

## Detecção de linguagem tendenciosa em decisões judiciais<sup>1</sup>

Alexandra Guedes Pinto\*, Catarina Vaz Warrot\*, Henrique Lopes Cardoso\*\*, Isabel Margarida Duarte\*, Rui Sousa-Silva\*

\*FLUP / CLUP

\*\*FEUP / LIACC

### Abstract

The linguistic expression of subjectivity is a complex phenomenon that has been the object of reflection by several sub-areas of Linguistics and, more recently, of Computational Linguistics. Linguistic subjectivity, in terms of the linguistic expression of the speaker's opinions and attitudes, affects all levels of discourse organization and is present, to different degrees, in diverse textual genres. Subjectivity and bias are connected, in the sense that the presence of bias in discourse has been related, both in Linguistics and Computational Linguistics, to the occurrence of signs of subjectivity.

Court decisions are an argumentative text genre that may convey traces of subjectivity but should not be biased. As a discourse that represents the State's position on social matters, it should reflect the principle of Equality. Nonetheless, a preliminary analysis of cases of gender violence reveals that this is not always the case. The research proposed in this paper aims to study the linguistic formulations that convey subjectivity and bias in court decisions on gender violence against women. The goal is to develop a linguistic model to detect these instances of bias, with a future possibility of application in a tool for automatic detection of gender bias in discourse, fueled by Artificial Intelligence (AI) and Natural Language Processing (NLP) techniques. A *corpus* of court decisions on gender violence has been extracted from the public access database of Instituto de Gestão Financeira e Equipamentos da Justiça (IGFEJ), and has been subject to analysis. A set of examples has been compiled in the analytical section of this study, demonstrating the possibility of connecting certain linguistic features, such as mitigation and intensification mechanisms, evidential expressions and counter-argumentative movements, to the presence of subjectivity and bias in discourse.

**Keywords:** Subjectivity, Linguistic Bias, Court decisions, Discourse Analysis, Natural Language Processing

**Palavras-chave:** Subjetividade, Bias linguístico, Sentenças judiciais, Análise do Discurso, Processamento de Linguagem Natural

### 0. Introdução

A expressão linguística da subjetividade, que pode existir em alguns géneros textuais e não ser tolerada noutros, é um fenómeno complexo e que pode expressar pontos de vista tendenciosos (Kerbrat-Orecchioni, 1980; Marques, *et al.*, 2019; Menegatti *et al.*, 2017; Pinto, 2016; Pinto, 2019; Pinto, 2021; Recasens *et al.*, 2013). Estes pontos de vista tendenciosos têm sido estudados a partir de várias perspetivas linguísticas, algumas

---

<sup>1</sup> A presente investigação foi apoiada por fundos nacionais portugueses e por fundos comunitários europeus atribuídos pela Fundação para a Ciência e a Tecnologia (Portugal) ao Centro de Linguística da Universidade do Porto através do programa de financiamento FCT-UIDB/00022/2020.



das mais recentes relacionando-se com a Linguística Computacional, com foco no desenvolvimento de modelos de deteção automática do fenómeno (De-Arteaga *et al.*, 2019; Prabhakaran, 2019; Recasens *et al.*, 2013; Riloff, 2003). Estes estudos têm conduzido a diferentes aplicações, incluindo a análise de sentimento (Pang & Lee, 2008), a deteção de discursos radicalizados nas redes sociais (Sousa-Silva, 2018) e em notícias hiper-partidárias (Cruz *et al.*, 2020), a identificação de marcadores de autoria (Sousa-Silva *et al.*, 2011), entre outros.

A expressão do preconceito de género tornou-se uma questão social particularmente relevante, que os governos tentam enquadrar através de medidas políticas e de leis voltadas para uma maior igualdade entre géneros. No site oficial da Comissão Portuguesa para a Cidadania e Igualdade de Género (CIG)<sup>2</sup>, podemos ler que a igualdade entre mulheres e homens é um princípio fundamental da Constituição da República Portuguesa e a promoção dessa igualdade é um dever fundamental do Estado. Nesse contexto, foi publicada no Jornal Oficial a Estratégia Nacional para a Igualdade e Não Discriminação “Portugal + Igual”<sup>3</sup>.

Reconhecendo esses princípios como um requisito para um futuro sustentável, a CIG definiu um programa de ação até 2030. Com efeito, a integração da dimensão da discriminação de género na política surge como um princípio de boa governança, onde o discurso oficial desempenha um papel decisivo como força motriz da mudança social. As decisões judiciais são uma das formas desse discurso oficial, em que diversos tipos de enviesamento (*bias*<sup>4</sup>) persistem, incluindo o enviesamento de género. Algumas das decisões judiciais que patenteavam muito explicitamente preconceitos de género contra a mulher atraíram a atenção dos *media*, tendo sido amplamente divulgadas, mas muitas outras permanecem desconhecidas, contribuindo para a manutenção de preconceito e discriminação e, conseqüentemente, para uma administração inadequada da justiça.

Este artigo tem como objetivo contribuir para a construção de um modelo linguístico que permita identificar a presença de enviesamento de género contra a mulher em decisões judiciais, aumentando, deste modo, a consciência social para este problema. A partir deste modelo poder-se-á, em futuros estudos, analisar a sua possível aplicação à anotação de *corpora* de dimensões representativas e à produção de uma ferramenta computacional de deteção automática deste enviesamento.

Para este efeito, na Secção 1 deste artigo, efetuamos um breve enquadramento da questão; na Secção 2, delineamos uma metodologia de pesquisa para abordar o problema, apresentando alguns exemplos concretos; finalmente, na Secção 3, efetuamos um breve balanço da reflexão.<sup>5</sup>

## 1. Enquadramento

### 1.1. Subjetividade

Sendo o nosso estudo voltado para a expressão da subjetividade no discurso, é relevante abordar alguns aspetos deste conceito. A subjetividade, no seu sentido comum definida como “1. carácter do que é subjetivo; 2. campo ou domínio das atividades subjetivas; 3. interpretação de um sujeito individual; 4. característica de uma opinião ou atitude marcada por sentimentos, impressões ou preferências pessoais”<sup>6</sup>, mantém relações com a subjetividade linguística, propriamente dita, ligada à forma como o sujeito enunciador marca a sua presença

<sup>2</sup> <https://www.cig.gov.pt/>

<sup>3</sup> <https://www.portugal.gov.pt/pt/gc21/comunicacao/comunicado?i=portugal-igual-estrategia-nacional-para-a-igualdade-e-a-nao-discriminacao-2018-2030>

<sup>4</sup> Usaremos no nosso estudo a palavra *bias*, pelo facto de a mesma se ter generalizado nos estudos sobre a matéria, a par das alternativas portuguesas *discurso tendencioso*, *ponto de vista enviesado* e *enviesamento*, que assumimos como sinónimas parciais da primeira.

<sup>5</sup> Note-se que, neste estudo, usaremos as expressões “decisões judiciais” e “sentenças judiciais” para representar quer as decisões de Tribunais de Primeira Instância - as sentenças - quer as decisões de Tribunais de Segunda Instância - os acórdãos. Na análise prática, usaremos um *corpus* constituído apenas por acórdãos.

<sup>6</sup> Definição extraída da *Infopédia* consultada em:

<https://www.infopedia.pt/dicionarios/lingua-portuguesa/subjetividade> em 20-05-2021.



no enunciado. A expressão da subjetividade (em sentido comum) no discurso é potenciada por um conjunto de instâncias linguísticas (que Benveniste (1966) teorizou em torno da Enunciação e do “aparelho formal da enunciação”), que originam, por sua vez, a subjetividade linguística.

