

Letras graúdas, letras miúdas

leitura e complexidade textual dos discursos
em Plenário 2007-2024

Pedro Duarte Blanco



Letras graúdas, letras miúdas:

**Leitura e complexidade textual dos discursos
em Plenário 2007-2024**

Pedro Duarte Blanco¹

¹ Consultor Legislativo do Senado Federal, especialidade Pronunciamentos. Mestre em História e Literatura pela Columbia University. Contato: <<https://pedblan.github.io>> / e-mail: pedblan@senado.leg.br

SENADO FEDERAL

DIRETORIA GERAL

Ilana Trombka – Diretora-Geral

SECRETARIA GERAL DA MESA

Danilo Augusto Barboza de Aguiar – Secretário Geral

CONSULTORIA LEGISLATIVA

Paulo Henrique de Holanda Dantas – Consultor-Geral

NÚCLEO DE ESTUDOS E PESQUISAS

Rafael Silveira e Silva – Coordenação

Alexandre Sidnei e Brunella Poltronieri Miguez – Revisão

João Cândido de Oliveira – Editoração

CONSELHO EDITORIAL

Eduardo Modena Lacerda

Pedro Duarte Blanco

Denis Murahovschi

Foto da Capa: Saulo Cruz/Agência Senado

O conteúdo deste trabalho é de responsabilidade dos autores e não representa posicionamento oficial do Senado Federal.

É permitida a reprodução deste texto e dos dados contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

Como citar este texto:

BLANCO, Pedro Duarte. **Letras graúdas, letras miúdas**: leitura e complexidade textual dos discursos em Plenário 2007-2024. Brasília: Núcleo de Estudos e Pesquisas/CONLEG/Senado, Fevereiro 2026 (Texto para Discussão nº 357). Disponível em: <www.senado.leg.br/estudos>. Acesso em: 11 fev. 2026.

Núcleo de Estudos e Pesquisas
da Consultoria Legislativa



Conforme o Ato da Comissão Diretora nº 14, de 2013, compete ao Núcleo de Estudos e Pesquisas da Consultoria Legislativa elaborar análises e estudos técnicos, promover a publicação de textos para discussão contendo o resultado dos trabalhos, sem prejuízo de outras formas de divulgação, bem como executar e coordenar debates, seminários e eventos técnico-acadêmicos, de forma que todas essas competências, no âmbito do assessoramento legislativo, contribuam para a formulação, implementação e avaliação da legislação e das políticas públicas discutidas no Congresso Nacional.

Contato:

conlegestudos@senado.leg.br

URL: www.senado.leg.br/estudos

ISSN 1983-0645

Letras graúdas, letras miúdas

leitura e complexidade textual dos discursos em Plenário 2007-2024

RESUMO

Este Texto para Discussão analisa a evolução da legibilidade e de traços morfossintáticos dos discursos proferidos no plenário do Senado Federal entre 2007 e 2024, a partir de técnicas computacionais de análise textual. Utilizam-se métricas clássicas de legibilidade e proporções de classes gramaticais extraídas com o pacote *textdescriptives* da biblioteca spaCy, aplicadas a um corpus de 37.455 discursos com mais de 300 palavras, excluídos os apartes. Os resultados indicam redução da dificuldade textual captada por métricas de legibilidade. Tal redução foi acompanhada de aumento na proporção de pronomes, verbos auxiliares e advérbios, e de queda relativa de adjetivos, preposições e nomes próprios. Esses padrões variam sistematicamente segundo partido e gênero dos oradores. A análise sugere que tais transformações não se explicam apenas por simplificação linguística, mas são compatíveis com a formação de um registro discursivo adaptado à circulação dos pronunciamentos nas redes sociais, associado a processos de personalização e democratização da linguagem parlamentar.

PALAVRAS-CHAVE: Discursos proferidos; Textdescriptives; Flesch Reading Ease; Biblioteca spaCy; Senado Federal; Plenário.

SUMARIO

1	INTRODUÇÃO	6
2	METODOLOGIA	6
	2.1. LEITURABILIDADE.....	8
	2.1.1. FLESCH READING EASE	9
	2.1.2. OUTRAS MÉTRICAS.....	10
	2.2. CLASSES GRAMATICAIS E SUAS FREQUÊNCIAS	14
	2.2.1. BLOCO A: CLASSES CUJA PROPORÇÃO CRESCEU	16
	2.2.2. BLOCO B: CLASSES CUJA PROPORÇÃO DIMINUIU	22
	2.3. CASO ESPECIAL: PRONOMES PESSOAIS RETOS	27
	2.4. SÍNTESE DOS RESULTADOS	29
3	DISCUSSÃO	31
	3.1. CONSIDERAÇÕES METODOLÓGICAS	31
	3.2. LEITURABILIDADE E COMPLEXIDADE TEXTUAL.....	34
	3.3. PESSOALIZAÇÃO E DEMOCRATIZAÇÃO.....	38
	3.3.1. PESSOALIZAÇÃO	38
	3.3.2. DEMOCRATIZAÇÃO	43
	3.4. DISCURSO, REDES SOCIAIS, POPULISMO	50
4	CONCLUSÃO	56
5	REFERÊNCIAS BIBLIOGRÁFICAS.....	57

