.), *Working with Portuguese Corpora*. Bloomsbury, 2014: 219-236.
- SANTOS, D. Literature studies in *Literateca*: between digital humanities and corpus linguistics. In: DOERR, M.; EIDE, Ø.; GRØNVIK, O.; KJELSVIK, B. (eds.), *Humanists and the digital toolbox: In honour of Christian-Emil Smith Ore*. Oslo: Novus forlag, 2019^a: 89-109.
<https://www.linguateca.pt/Diana/download/DianaSantosFSCEO.pdf>
- SANTOS, D. Distant reading health: A pilot study on health and disease in lusophone literature. Presentation at *Illness and disability in literary and cultural texts: an international seminar* (Univ. of Oslo, 17 June 2019), 2019b.
<https://www.linguateca.pt/Diana/download/DRHealth.pdf>
- SANTOS, D.; SIMÕES, A. Towards a computational environment for studying literature in Portuguese. *DH Budapest 2019, Digital Humanities Conference* (Setembro de 2019).
<https://www.linguateca.pt/Diana/download/PresentationBudapestSantosSimoes.pdf>
- SANTOS, D.; SOARES DA SILVA, A.; MOTA, C. Guarda-fatos: notas sobre a anotação do campo semântico do vestuário em português. 21 de Fevereiro de 2011. [1.a versão: 26 de Outubro de 2009]
<https://www.linguateca.pt/acesso/GuardaFatos.pdf>
- SANTOS, D.; TAGNIN, S. E. O.; TEIXEIRA, E. D. CorTrad and Portuguese-English translation studies: investigating colours. In: EBELING, S. O.; EBELING, J.; HASSELGAARD, H. (eds.), *Aspects of corpus linguistics: compilation, annotation, analysis*. Helsinki: Research Unit for Variation, Contacts, and Change in English, 2012. e-ISSN: 1797-4453.
http://www.helsinki.fi/varieng/series/volumes/12/santos_tagnin_teixeira/
- SANTOS, D.; PIRES, E.; LOPES, J. M.; FUÃO, R. S.; FREITAS, C. Periodização automática: Estudos linguístico-estatísticos de literatura lusófona. *Linguamática*, n. 12.1, 2020, pp. 80-95.
<https://linguamatica.com/index.php/linguamatica/article/view/314/465>
- SOARES DA SILVA, A.; DUARTE, M. O léxico do vestuário no Português Europeu e no Português Brasileiro: convergência ou divergência? *Revista Portuguesa de Humanidades*, n. 9, pp. 117-136, 2005.
- SOARES DA SILVA, A. Integrando a variação social e métodos quantitativos na investigação sobre linguagem e cognição: para uma sociolinguística cognitiva do português europeu e brasileiro. *Revista de Estudos Linguísticos*, n. 16.1, jan./jun. 2008, Belo Horizonte, pp. 49-81.

<http://www.periodicos.letras.ufmg.br/index.php/relin/article/view/2480>

WITKOWSKI, S. R.; BROWN, C. H. Climate, clothing and body-part nomenclature. *Ethnology*, n. 24.3, 1985, pp. 197-214.

Recebido em: 29/05/2020

Aceito em: 22/09/2020

Publicado em janeiro de 2021