Estes elementos estão disponíveis nas línguas naturais, enquanto sistemas, e possibilitam a inscrição do sujeito empírico como sujeito enunciador, permitindo a conversão das línguas em discurso. A díade de pronomes pessoais *eu-tu* é um dos elementos mais importantes deste conjunto de formas, mas as possibilidades de marcação do sujeito empírico no enunciado são vastíssimas, recobrando toda a gramática das línguas naturais.

Ao constituir-se como sujeito enunciador, o locutor deixa, inevitavelmente, marcas da sua presença no discurso, isto é, marcas de *subjetividade*, na aceção específica que aqui pretendemos realçar - “la capacité du locuteur à se poser comme «sujet»” (Benveniste, 1966: 269) - e que foi desenvolvida por muitos autores subsequentes a Benveniste da área da Linguística da Enunciação e não só.

A subjetividade no discurso foi, desde sempre, um tema de interesse para os linguistas. Benveniste chamou a atenção para as mais básicas formas de inscrição do sujeito na linguagem. Os trabalhos sobre modalidade e modalização também problematizaram esta questão central, ao analisarem a manifestação de opiniões e de atitudes do enunciador no seu discurso (Oliveira, 2004; Oliveira & Mendes, 2013). Kerbrat-Orecchioni (1980) propõe o termo “subjectivèmes”, uma categoria que integra todas as instâncias gramaticais que reenviam para as marcas do sujeito no enunciado.

As diversas teorias sobre as funções da linguagem (do funcionalismo linguístico à pragmática) aceitam a existência de proposições em que a presença do sujeito é forte, dando nascimento à função emotiva e aos atos ilocutórios expressivos. Da mesma forma, no campo da linguística textual, a subjetividade foi confirmada como um índice que diferencia géneros. Mais recentemente, as abordagens de fenómenos de cortesia e atenuação e de intensificação (Albelda, 2005; Briz & Albelda, 2013), em vários modos (oral, escrito) e diversos tipos de discurso (científico, político) mostraram que os géneros textuais são vistos como formas privilegiadas de o sujeito construir o seu *ethos* e a sua relação com o interlocutor. A subjetividade está, por sua vez, intimamente relacionada com a argumentação, na medida em que muitos dos enunciados com fortes marcas de subjetividade são simultaneamente enunciados com um alto grau de argumentação (sendo o seu objetivo agir sobre o interlocutor).

Num estudo que se propõe abordar formas de subjetividade e enviesamento num determinado género textual, tornam-se relevantes todas as formas de subjetividade linguística, de marcação do sujeito enunciador no seu enunciado, enquanto vestígios de uma subjetividade empírica, vestígios de uma interpretação individual de um sujeito ou de uma opinião ou atitude marcada por sentimentos, impressões ou preferências pessoais<sup>7</sup>.

Compreendemos, pelo exposto acima, que a subjetividade linguística num dado discurso não se define em termos de presença/ausência, mas, sim, de uma maior ou menor presença, ou seja, em termos de grau de manifestação num dado discurso, e, que um modelo de deteção de formas de subjetividade linguística terá de ser sensível a dados contextuais, como, por exemplo, o do género textual e o do tipo de discurso em causa.

## 1.2. *Bias* ou enviesamento<sup>8</sup>

Nas palavras de Beukeboom e Burgers (2017: 2), a categorização social e os estereótipos desempenham um papel fundamental na perceção, no julgamento e na interação social. Embora os estereótipos sejam funcionais, permitindo-nos compreender o complexo ambiente social em que vivemos, o seu uso pode promover preconceitos e discriminação, quando os indivíduos são tratados com base em expectativas estereotipadas genéricas, em vez de informações individualizantes.

<sup>7</sup> Conforme a definição dicionarizada de *subjetividade* apresentada no início da subsecção 1.1.

<sup>8</sup> Ver o exposto na nota 2.



O preconceito e a discriminação surgem, pois, de associações estereotipadas generalizadas (negativas ou positivas) que os indivíduos assimilam sobre as categorias sociais: “Prejudice and discrimination generally refer to situations in which individuals are treated, described, and/or pre-judged based on such generic stereotypic expectancies and sentiments, rather than on available individuating information.” (Beukeboom & Burgers, 2017: 2).

O discurso é uma das formas privilegiadas de manutenção ou de contestação dos estereótipos, tornando-se, assim, relevantes os estudos sobre as suas manifestações discursivas: “Research on biased language use reveals the communicative and linguistic processes through which stereotypes are formed and maintained.” (*ibidem*).

Um “*bias* linguístico” é definido, nas palavras de Beukeboom e Burgers (2017: 3), como uma assimetria sistemática na escolha de palavras para a caracterização de um dado fenómeno social, escolha essa que reflete as categorizações coletivas acerca desse fenómeno<sup>9</sup>. Segundo os autores, o “*bias* linguístico” manifesta-se em três planos diferentes:

Three types of biases are distinguished in the literature that reveal, and thereby maintain, social category knowledge and stereotypes: (a) biases in labeling, (b) bias in what we communicate about, and (c) biases in how we formulate information about categorized individuals. (Beukeboom & Burgers, 2017:3)

Essas três formas relacionam-se, pois, com as etiquetas que usamos para categorizar os indivíduos ou os fenómenos; com as informações que comunicamos sobre eles – preferencialmente, informações consistentes com os estereótipos, ao invés de informações inconsistentes com os estereótipos; e, finalmente, com o modo como são formuladas as informações acerca dos indivíduos ou dos fenómenos categorizados. Segundo os autores, as diferenças de formulação (por exemplo, na abstração da linguagem, nas explicações, no uso de negações, na ironia) revelam, subtilmente, se o comportamento de uma categoria-alvo está de acordo com o estereótipo ou não, sendo que a informação consistente com o estereótipo tende a ser formulada de forma diferente da informação inconsistente com o estereótipo (*ibidem*: 5)

A presença de *bias* ou ponto de vista tendencioso no discurso tem sido relacionada, tanto em Linguística como em Linguística Computacional, com a ocorrência de indícios de subjetividade. Isso mesmo nos dizem Recasens *et al.* (2013), num estudo em que trabalham a presença de *bias* em textos da Wikipédia, pertencentes, pois, ao tipo de discurso de divulgação científica:

Bias is linked to the lexical and grammatical cues identified by the literature on subjectivity (Wiebe *et al.*, 2004; Lin *et al.*, 2011), sentiment (Liu *et al.*, 2005; Turney, 2002), and especially stance or “arguing subjectivity” (Lin *et al.*, 2006; Somasundaran and Wiebe, 2010; Yano *et al.*, 2010; Park *et al.*, 2011; Conrad *et al.*, 2012). For example, like stance, framing bias is realized when the writer of a text takes a particular position on a controversial topic and uses its metaphors and vocabulary. But unlike the product reviews or debate articles that overtly use subjective language, editors in Wikipedia are actively trying to avoid bias, and hence biases may appear more subtly. (Recasens *et al.*, 2013: 1650-1651)

Assim, embora as pistas relativas à subjetividade linguística possam ajudar a detetar a presença de *bias* ou enviesamento, atendendo a que este tende a ocorrer em sequências em que o sujeito enunciador assume um determinado posicionamento acerca de um determinado fenómeno, não é válido presumir que todos os excertos textuais com indícios de subjetividade, ainda que num grau mais marcado da mesma, em que muitos indícios se acumulam numa janela curta de texto, pressupõem a existência de *bias* ou enviesamento. A relação entre os

<sup>9</sup> Tradução nossa, correspondente ao original: “a systematic asymmetry in word choice that reflects the social category cognitions that are applied to the described group or individual(s)”.



dois fenómenos poderá não corresponder, pois, a uma implicação bilateral, no sentido em que a presença de *bias* ou enviesamento normalmente estará associada à presença de subjetividade, mas a presença de subjetividade pode não implicar a presença de *bias* ou enviesamento.