1 INTRODUÇÃO

Em “Plenário, palanque, estúdio”, meu artigo anterior, elaborei hipóteses que vinculavam as mudanças nos discursos no plenário do Senado com a expansão do uso das redes sociais pelos legisladores. A ideia central foi a de que os pronunciamentos já não visam primordialmente ao debate parlamentar; tornaram-se, antes, oportunidades para a gravação de vídeos para os novos tipos de mídia. Em outras palavras, a tribuna do plenário passou a ser usada como “estúdio” (Blanco, 2025b).

Neste Texto para Discussão, busco aprofundar e qualificar essa hipótese, investigando como esse processo transcorre em nível linguístico.

A inspiração do projeto veio de editorial veiculado no jornal *Washington Post*. A coluna atribuiu o sucesso de Donald Trump nas eleições presidenciais de 2016 ao fato de que o político norte-americano se expressa de maneira simples (Smith, 2016). Essa tese, baseada em métricas de complexidade textual, fundamentou leituras retóricas, históricas e sociolinguísticas sobre a política contemporânea, em trabalhos como os de Degani (2016).

Em que medida o raciocínio se aplica à política brasileira? Para responder a essa pergunta, apliquei algumas dessas métricas a um banco de discursos proferidos no plenário do Senado, usando ferramentas de análise computacional, e refleti sobre os resultados à luz de teorias de linguística e análise de discurso. Descrevo com detalhes essa metodologia na seção seguinte.

2 METODOLOGIA

A metodologia da pesquisa consistiu na aplicação de técnicas computacionais e métodos de estatística descritiva a discursos proferidos no plenário do Senado Federal entre 2007 e 2024. O banco de dados foi formado a partir de requisições ao Portal de Dados Abertos do Senado Federal (<<https://www12.senado.leg.br/dados-abertos>>) e complementado com a transcrição automatizada dos pronunciamentos em vídeo, feitos nas sessões de deliberação remota realizadas durante o período da pandemia.

Depois da etapa de coleta, submeti os textos a um procedimento de normalização taquigráfica, visando a excluir os anexos dos discursos. Estes são textos cuja inclusão nos anais do Senado Federal é requerida pelos oradores.

Os anexos foram retirados do estudo porque seu peso nas notas taquigráficas poderia provocar um viés relevante na análise estatística. Pelo mesmo motivo, excluí da apreciação deste trabalho os discursos com apartes. As observações dos taquígrafos – em geral, bastante breves – foram, contudo, mantidas.

Em um terceiro passo, os discursos foram analisados por meio do pacote *textdescriptives* (Hansen; Olsen; Enevoldsen, 2023), disponível aos usuários da biblioteca de processamento de linguagem natural *spaCy*. Em linhas gerais, *textdescriptives* faz estatísticas sobre morfologia, sintaxe e outros aspectos linguísticos de textos, além de analisá-los sob as métricas de complexidade textual mencionadas na introdução deste artigo. Por meio de *scripts* Python, apliquei os métodos de *textdescriptives* a **todos os pronunciamentos com mais de 300 palavras** existentes no banco de dados, entre aqueles **que não receberam apartes**. No total, analisei 37.455 discursos proferidos por 325 senadores, apresentando os resultados da pesquisa em gráficos com as tendências centrais, ao longo da série histórica, de cada componente medido. Os discursos, bem como os resultados das análises computacionais, encontram-se disponíveis na plataforma Kaggle¹ (Blanco, 2025a).

Preparei estudos de caso para ilustrar os resultados da pesquisa. Esses estudos abordam o comportamento de medianas agrupadas segundo três categorias: parlamentares, seus partidos e seus gêneros. Ainda fiz um exame mais aprofundado da incidência de pronomes pessoais retos nos discursos, cuja evolução temporal apresenta um padrão expressivo.

Na parte de discussão, analiso os resultados à luz de dois trabalhos de linguística. O primeiro consiste nas teorias de análises de *corpus*, registro e estilo, presentes na obra de Douglas Biber e Susan Conrad (Biber; Conrad, 2019). O segundo é o método da análise crítica de discurso elaborado por Norman Fairclough (Fairclough, 1996). Ao final, reforço meus pontos de vista com pesquisas por expressões específicas, utilizando-as como heurística para as práticas discursivas dos senadores.

Vale destacar o emprego de ferramentas de inteligência artificial neste trabalho. Para elaborar os *scripts* de coleta e análise de textos, usei os modelos de linguagem GPT-5.2, da empresa OpenAI, e Gemini 2.5, da empresa Google. Para revisão crítica, utilizei, mais uma vez, o modelo GPT-5.5. Todos esses procedimentos foram controlados e revisados por mim, razão pela qual assumo plena responsabilidade por qualquer erro.

¹ Disponível em <<https://www.kaggle.com/datasets/pedrouarteblanco/falando-nela>>.

2.1. Leiturabilidade

Uma das primeiras aplicações das métricas de leiturabilidade deu-se no apoio a pesquisas sobre o sistema educacional norte-americano. As pesquisas dos educadores Edgar Dale e Jeanne Chall fazem referência explícita (Dale; Chall, 1949) a essa questão. Muitas das escalas empregadas para avaliar a complexidade de um texto se baseiam na escolaridade, em anos, tida como necessária para compreendê-lo.

É por isso que a maior parte das métricas apresentadas neste trabalho tem uma relação diretamente proporcional com a dificuldade de se ler o texto avaliado – em outras palavras, quanto mais alto o valor atribuído a um texto, mais difícil ele é para o aluno leitor. A exceção é o índice *Flesch Reading Ease*, talvez a mais conhecida dessas métricas, que mede a facilidade de se ler determinado trecho.

De maneira geral, as medidas de leiturabilidade usadas na pesquisa não estão adaptadas a padrões específicos da língua portuguesa, o que impossibilita seu uso na comparação do material coletado com discursos proferidos em outras línguas. O objetivo do trabalho, no entanto, é o de comparar apenas discursos em português; por isso, a calibragem da métrica não foi vista como um impedimento à análise que proponho neste artigo. Mas é importante ter esse fato em consideração, para que as métricas apresentadas não sejam extrapoladas indevidamente.

Feita essa ressalva, os gráficos a seguir mostram a trajetória dos discursos ao longo da série histórica nas seguintes métricas: *Flesch Reading Ease*, *Gunning Fog*, *Simplified Measure of Gobbledygook* (SMOG), *Automated Readability Index* e *Coleman-Liau*. Além da mediana anual dos discursos, esses gráficos apresentam, para auxiliar a visualização, a métrica correspondente a um discurso clássico da história recente: a fala de promulgação da Constituição Federal de 1988, proferida pelo deputado Ulysses Guimarães, presidente da Assembleia Nacional Constituinte (Guimarães, 2023). Esse discurso é referido na legenda dos gráficos pela frase “Ódio e nojo à ditadura”. Ao final, apresenta-se um gráfico com a variação anual de todos esses índices em relação à respectiva mediana global – em outras palavras, o “delta”.

2.1.1. Flesch Reading Ease

A métrica de Flesch empregada em textdescriptives baseia-se no comprimento das sentenças e no número de sílabas das palavras. Ela é calculada por meio da seguinte fórmula:

$$\text{Flesch Reading Ease} = 206.835 - (1.015 \times \text{Comprimento médio das sentenças}) - (84.6 \times \text{número médio de sílabas por palavra})$$

(“Readability”, [S.d.])²

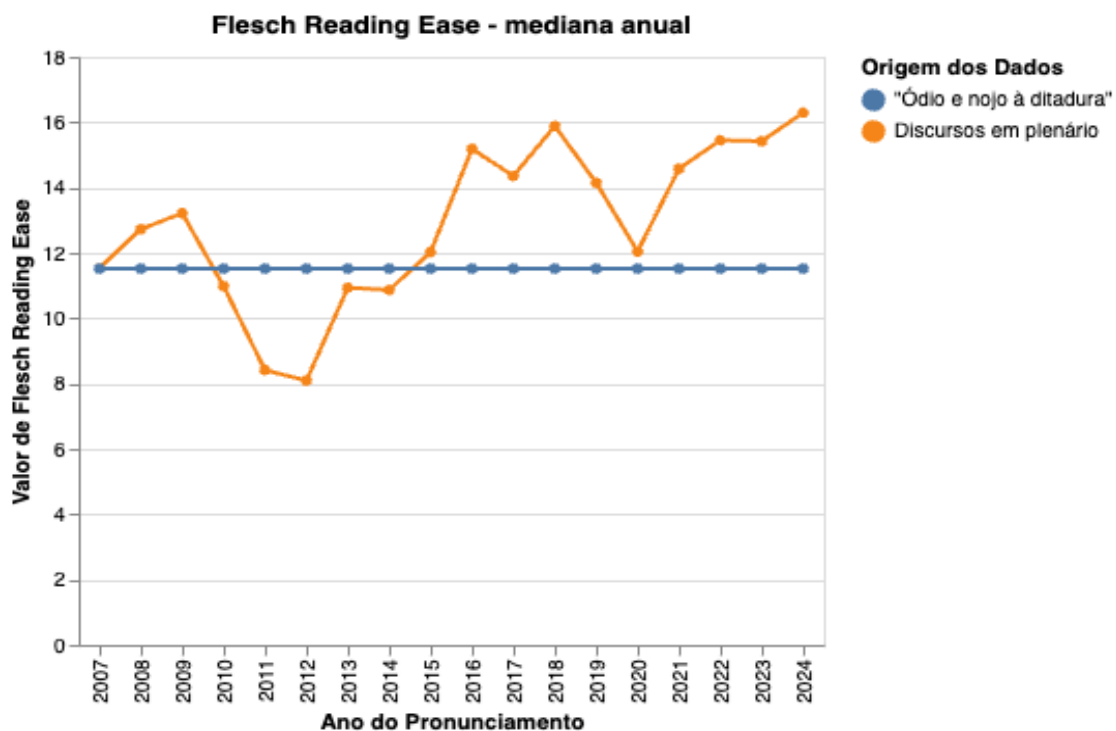


Figura 1 – Flesch Reading Ease, mediana anual. Fonte: elaboração própria a partir de textdescriptives.