### 1.3. Subjetividade e *Bias* no PLN e na Linguística Computacional

Recentemente, assistiu-se a um interesse crescente da Linguística Computacional e do PLN (Processamento de Linguagem Natural) pela construção da subjetividade no discurso, trazendo o desenvolvimento de modelos para a sua análise automática. Esses modelos provaram ser particularmente relevantes no tratamento de grandes volumes de dados, como os gerados na web, possibilitando detetar automaticamente a subjetividade nos textos, com a subsequente extração, classificação e sumarização desta informação.

Construções linguísticas e sociais como a expressão de sentimentos (Wiebe *et al.*, 2004), da subjetividade (Riloff & Wiebe, 2003), de opiniões (Pang & Lee, 2008), de dados factuais, de crenças, de cortesia, de respeito, de poder (Prabhakaran *et al.*, 2019), de enviesamento/*bias* (De-Arteaga *et al.*, 2019; Menegatti *et al.*, 2017; Recasens *et al.*, 2013), entre outros, têm sido estudadas em pesquisas computacionais com resultados muito interessantes e úteis.<sup>10</sup>

Estes estudos mostraram-se promissores quanto à possibilidade de associação regular entre determinadas construções linguísticas, subjetividade marcada e enviesamento/*bias* potencial. Entendendo a subjetividade de uma forma integrada, como a expressão de opiniões, avaliações, sentimentos e especulações em linguagem natural – veja-se a proposta de Wiebe *et al.* (2004: 2) “Subjectivity in natural language refers to aspects of language used to express opinions, evaluations, and speculations” - os estudos em Linguística Computacional, que focalizam esta questão, almejam construir um modelo que seja capaz de detetar automaticamente a expressão de subjetividade nos textos. O isolamento de pistas linguísticas, tais como palavras (os chamados léxicos de sentimento) e categorias gramaticais é o caminho utilizado, a par do cruzamento com outras estratégias, como por exemplo, a densidade de ocorrência de pistas linguísticas de subjetividade num dado cotexto, como salientam Wiebe *et al.* (2004: 277):

Clues of subjectivity are generated and tested, including low-frequency words, collocations, and adjectives and verbs identified using distributional similarity. The features are also examined working together in concert (...) the density of subjectivity clues in the surrounding context strongly affects how likely it is that a word is subjective.

De facto, a gramaticalização de opiniões e de atitudes do falante<sup>11</sup> estende-se a todos os níveis linguísticos (fonológico, morfológico, lexical, sintático, semântico, discursivo) e apenas um modelo integrado pode responder ao desafio de detetar manualmente ou automaticamente a presença da subjetividade e de *bias* ou enviesamento no discurso.

Recasens *et al.* (2013), comparando várias reedições humanas de um mesmo artigo na Wikipedia, constataram a presença de duas classes de *bias*: uma classe que os autores intitularam de *framing bias*, que acontece quando os artigos revelam a existência de atos expressivos de elogio ou de palavras que reenviam para um determinado ponto de vista e, uma outra classe, que os autores nomearam de *epistemological bias*, que

<sup>10</sup> Não podemos, no escopo deste artigo, efetuar a recensão dos vários estudos mencionados, que contribuindo parcelarmente para a constituição de um corpo teórico e experimental na área dos estudos sobre subjetividade e *bias* ou enviesamento em Linguística Computacional, utilizam, cada um deles, metodologias e *corpora* diferentes. Focalizaremos, na presente secção, alguns dos estudos que foram particularmente influentes na nossa análise.

<sup>11</sup> Assumimos, neste contexto, a aceção de *gramaticalização* conforme Oliveira (2003: pp. 245-247).



acontece quando, por exemplo, se apresentam proposições controversas como proposições genericamente aceites e verdadeiras. Várias pistas linguísticas foram identificadas como indiciadoras destes dois tipos de *bias* neste género de divulgação científica que os autores situam dentro do discurso científico, nomeadamente determinados tipos de verbos e de intensificadores. A possibilidade aventada pela pesquisa de encontrar uma associação sistemática entre determinadas categorias e construções linguísticas e a presença de subjetividade marcada e *bias* ou enviesamento potencial é o aspeto que pode favorecer um tratamento automático desta questão, viabilizado pelos estudos em Linguística Computacional, em PLN e em aprendizagem computacional.

Inúmeros trabalhos desta área da Linguística Computacional têm focalizado a polarização semântica de material lexical, associando de forma sistemática e automática os nomes, por exemplo, a sentimentos positivos ou negativos. É o caso de Rillof e Wiebe (2003) e Wiebe *et al.* (2004).

Todavia a complexidade do fenómeno é reconhecida por todos os que produzem estudos nesta área de interface. Nas palavras de Balahur *et al.*:

Although the subjectivity and sentiment analysis research fields have been highly dynamic in the past years, dealing with subjectivity and sentiment in text has proven to be a complex, interdisciplinary problem that remains far from being solved. Its challenges include the need to address the issue from different perspectives and at different levels, depending on the characteristics of the textual genre, the language(s) treated and the final application for which the analysis is done. (Balahur *et al.*, 2011: iii)

Um dos aparatos científicos para o estudo da construção linguística da subjetividade é a área da modalidade, situada na interface entre a Semântica e a Pragmática. Oliveira e Mendes (2013: 623) definem a modalidade como: “a forma de exprimir, por meios linguísticos, atitudes e opiniões dos falantes sobre o conteúdo proposicional dos enunciados que produzem”; Werner *et al.* (2015), num estudo em que se propõem anotar, em *corpora*, ocorrências de envolvimento epistémico, no sentido apresentado pelos autores de “crença do locutor numa dada proposição”, salientam o carácter desafiante do fenómeno, tanto do ponto de vista linguístico, como do ponto de vista do processamento de linguagem natural:

Level of committed belief is a modality in natural language, it expresses a speaker/writers belief in a proposition. Initial work exploring this phenomenon in the literature both from a linguistic and computational modeling perspective shows that it is a challenging phenomenon to capture, yet of great interest to several downstream NLP applications. (Werner *et al.*, 2015: 32)

A passagem das descrições qualitativas destes fenómenos, em *corpora* pequenos, para descrições quantitativas, que possam alimentar uma análise automática, é uma passagem desafiante, mas promissora. A impossibilidade de relacionar, em abstrato, uma determinada forma linguística com a expressão da subjetividade, e de um posicionamento potencialmente enviesado, dificulta esta passagem. Para além disto, as condicionantes pragmáticas e contextuais do discurso são relevantes na determinação dos valores que as formas assumem. Tal como indicam Kiesling *et al.* (2018: 709): “Furthermore, the stancetaking properties of individual words and phrases are always shaped by the pragmatic context, including discourse and (especially in social media) extralinguistic factors.”

Neste contexto, apenas um volume significativo de anotações manuais de *corpora* relativas a fenómenos de subjetividade e *bias* ou enviesamento pode potenciar e validar eventuais abordagens automáticas da questão.

As sentenças e acórdãos judiciais são um género pouco (ou nada) tratado do ponto de vista refletido acima. Se, por um lado, a subjetividade é inerente ao género, que possui uma natureza argumentativa essencial, a expressão de *bias* ou enviesamento não é esperada nem deve ser aceitável. Este tipo de discurso exige um



modelo explicativo que difere de outros já construídos para outros tipos de discurso e de géneros de texto, como artigos de opinião ou publicações nas redes sociais, onde a expressão de opinião pode ser uma marca de género, como acontece, por exemplo, nas interações polémicas ocorrentes nas redes, sendo a expressão de pontos de vista tendenciosos mais expectável. Apenas a integração de vários níveis linguísticos pode ajudar na deteção e análise do *bias* ou enviesamento nas decisões judiciais: o nível lexical, através de um léxico específico presente nas decisões judiciais (no caso ocorrente, nas decisões judiciais de violência de género contra a mulher), como outros autores fizeram para outros géneros de texto (Wiebe *et al.*, 2004); o nível morfológico, incluindo a estrutura interna das palavras e a sua modificação, minimizando e intensificando expressões; o nível sintático-semântico, nos casos de predicados ativadores de implícitos (Karttunen, 1971; Kiparsky, 1970) e, finalmente, o nível pragmático e discursivo, envolvendo a expressão de modalidades, de atos ilocutórios, de marcas de argumentação e de movimentos argumentativos, entre outros.

#### 1.4. O *bias* ou enviesamento de género

Questões relativas ao *bias* ou enviesamento de género têm sido exploradas a partir de diferentes quadros teóricos em diferentes áreas das ciências sociais. A categoria *género* foi desenvolvida pelas teorias do feminismo contemporâneo sob a perspectiva de compreender e responder, dentro de parâmetros científicos, à situação de desigualdade entre os sexos e ao modo como esta questão opera na realidade e interfere no conjunto das relações sociais. Varikas (1989: 28) afirma que, ao tomar o termo emprestado da gramática e da linguagem, as feministas postularam a necessidade de separar o sexo biológico, mais ou menos dado pela natureza, do sexo social, produto de uma construção social permanente, que forma em cada sociedade humana, a organização das relações entre os homens e as mulheres.

No glossário disponível no site da Comissão para a Igualdade no Trabalho e no Emprego<sup>12</sup>, podemos ler que o conceito de género se refere às representações (conjunto de crenças e ideias feitas) em torno do sexo que são geradas em sociedade. Com efeito, os seres humanos só se constroem como tal em relação com os outros. Saffioti (1992: 210) considera que não se trata de perceber apenas corpos que entram em relação com o outro. É a totalidade formada pelo corpo, pelo intelecto, pela emoção, pelo carácter do EU, que entra em relação com o outro.

No mesmo glossário referido acima, (p.299) menciona-se que, quando consideramos a forma como se estabelecem as diferenças entre os sexos, se verifica que a diferença feminina se estende ao coletivo das mulheres (sejam elas ricas ou pobres, estéreis ou férteis, elas serão vistas sobretudo como mães e como seres socialmente dependentes) enquanto que os homens, pelo contrário, podem ser vistos como diferentes entre si (por razões de classe, de cor da pele, de nacionalidade, etc.). Por isso, também, a feminilidade (significados associados ao sexo feminino) é mais normativa para as mulheres, do que a masculinidade é para os homens, delimitando os contextos que são “próprios” às mulheres (a casa, a esfera privada, etc.) e recordando-lhes, constantemente, “como devem comportar-se” para “serem femininas”.

As considerações sobre género relacionam-se também com conceitos tais como: *diferença de género*, *igualdade de género*, *desigualdade de género*. Segundo o relatório apontado<sup>13</sup>, a *diferença de género* corresponde ao reconhecimento da existência de valores constitutivos da identidade feminina e da identidade masculina que devem estar presentes e manifestar-se em igualdade em todas as esferas e dimensões da vida. (p.303); a *igualdade de género* diz respeito à igualdade entre as mulheres e os homens, materializando-se na ausência de assimetrias entre umas e outros em todos os indicadores relativos à organização social, ao exercício de direitos e de responsabilidades, à autonomia individual e ao bem-estar. Pressupõe o reconhecimento do igual

<sup>12</sup> [http://cite.gov.pt/Formar\\_Iguald/PDFs\\_Manual/M07\\_Glossario.pdf](http://cite.gov.pt/Formar_Iguald/PDFs_Manual/M07_Glossario.pdf)

<sup>13</sup> Ver acima, nota de rodapé 6.



valor social das mulheres e dos homens e do respetivo estatuto na sociedade e implica a participação equilibrada de homens e mulheres em todas as esferas da vida, incluindo a participação económica, política, social e na vida familiar, sem interditos nem barreiras em razão do sexo. Este conceito sublinha a liberdade que todos os seres humanos têm de desenvolver as suas capacidades e de fazer as suas escolhas sem as limitações impostas pelos papéis sociais de género e considera, valoriza e trata os comportamentos, aspirações e necessidades das mulheres e dos homens de igual forma. Por fim, a *desigualdade de género* pressupõe a existência de uma hierarquia de direitos, estatuto e dignidade entre as mulheres e os homens, quer a nível da lei quer a nível dos factos. Remete igualmente para as assimetrias nos indicadores sociais entre a situação das mulheres e dos homens, tanto na esfera pública como na esfera privada; a discriminação em função do sexo ou do género e o prejuízo de qualquer natureza decorrente de normas jurídicas, práticas sociais ou comportamentos individuais que é sofrido por uma pessoa em função do sexo ou em função do género. (p. 303).

A resolução 48/104, de 20 de dezembro de 1993 sobre a eliminação da violência contra as mulheres, proclamada pela Assembleia Geral das Nações Unidas, reconhece o seguinte:

(...) a violência contra as mulheres constitui uma manifestação de relações de poder historicamente desiguais entre homens e mulheres, que conduziram ao domínio e à discriminação das mulheres por parte dos homens e impediram o progresso pleno das mulheres, e que a violência contra as mulheres constitui um dos mecanismos sociais fundamentais através dos quais as mulheres são forçadas a assumir uma posição de subordinação em relação aos homens.<sup>14</sup>

A nossa proposta, neste estudo, adota uma perspetiva discursiva, ao analisar a identidade de género como uma construção histórico-discursiva e o discurso como uma forma privilegiada de representação de estereótipos. As categorizações sociais e os estereótipos, que podem promover a discriminação e o preconceito, são sustentadas através do discurso, e surgem (muitas vezes de forma subtil e implícita) em “preconceitos linguísticos” (Beukeboom & Burgers, 2017), no sentido de manifestações de preconceitos em estruturas linguísticas que são destacadas no presente estudo.

Partindo de uma perspetiva linguística, a discriminação de género já foi abordada em trabalhos anteriores dos autores, tanto no que diz respeito a decisões judiciais (Pinto, 2019; Pinto, 2021), como no discurso mediático (Marques *et al.*, 2019), a par de outras formas de enviesamento (Pinto, 2016).