Percebe-se que a legibilidade Flesch dos discursos em plenário teve uma trajetória oscilante: duas depressões – no biênio 2010-2011 e em 2020³ –,

² Na versão para a língua portuguesa feita pelo Núcleo Interinstitucional de Linguística Computacional, da Universidade de São Paulo, a fórmula de Flesch é a seguinte: $248,835 - [1,015 \times (\text{média de palavras por sentença})] - [84,6 \times (\text{média de sílabas por palavra})]$.

Percebe-se que a adaptação consistiu na alteração da constante aditiva do segundo termo da equação, mantendo-se inalterados os coeficientes associados ao comprimento médio das sentenças e ao número médio de sílabas por palavra. Por esse motivo, não observei maiores problemas em usar a métrica de textdescriptives para os propósitos deste trabalho.

picos locais em 2009, 2016 e 2018 e recorde em 2024. Ao final, variou positivamente quase 20% ao longo da série histórica, o que significa aumento na facilidade de leitura. Esse comportamento é respaldado pelos demais índices, cuja trajetória é inversa – como observado, estes são inversamente proporcionais ao índice de Flesch.

2.1.2. Outras métricas

As demais métricas apresentadas neste trabalho visam ao controle do resultado obtido por Flesch Ease. Elas são resumidas na tabela abaixo:

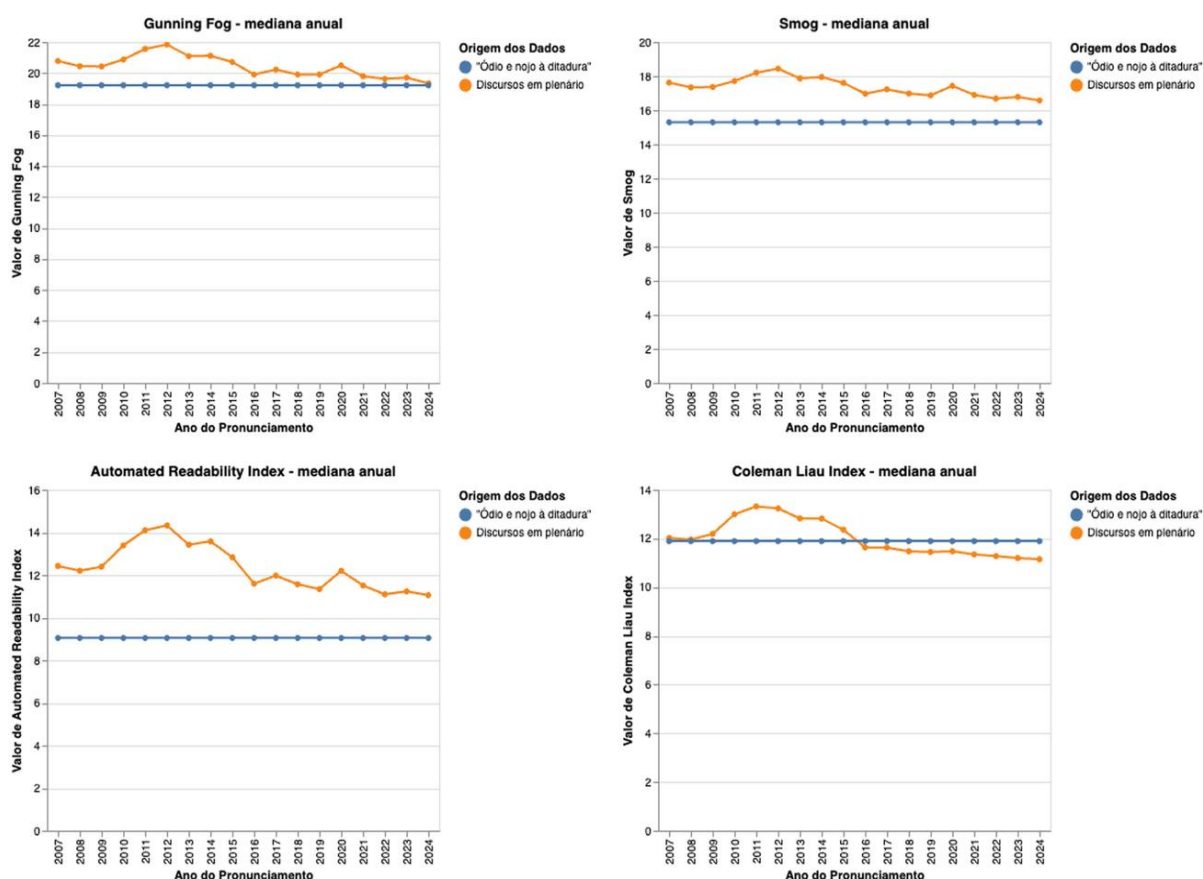
Métrica	O que mede	Princípio linguístico central	Interpretação do valor	Fórmula
<i>Gunning Fog Index</i>	Grau de dificuldade do texto em termos de escolaridade necessária para compreensão.	Comprimento médio das sentenças e proporção de palavras longas (mais de duas sílabas).	Valores mais altos indicam textos mais complexos e exigentes do ponto de vista educacional.	$GF = 0,4 \times (\text{Comprimento médio das sentenças} + \text{Número de palavras difíceis}).$
<i>SMOG (Simplified Measure of Gobbledygook)</i>	Idem acima.	Frequência de palavras longas em relação ao número de sentenças.	Valores mais altos indicam maior complexidade lexical e sintática.	$1.043 \times \sqrt{(30 \times (\text{número de palavras difíceis} / \text{número de sentenças})) + 3.1291}.$
<i>Automated Readability Index (ARI)</i>	Dificuldade de leitura associada à densidade gráfica do texto.	Número médio de caracteres por palavra e de palavras por sentença.	Valores mais altos indicam textos mais densos e menos acessíveis.	$RI = 4,71 \times (n^2 \text{ de caracteres} / n^2 \text{ de palavras}) + 0,5 \times (n^2 \text{ de palavras} / n^2 \text{ de sentenças}) - 21,43.$
<i>Coleman–Liau Index</i>	Idem acima.	Proporção de caracteres e sentenças por número de palavras.	Valores mais altos indicam maior complexidade formal do texto.	$CLI = 0,0588 \times L - 0,296 \times S - 15,8^4$