## 2. Metodologia e Análise de acórdãos

Em Portugal, o Instituto de Gestão Financeira e Equipamentos da Justiça (IGFEJ) disponibiliza uma base de dados de decisões judiciais de acesso público, de onde é possível extrair acórdãos segmentados por vários critérios de busca<sup>15</sup>. Pesquisando neste banco de dados temas relacionados com a violência de género contra a mulher, é possível constituir um *corpus* relevante para uma análise focada nos fenómenos indutores deste *bias* ou enviesamento.<sup>16</sup>

<sup>14</sup> Resolução disponível para consulta em: <http://gddc.ministeriopublico.pt/sites/default/files/declaracaoviolenciamulheres.pdf>

<sup>15</sup> A Base de Dados Jurídico-Documentais está disponível para consulta em: <http://www.dgsi.pt/>

<sup>16</sup> A pesquisa na base de dados Jurídico-Documentais, cujo número total de acórdãos está em permanente atualização, pode ser executada a partir de critérios diversos: por termos, por campo, por descritor e ainda por “pesquisa livre”. Na recolha para o presente estudo, utilizamos descritores já disponíveis no motor de busca, a saber “violência doméstica”; “abuso sexual”; “coerção sexual”; “violação”; “maus tratos”, uma vez que estes descritores, já predefinidos na plataforma, configuram categorias estereotipadas de o sistema judicial catalogar o tipo de crimes em causa. Outros descritores poderiam (e poderão vir a) ser elegíveis para a recolha de *corpus* relevante. No presente estudo, meramente exploratório, os descritores utilizados revelaram-se produtivos na devolução de casos relacionados com o tema da violência de género contra a mulher. O descritor “violência de género” não se encontra ainda disponível como categoria de pesquisa nas decisões judiciais compiladas na referida base de dados.



No presente estudo, de natureza exploratória, apresentaremos algumas ocorrências textuais, extraídas de acórdãos disponíveis na Base de Dados acima referenciada, que comprovam a existência de pistas de subjetividade e *bias* ou enviesamento neste género de texto. Procederemos, assim, à categorização linguística das ocorrências e ao respetivo comentário.

A consideração de um *corpus* de maior dimensão e a sua anotação, segundo fenómenos linguísticos indutores de *bias* ou enviesamento, poderia, eventualmente, permitir a passagem de uma análise qualitativa, aplicada a casos isolados, para uma análise quantitativa, aplicada a *corpora* de dimensões significativas. Esta transposição da análise para uma escala relevante poderia viabilizar, por sua vez, a ativação do PLN e de técnicas de aprendizagem computacional, para desenvolver uma ferramenta de deteção automática, não supervisionada, de pontos de vista tendenciosos em decisões judiciais.

No presente estudo não cabe, todavia, a exposição dos procedimentos que permitiriam o alargamento da pesquisa qualitativa para a pesquisa quantitativa, ficando esta possibilidade para eventuais estudos futuros.

## 2.1. Constituição do *corpus*

Para seleccionar sentenças judiciais da Base de Dados referida, em que as questões de género surjam mais claramente, certos indicadores de pesquisa podem ser utilizados, tais como “violência doméstica”, “violação”, “maus tratos”, “coerção sexual” ou “abuso sexual”.<sup>17</sup> A questão da proteção de dados é salvaguardada pelo facto de as decisões do tribunal já terem sido expurgadas de quaisquer dados pessoais antes da sua inclusão no banco de dados. Além disso, certas partes das decisões judiciais são particularmente relevantes em termos de expressão da subjetividade e do *bias* ou enviesamento, tais como a «indicação dos fatos provados e não provados» e a «fundamentação». A primeira visa expor os fatos comprovados, enquanto a última visa enquadrar e justificar a decisão do juiz ou do painel de juízes.

O *corpus* deste estudo inclui apenas excertos de decisões de Tribunais de Segunda Instância, pelo facto de serem estas as decisões disponíveis na Base de Dados em questão. Os excertos foram extraídos de cinco acórdãos que cobrem um intervalo temporal entre 2007 e 2018 e provêm do Tribunal da Relação do Porto, do Tribunal da Relação de Lisboa e do Supremo Tribunal de Justiça.

A recolha dos exemplos corresponde a uma amostragem não-casual e por conveniência.<sup>18</sup> Os resultados não se pretendem representativos, mas, tão somente, ilustrativos, uma vez que a seleção de casos obedeceu a critérios de relevância para o estudo.

## 2.2. Categorias de análise

A estabilização de um modelo integrando fenómenos linguísticos indutores de subjetividade e *bias* ou enviesamento é crucial para esta pesquisa. A necessidade de identificar linguagem tendenciosa incorporada em estruturas discursivas sofisticadas como as usadas em decisões judiciais, onde se espera que os juízes desenvolvam movimentos argumentativos, implica um conjunto de categorias linguísticas ajustadas especificamente a este género de texto. Outras taxonomias foram utilizadas como ponto de partida, nomeadamente as relativas à subjetividade (Kerbrat-Orecchioni, 1980), à atenuação (Briz & Albelda, 2013; Hyland, 1998), à intensificação (Albelda, 2005), assim como trabalhos anteriores sobre a análise de sentimentos,

<sup>17</sup> Não está excluída a hipótese de procurar material relevante nesta Base de Dados, através de outros termos de pesquisa, que possam gerar resultados igualmente interessantes. Aquando da apresentação deste trabalho no ENAPL recebemos a sugestão de ensaiar termos de pesquisa como “divórcio”, “tutela parental”, entre outros, que gostaríamos de explorar em estudos futuros.

<sup>18</sup> Reforçamos que, neste estudo exploratório, utilizamos apenas acórdãos em que, depois de uma análise prévia dos mesmos, confirmamos a existência de estruturas linguísticas que servem a manifestação de juízos subjetivos e preconceitos, facto pelo qual designamos a amostra como não-casual e por conveniência, ou seja, selecionada expressamente para comprovar que certas estruturas linguísticas se vocacionam para a expressão de pontos de vista tendenciosos nas decisões judiciais.



da subjetividade e a deteção de enviesamento/*bias* em Linguística Computacional (Recasens *et al.*, 2013; Riloff *et al.*, 2003; Wiebe *et al.*, 2004).

Neste estudo particularizaremos apenas as seguintes instâncias:

- Intensificação através de quantificadores e outros recursos;
- Atenuação através de minimizadores e movimentos argumentativos concessivos contra-argumentativos;
- Elevação da força modal epistémica através de construções de evidencialidade;
- Elevação da força modal apreciativa/avaliativa através de léxico de polaridade semântica negativa.

### 2.3. Análise de ocorrências

Da análise já empreendida a alguns acórdãos noutros estudos, foi possível detetar a contribuição de vários recursos linguístico-discursivos para a construção de sequências potencialmente tendenciosas. Nessas sequências, mecanismos como a acumulação de conectores argumentativos, a presença de atos assertivos avaliativos e de atos expressivos, com um léxico de polaridade semântica marcada, a par de uma distribuição recorrente de papéis semânticos, como os de agente e de paciente, aos mesmos atores sociais, entre outros fenómenos, parecem desempenhar um papel preponderante.<sup>19</sup>

Para a deteção do enviesamento linguístico, podemos considerar o processo de análise como uma sequência de duas etapas: (i) a identificação de palavras ou sequências potencialmente tendenciosas; e (ii) a inserção dessas palavras ou sequências na respetiva categoria linguística. A lista de exemplos que abaixo transcrevemos, referenciados aos acórdãos de onde foram extraídos, e organizados de acordo com o fenómeno linguístico a relevar, corresponde às etapas (i) e (ii), a que acrescentamos, no caso deste estudo, uma etapa (iii), com um breve comentário sobre o contributo dos mecanismos linguísticos destacados para a veiculação da subjetividade e do *bias* ou enviesamento potencial.