Tabela 1 – Métricas de legibilidade textual. Fonte: elaboração própria, a partir de análise com textdescriptives (Cherian; Jha, 2024; “Readability”, [S.d.]).

³ A depressão de 2020 na métrica de Flesch pode resultar de um viés importante de registro (v. Discussão, *infra*). No banco de dados, a maior parte dos discursos do ano de 2020 constitui-se de transcrições feitas com inteligência artificial (modelo Whisper, da empresa OpenAI). Essas se referem aos vídeos em que as falas foram gravadas durante as sessões remotas realizadas durante a pandemia de covid-19. É possível que o modelo tenha usado critérios distintos para determinar a pontuação (e, portanto, o comprimento das sentenças). Mas também é possível que, ao gravar as falas, os parlamentares tenham usado frases compridas de propósito, como no estilo das frases de arrastão (Garcia, 2010).

⁴ “L” corresponde ao número médio de caracteres por 100 palavras e “S” ao número médio de sentenças por 100 palavras. Na versão implementada por textdescriptives, a adotada neste trabalho, esses valores são calculados com base no texto integral, e não apenas em um recorte fixo de 100 palavras.

Como observado, as demais métricas de leitura são inversamente proporcionais à de Flesch. Os resultados da aplicação dessas métricas corroboram a análise anterior. Ao longo do período estudado, houve redução da complexidade textual dos discursos proferidos em plenário. É o que indica o painel abaixo:



Painel 1 – Índices Gunning Fog, SMOG, ARI e Coleman-Liau. Medianas anuais.
Fonte: elaboração própria a partir de textdescriptives.

A principal diferença entre os índices é o comportamento em relação ao texto que serviu como parâmetro de visualização, o discurso de promulgação da Constituição Federal. A trajetória do índice Gunning Fog encontra-se próxima à marca de referência, o que talvez reflita a frequência alta de palavras com mais de duas sílabas na língua portuguesa. Coleman-Liau, ao cruzar a linha do discurso de promulgação, dá outra perspectiva a respeito do tema, mostrando queda substancial no tamanho de palavras e sentenças, principalmente depois de 2016. Em quase todos os casos, a redução de complexidade começa a partir do biênio 2011-2012.

Os gráficos acima dão uma ideia do padrão discursivo praticado no plenário do Senado. Como, no entanto, as métricas extraídas de `textdescriptives` não necessariamente se encontram calibradas para o português, o propósito deste estudo seria mais bem servido por um gráfico que mostrasse a mudança ao longo dos anos na apuração de cada uma das métricas. O gráfico abaixo, que representa a diferença da mediana anual frente à mediana global do período, tem essa finalidade:

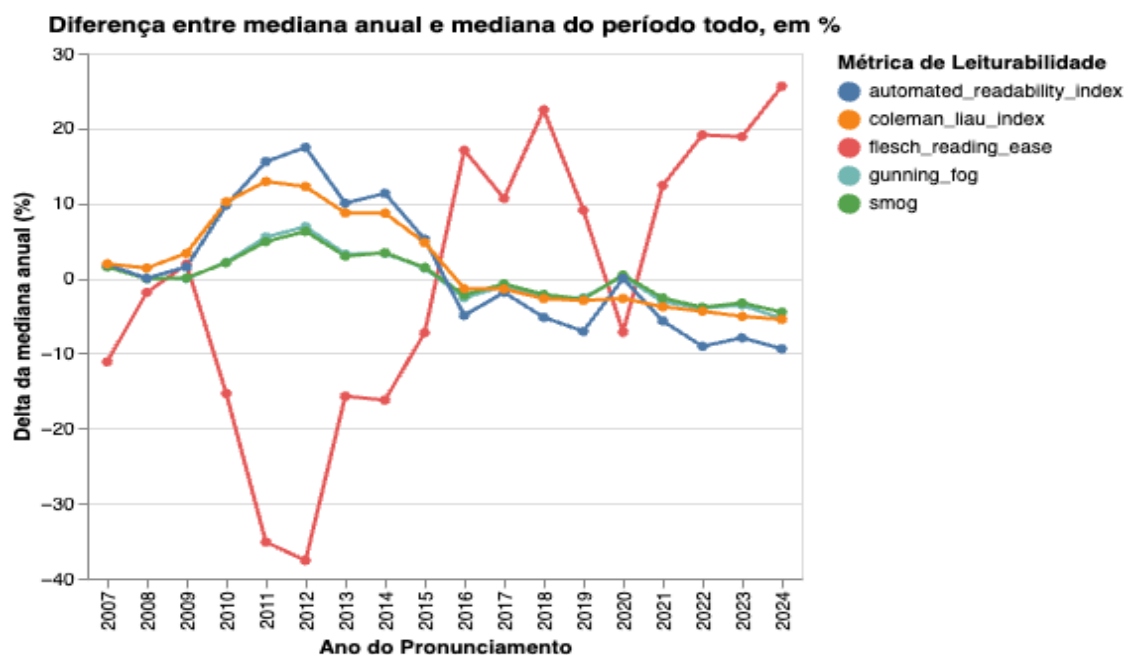
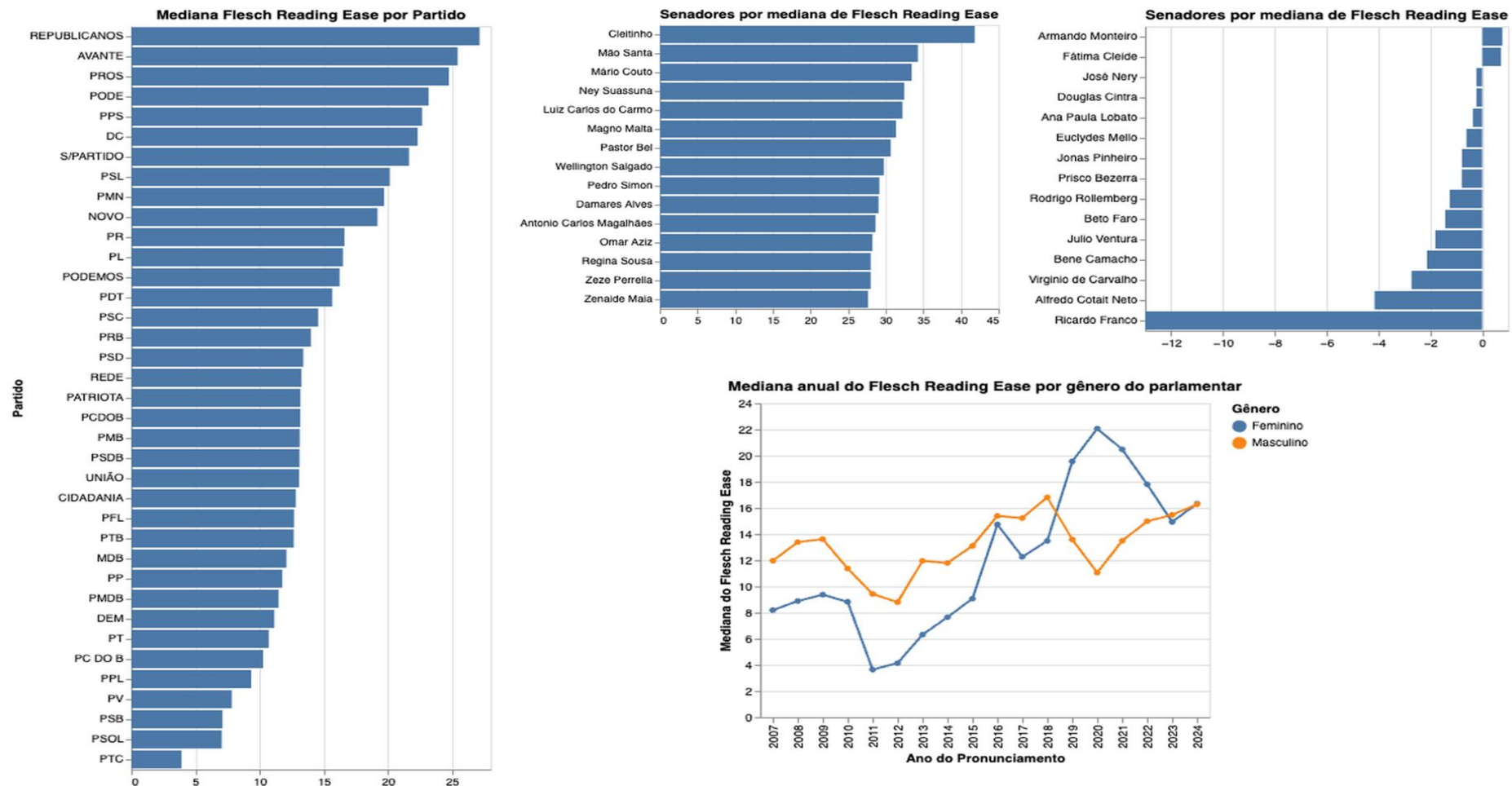