#### 2.3.1 Exemplos e categorização

##### 2.3.1.1 Intensificação através de quantificadores e outros recursos

- (1) *Percebe-se facilmente (...) que os factos descritos nos n.ºs 1 a 3 não têm **qualquer** significado para a medida da culpa dos arguidos. (TRP-Processo 355/15.2 GAFLG.P1 de 11 out. 2017)*<sup>20</sup>
- (2) *Não merece, pois, **qualquer** reparo (...) o juízo probatório e valorativo efectuado pelo tribunal. (TRP-Processo 353/17.1SLPRT. P1 de 31 out. 2018)*
- (3) *A opção do tribunal (...) não é, por si só, merecedora de **qualquer** reparo ou crítica. (TRP-Processo 353/17.1SLPRT. P1 de 31 out. 2018)*
- (4) *(...) embora negasse conhecimento de **qualquer** interesse afetivo, mútuo, ou de outra ordem, entre os dois, não adiantando **qualquer** outro facto que nos permitisse, em conjunto com aquele, retirar **qualquer** tipo de conclusão pertinente. (TRP-Processo 3897/16.9JAPRT.P1 de 27 jun. 2018)*
- (5) *(...) atento o facto de não haver notícia da posteriores deslocações da ofendida ao “F...” ou de **qualquer** tipo de contacto entre ela e os arguidos, com núcleos de vida familiar e profissional perfeitamente afastados. (TRP-Processo 3897/16.9JAPRT.P1 de 27 jun. 2018)*
- (6) *Os arguidos não têm **qualquer** percurso criminal. (TRP-Processo 3897/16.9JAPRT.P1 de 27 jun. 2018)*

<sup>19</sup> Cf. Pinto, 2019 e Pinto, 2021, pp. 27-46.

<sup>20</sup> Os acórdãos encontram-se referenciados quanto ao tribunal de que provêm, à identificação processual e à data.



- (7) *Ao arguido não é conhecida **qualquer** condenação criminal, seja por factos anteriores ao crime em causa, seja por factos posteriores, sendo que tem a sua vida familiar e profissional estabilizadas.* (TRP-Processo 07 14613 de 7 nov. 2007)
- (8) *Mas não se provou que dele resultasse para a ofendida **qualquer** sofrimento agudo, sendo que não só se não apurou o sentido deste “marcar-te a cara”, como ela não viu nessa expressão **qualquer** ameaça de mal físico (...).* (TRP-Processo 07 14613 de 7 nov. 2007)
- (9) *O quadro traçado na acusação **está longe, muito longe mesmo, de corresponder** à realidade dos factos provados.* (TRP-Processo 353/17.ISLPRT. P1 de 31 out. 2018)
- (10) *Este caso de maus tratos **está longe de ser** dos mais graves que surgem nos tribunais.* (TRP-Processo 353/17.ISLPRT. P1 de 31 out. 2018)
- (11) *Este caso **está longe de ter** a gravidade com que, geralmente, se apresentam os casos de maus tratos no quadro da violência doméstica.* (TRP-Processo 355/15.2 GAFLG.P1 de 11 out. 2017)

### 2.3.1.2. Atenuação através de minimizadores e movimentos argumentativos concessivos contra-argumentativos <sup>21</sup>

- (12) (...) os factos, apreciados na sua globalidade, **não revelam uma carga de ilicitude particularmente acentuada**, confinando-se àquilo que é a situação mais comum no quadro geral da violência doméstica. (TRP-Processo 353/17.ISLPRT. P1 de 31 out. 2018)
- (13) *O arguido **não revela graves problemas** de inserção social.* (TRP-Processo 353/17.ISLPRT. P1 de 31 out. 2018)
- (14) (...) **não houve dolo muito intenso**, pois o arguido estava em alguma medida sob o efeito de bebidas alcoólicas. (STJ-Processo 195/14.6JAPDL.S1 de 13 de jul. 2017)
- (15) *Em função desses fatores a culpa situa-se em patamar **um pouco** acima da média, permitindo que a pena se fixe na zona intermédia da moldura penal.”* (STJ-Processo 195/14.6JAPDL.S1 de 13 de jul. 2017)
- (16) *O grau de ilicitude do facto pode considerar-se **um pouco** acima da média (...).* (TRP-Processo 07 14613 de 7 nov. 2007)
- (17) (...) em relação a cada um dos acusados, **não se evidenciam particulares factores de risco** associados a problemas comportamentais (TRP-Processo 3897/16.9JAPRT.P1 de 27 jun. 2018)
- (18) *os danos físicos provocados **não assumem especial gravidade** considerando o período de cura das lesões provocadas* (TRP-Processo 3897/16.9JAPRT.P1 de 27 jun. 2018)
- (19) *A leitura dos factos espelha personalidades com **escassíssimo** pendor para a reincidência* (TRP-Processo 3897/16.9JAPRT.P1 de 27 jun. 2018)
- (20) *Com efeito, a violação, envolveu um acto de cópula e outro de coito anal, ocorridos no interior de um automóvel, **sem violência particularmente grave**. O arguido **limitou-se a intimidar** a ofendida, dizendo-lhe que se resistisse “seria pior”, para além de, fazendo **uso moderado** da força física, ir vencendo as **pequenas resistências** que ela, não obstante a ameaça, foi opondo.* (TRP-Processo 0714613 de 7 nov. 2007)
- (21) (...) *efeito das bebidas alcoólicas diminui a sua capacidade de avaliar a gravidade dos seus actos, pelo que a culpa, **não obstante** ser na forma de dolo, **não atinge patamares elevados**.* (STJ-Processo 195/14.6JAPDL.S1 de 13 de jul. 2017)

<sup>21</sup> Relevamos que alguns dos exemplos incluídos na categoria 2.3.1.2 incluem também fenómenos de quantificação ilustrados na categoria 2.3.1.1. Por outro lado, as construções concessivas são consideradas na categoria 2.3.1.2, enquanto instâncias de atenuação, segundo a tipologia de mecanismos de atenuação de Briz e Albelda (2013).



- (22) (...) **apesar da censurabilidade das suas condutas**, os danos físicos provocados **não assumem especial gravidade** considerando o período de cura das lesões provocadas (TRP-Processo 3897/16.9JAPRT.P1 de 27 jun. 2018)

### 2.3.1.3. Elevação da força modal epistémica através de estruturas de evidencialidade

- (23) **É óbvio que**, se o arguido foi internado devido ao seu estado de depressão, esta teria que ser profunda. (TRP-Processo 355/15.2 GAFLG.P1 de 11 out. 2017)
- (24) **Tal como é de primeira evidência** que uma depressão não se cura de um dia para o outro. (TRP-Processo 355/15.2 GAFLG.P1 de 11 out. 2017)
- (25) (...) **a verdade é que**, nos últimos tempos, têm-se acentuado os sinais de uma tendência de sentido contrário, em que a mais banal discussão ou desavença entre marido/companheiro/namorado e mulher/companheira/namorada é logo considerada violência doméstica. (TRP-Processo 353/17.ISLPRT.P1 de 31 out. 2018)

### 2.3.1.4. Elevação da força modal apreciativa/avaliativa através de léxico de polaridade semântica negativa

- (26) o **adultério** da mulher é um **gravíssimo atentado** à honra e dignidade do homem. **Sociedades existem em que a mulher adúltera é alvo de lapidação até à morte**. (TRP-Processo 355/15.2 GAFLG.P1 de 11 out. 2017)
- (27) Foi a **deslealdade** e a **imoralidade** sexual da assistente que fez o arguido X cair em profunda **depressão** e foi nesse estado **depressivo** e **toldado** pela **revolta** que praticou o **acto de agressão**, como bem se considerou na sentença recorrida (TRP-Processo 355/15.2 GAFLG.P1 de 11 out. 2017)<sup>22</sup>

## 2.3.2. Comentário

Como referimos na secção 2.1 acima, relativa à constituição do *corpus*, a amostragem neste estudo foi não casual, na medida em que apenas recolhemos sequências textuais de um número limitado de acórdãos que poderiam ser portadoras de *bias* ou enviesamento, neste caso específico, de enviesamento de género contra a mulher.