Figura 2 – Leiturabilidade. Diferença entre mediana anual e mediana global, todos os índices. Fonte: elaboração própria a partir de `textdescriptives`.

O gráfico evidencia a relação inversamente proporcional entre *Flesch Reading Ease* e os demais índices. Enquanto o primeiro é uma escala de facilidade de leitura, os demais visam a apontar escolaridade mínima para a compreensão do texto – em outras palavras, são métricas de dificuldade de leitura. Os índices parecem convergir para o mesmo resultado: a redução na complexidade dos discursos medida pelos índices de leiturabilidade.

O painel abaixo finaliza a exposição desses índices por meio de três estudos de caso: comparações agregadas de *Flesch Reading Ease* segundo partido (gráfico de barras horizontais à esquerda da página), orador (gráficos de barras horizontais ao centro e à direita da página) e gênero (gráfico de linhas) dos senadores.

Estudos de caso: Flesch Reading Ease



Painel 2 – Flesch Reading Ease. Mediana agregada por partido, orador e gênero do orador. Fonte: elaboração própria a partir de textdescriptives.

A análise revela um dos padrões que mais se repetem neste artigo. Na comparação por partido, tendo em vista o período todo, constata-se a presença de partidos de direita e centro-direita (PODEMOS⁵, NOVO, PL⁶, PSL) na metade superior da tabela. Isso indica que o pronunciamento típico de parlamentares desses partidos tende a ter leitura mais simples. Em contrapartida, partidos de esquerda e centro-esquerda (PT, PSOL, PSB, PC do B) figuram na metade inferior, o que sugere um discurso típico mais difícil.

Em relação a oradores, percebe-se, entre aqueles com *Flesch* alto, vários senadores recentes com intensa atividade em plenário, como Cleitinho, Magno Malta e Damares Alves. Entre os senadores com baixo grau da métrica *Flesch*, percebem-se parlamentares de legislaturas mais antigas, ou com poucos discursos – com exceções notáveis na figura dos senadores Rodrigo Rollemberg (237 discursos), Armando Monteiro (83) e Fátima Cleide (96).