A opção por procurar manifestações deste tipo de *bias* ou enviesamento em acórdãos de violência contra a mulher justifica-se devido à prevalência de determinados estereótipos (*bias* sociais) que parecem tender para a naturalização de certas formas de violência; para a desculpabilização do seu exercício e, muitas vezes, para a responsabilização da mulher pela violência exercida sobre ela. Estes *biases* ou enviesamentos, correspondendo a categorizações coletivas de fenómenos sociais e dos seus atores, são, muitas vezes, perpetuados pela *voz* do Estado, que, ao mesmo tempo que defende e cria comissões para a promoção da Igualdade de género (ver acima secção introdutória), reconhecendo que a assunção de políticas ativas para a Igualdade social é um princípio de boa governança, permite que o poder judicial divulgue acórdãos onde fala uma *voz* altamente tendenciosa quanto a estas questões, que vulgariza e atenua os atos de violência exercidos contra a mulher, culpabilizando-a pelas agressões sofridas.

Recorremos às palavras de Nascimento (2017:9) para sintetizar alguns dos *bias* ou pontos de vista tendenciosos, que poluem a avaliação de casos de violência contra a mulher (nomeadamente casos de violação):

<sup>22</sup> No exemplo (27) destaca-se também a construção avaliativa: “como **bem** se considerou na sentença recorrida”.



Comportamentos comumente associados à Cultura do Estupro incluem a culpabilização da vítima; a objetificação sexual da mulher; a segregação de género; a crença em mitos do estupro, ou seja, crenças estereotipadas e preconceituosas sobre as razões de estupros, estupradores e vítimas; a descrença na voz da vítima; a trivialização do estupro ou a negação deste; a recusa em reconhecer o dano causado à vítima por algumas formas de violência sexual; a apatia das instituições ao lidar com o crime, ou a combinação entre esses comportamentos. (Nascimento, 2017:9)

Nas ocorrências reunidas em 2.3.1.1, confirmamos o papel relevante do quantificador universal “qualquer” na argumentação dos juízes, já que a sua essência semântica conduz a anular outras hipóteses de interpretação das proposições para além da defendida por estes. Serve assim como um mecanismo de intensificação que, nos casos listados de (1) a (8), contribui para a minimização da culpa dos arguidos e para a minimização do dolo infligido na vítima. As ocorrências relatadas de (9) a (11), exemplificam o uso da construção “estar longe de V”<sup>23</sup> que, marcando uma forma sintático-semântica enfática, comparativamente com a negação simples – cf. “está longe de ter” vs “não tem” – atua também como mecanismo de intensificação da asserção, que, nos contextos referidos, contribui para minimizar a gravidade das situações em julgamento ou descredibilizar a voz da vítima ou do magistrado recorrente.

As ocorrências reunidas em 2.3.1.2 ilustram a atuação de vários mecanismos de atenuação. São, na sua maioria, atenuadores, que, recuperando as propostas de Briz e Albelda (2013), podemos caracterizar como modificadores morfológicos externos, quantificadores minimizadores, aproximativos ou difusores do significado proposicional ou extraproposicional. Assim, no conjunto de (12) a (20), temos exemplos como “(não) particularmente acentuada”, “(não) muito intenso”, “(não) graves problemas”, “(não) particulares fatores de risco”, “(não) especial gravidade”, “(não) particularmente grave”<sup>24</sup>, “um pouco acima da média”, “uso moderado”, “pequenas resistências”. Estes atenuadores contribuem para a mitigação da gravidade do ato. As construções concessivas/contra-argumentativas exemplificadas em (21) e (22) cumprem também a mesma função de mitigação da gravidade do ato, na medida em que executam um movimento argumentativo composto de concessão, em que os juízes reconhecem alguma gravidade na violência exercida, imediatamente seguido de contra-argumentação, em que apresentam um argumento mais forte para a conclusão pretendida, sendo este argumento que prevalece.

As instâncias em 2.3.1.3 exemplificam a elevação da força das asserções por recurso à voz da evidência. A voz da evidência é uma voz a que se atribui relevância e credibilidade social. Quando o juiz comunica que os argumentos ou as conclusões apresentadas não são apenas suas, mas, sim, pertença da comunidade, estes argumentos aparecem como incontestáveis. A questão é que, muitas vezes, neste género de texto, se atribuem determinadas asserções à voz da *doxa*, quando nada habilita a assumir que a *doxa* considere aquelas proposições como verdades genéricas, gerando-se assim uma espécie de argumentação falaciosa.

Finalmente, as instâncias elencadas no ponto 2.3.1.4 comprovam como a densidade de ocorrência de lexemas de polaridade semântica negativa (neste caso) pode constituir um indício de uma sequência tendenciosa. Os lexemas de polaridade semântica marcada, adjetivos e advérbios avaliativos, e nomes e verbos com carga avaliativa, indiciam a presença de modalidade apreciativa/avaliativa nos textos e, logo, a tomada de posição por parte do enunciador.

<sup>23</sup> Na construção “estar longe de V” V representa a variável gramatical “verbo”.

<sup>24</sup> Na sua maioria, estes exemplos constituem casos de litotes.



### 3. Conclusões

Baseando-se em trabalhos e modelos anteriores, sobre o estudo da subjetividade em Linguística e em Linguística Computacional, esta breve reflexão teve como objetivo transferir o estudo da subjetividade e do enviesamento/*bias* linguístico para o ainda não explorado género das decisões judiciais.

Aplicada à análise de decisões judiciais publicadas ou em curso, uma reflexão como a apresentada neste trabalho pode abrir caminho para a deteção de palavras, categorias e estruturas linguísticas potencialmente tendenciosas. Se alargada a um universo maior de exemplares, o estudo poderá contribuir para o desenvolvimento de uma aplicação computacional para a deteção automática de discurso tendencioso em decisões judiciais, já escritas ou ainda em construção.

Embora o presente estudo assuma um carácter exploratório, decorrendo de um número pequeno de ocorrências, acreditamos que identifica alguns dos critérios linguísticos subjacentes à expressão de enviesamento. Não obstante esta conclusão, será importante a replicação desta análise em *corpora* mais extensos, já que essa extensão ampliará a perspetiva sobre o fenómeno e, dessa forma, ajudará a gerar uma maior consciência científica e social do preconceito nos acórdãos jurídicos, mais especificamente nos casos de violência de género contra a mulher, contribuindo, eventualmente, para a construção de um discurso mais igualitário neste género de texto.

### Referências

- Albelda, Marta (2005) *La intensificación en el español coloquial*. Valencia: Universidad de Valencia.
- Balahur, Alexandra, Ester Boldrini, Andres Montoyo & Patricio Martínez-Barco (eds) (2011) *Proceedings of the Workshop on Computational Approaches to Subjectivity and Sentiment Analysis WASSA*. Portland, Oregon, USA, Association for Computational Linguistics (ACL), ISBN-13 9781937284060
- Benveniste, Émile (1966) *Problèmes de linguistique générale*, vol. I. Paris: Gallimard.
- Beukeboom, Camiel & Christian Burgers (2017) Linguistic bias. In. Howard Giles & Jake Harwood (eds.) *Oxford Research Encyclopedia of Communication*. United States: Oxford Research Encyclopedias, Oxford University Press, pp. 1-21. <https://doi.org/10.1093/acrefore/9780190228613.013.439>
- Briz, Antonio & Marta Albelda Marco (2013) Una propuesta teórica y metodológica para el análisis de la atenuación lingüística en español y portugués. La base de un proyecto com ún (es.por.atenuaciOn). *Onomázein Revista de lingüística, filología y traducción* 28, pp. 288–319. <https://doi.org/10.7764/onomazein.28.21>
- Cruz, André Ferreira, Gil Rocha & Henrique Lopes Cardoso (2020) On document representations for detection biased news articles. *Procs. 35th Annual ACM Symposium on Applied Computing*. p. 892–899. SAC '20, Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3341105.3374025>
- De-Arteaga, Maria, Alexey Romanov, Hanna Wallach, Jennifer Chayes, Christian Borgs, Alexandra Chouldechova, Sahin Geyik, Krishnaram Kenthapadi & Adam Tauman Kalai (2019) Bias in bios: A case study of semantic representation bias in a high-stakes setting. *Procs. Conf. on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, pp. 120-128. <https://doi.org/10.1145/3287560.3287572>
- Hyland, Ken (1998) *Hedging in Scientific Research Articles*. John Benjamins.
- Karttunen, Lauri (1971) Implicative verbs. *Language* 47 (2), 340–358.
- Kerbrat-Orecchioni, Catherine (1980) *L'énonciation: De la subjectivité dans le langage*. Paris: Armand Colin.
- Kiparsky, Paul & Carol Kiparsky (1970) Fact. *Progress in Linguistics*. De Gruyter Mouton, pp. 143–173.
- Kiesling, Scott, Umashanthi Pavalanathan, Jim Fitzpatrick, Xiaochuang Han & Jacob Eisenstein (2018) Interactional Stancetaking in Online Forums. *Computational Linguistics* 44 (4), pp. 683–718. doi: [https://doi.org/10.1162/coli\\_a\\_00334](https://doi.org/10.1162/coli_a_00334)



- Marques, Aldina, Isabel Margarida Duarte, Alexandra Guedes Pinto & Catarina Pinho (2019) A construção da identidade da mulher em revistas do Estado Novo. *Ex aequo*, pp. 71–88.
- Menegatti, Michela & Monica Rubini (2017) Gender bias and sexism in language. In: Jon F. Nussbaum (ed.) *Oxford Research Encyclopedia of Communication*. Oxford University Press.
- Nascimento, Ana Luíza Tinoco (2017) “*Cultura do Estupro*” e a culpabilização da vítima ou o arquétipo da Condessa Szemioth. Dissertação de Mestrado, Faculdade de Direito da Universidade de Coimbra. Recuperado de <https://estudogeral.sib.uc.pt/handle/10316/81094>
- Oliveira, Fátima (2003) Modalidade e modo. In: Maria Helena Mira Mateus *et al*, *Gramática da Língua Portuguesa*. 5ª edição revista e aumentada. Lisboa: Caminho, pp. 245-247.
- Oliveira, Fátima & Amália Mendes (2013) Modalidade. In: Eduardo Buzaglo Paiva Raposo *et al*, *Gramática do Português*, vol. I. Lisboa: Fundação Calouste Gulbenkian, pp. 623-669.
- Pang, Bo & Lillian Lee (2008) Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval* 2, pp. 1–135. <https://doi.org/10.1561/15000000011>.
- Pinto, Alexandra Guedes (2016) A retórica do eu e do outro - the *othering*: A gramática da identidade no discurso político. *Estudos do Discurso: Caminhos e Tendências*. São Paulo: Editora Paulistana. pp. 25-53.
- Pinto, Alexandra Guedes (2019) A construção ideológica da mulher num acórdão sobre violência doméstica. *Savoir et pouvoir dans un monde polycentrique: les discours aux prismes des langues, des cultures et des espaces*. Paris: Congrès DNC3-ALED (Working Paper).
- Pinto, Alexandra Guedes (2021) A construção da identidade da mulher num acórdão sobre violência doméstica. *Revista de Estudos Linguísticos da Universidade do Porto - N° Especial* – pp. 27 – 46.
- Prabhakaran, Vinodkumar, Ben Hutchinson & Margaret Mitchell (2019) Perturbation sensitivity analysis to detect unintended model biases. *Procs. 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int. J. Conf. on Natural Language Processing (EMNLP-IJCNLP)*. ACL, Hong Kong, China, pp. 5740-5745. <https://doi.org/10.18653/v1/D19-1578>
- Recasens, Marta, Cristian Danescu-Niculescu-Mizil & Dan Jurafsky (2013) Linguistic models for analyzing and detecting biased language. *Procs. 51st Annual Meeting of the Association for Computational Linguistics*. vol. 1. ACL, Sofia, Bulgaria, pp. 1650-1659.
- Riloff, Ellen & Janyce Wiebe (2003) Learning extraction patterns for subjective expressions. *Procs. 2003 Conf. on Empirical Methods in Natural Language Processing*. pp. 105–112.
- Saffioti, Heleieth (1992) Rearticulando gênero e classe social. In: Albertina de Oliveira Costa & Cristina Bruschini (Orgs.) *Uma Questão de gênero*. São Paulo; Rio de Janeiro: Rosa dos Tempos.
- Saffioti, Heleieth (1997) Violência de gênero: lugar da práxis na construção da subjetividade. *Revista Lutas Sociais*, São Paulo, n. 2.
- Sousa-Silva, Rui (2018) Computational forensic linguistics: An overview of computational applications in forensic contexts. *Language and Law* 5, pp. 118–143.
- Sousa Silva, Rui, Gustavo Laboreiro, Luís Sarmento, Tim Grant, Eugénio Oliveira & Belinda Maia (2011) ‘twazn me!!! (’ automatic authorship analysis of micro-blogging messages. In: Rafael Muñoz, Andrés Montoyo & Elisabeth Métais (eds.) *Natural Language Processing and Information Systems*. Springer, pp. 161-168.
- Varikas, Eleni (1989) Jornal das damas: feminismo no sec. XIX na Grécia. *Seminário relações sociais de gênero versus relações de sexo*. São Paulo: FFLCH/USP.
- Werner, Gregory, Vinodkumar Prabhakaran, Mona Diab & Owen Rambow (2015) Committed Belief Tagging on the FactBank and LU Corpora. A Comparative Study Proceedings of the Second Workshop on Extra-Propositional Aspects of Meaning in Computational Semantics (*ExProM*).
- Wiebe, Janyce, Theresa Wilson, Rebecca Bruce, Matthew Bell & Melanie Martin (2004) Learning subjective language. *Computational Linguistics* 30 (3), pp. 277–308. <https://doi.org/10.1162/0891201041850885>.