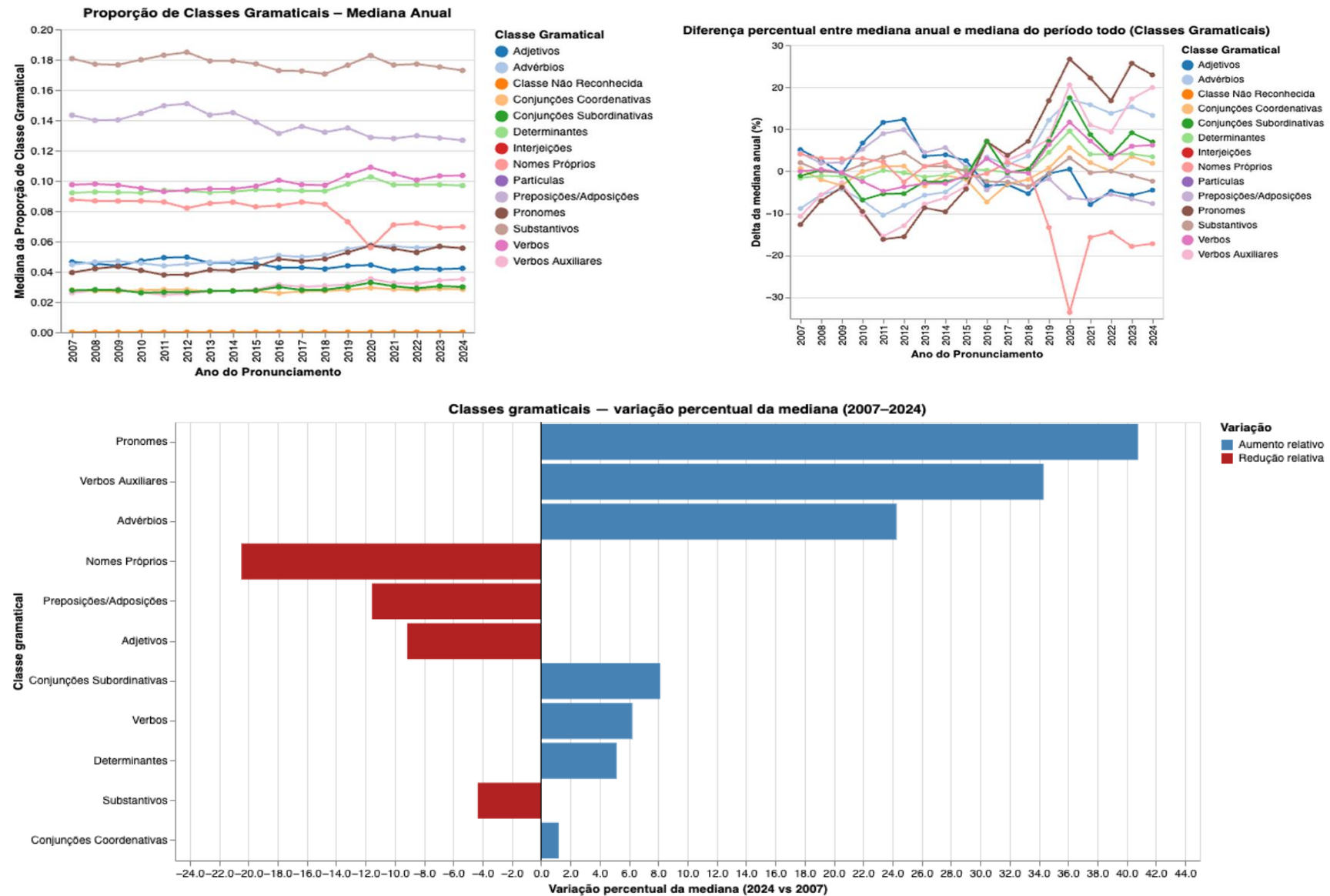
Finalmente, na análise por gênero do orador, percebe-se que a mediana anual de leiturabilidade dos senadores homens foi superior entre 2007 e 2019, ano em que as senadoras tomaram a precedência. Essa situação perdurou até 2023, quando a leiturabilidade Flesch dos discursos de homens e mulheres passou a se igualar.

2.2. Classes gramaticais e suas frequências

O estudo deste aspecto dos discursos valeu-se das proporções de classes gramaticais ante o total de palavras. O painel abaixo mostra três perspectivas a respeito do tema. Em primeiro lugar, a proporção anual de cada classe. Segundo, a diferença entre a mediana anual e a mediana global. Terceiro, a diferença, relativa à mediana anual de cada classe, entre os anos do início e do fim da série histórica – isto é, a variação da série como um todo.

⁵ As nomenclaturas “PODE” e “PODEMOS” dizem respeito ao mesmo partido. “PODE” foi a sigla adotada pelo Portal de Dados Abertos do Senado Federal entre 2017, quando o antigo Partido Trabalhista Nacional registrou a mudança de nome para “PODEMOS”, até 2019. A partir de 2018, “PODEMOS” também passou a ser usado pelo Portal.

⁶ Vale destacar que o PL passou parte do período estudado na base de apoio do primeiro e segundo governo do presidente Lula – isto é, mantendo-se politicamente à esquerda. Nesse período, porém, o partido não tinha grande expressão no Senado Federal, apesar de contar com o vice-presidente José Alencar em suas fileiras. Ao final da série histórica, no entanto, o PL migrou à direita, liderado pelo grupo político do ex-presidente Jair Bolsonaro. Nesse contexto, passou a ter presença relevante nos órgãos do Congresso Nacional. É provável, portanto, que a mediana contemple mais pronunciamentos dessa segunda fase da história do partido.



Painel 3 – Classes gramaticais. Proporção anual (mediana), variação anual (diferença entre mediana anual e mediana global) e variação da mediana anual, em porcento (anos 2007 e 2024). Fonte: elaboração própria a partir de textdescriptives.

À primeira vista, a maior parte das estatísticas de classe gramatical indica padrão morfossintático estável. O gráfico de proporções anuais (acima, à esquerda do painel) mostra a prevalência de substantivos, preposições e verbos, em um padrão comum neste tipo de texto. Essa figura evidencia mudanças na proporção de nomes próprios, pronomes e preposições, mas se vê pouco da evolução histórica dessas métricas: é preciso observar a variação ao longo do tempo para fazer uma análise mais detalhada.

O gráfico à direita possibilita essa leitura, ao mostrar a diferença entre a mediana anual e a mediana da série histórica inteira, isto é, o “delta”. O gráfico indica dois períodos de redistribuição entre as classes gramaticais. O primeiro observa-se entre os anos 2010 e 2012; o segundo, de 2018 em diante. O ano de 2016 representa um ponto de inflexão, em que as linhas se cruzam, como em um “X”.

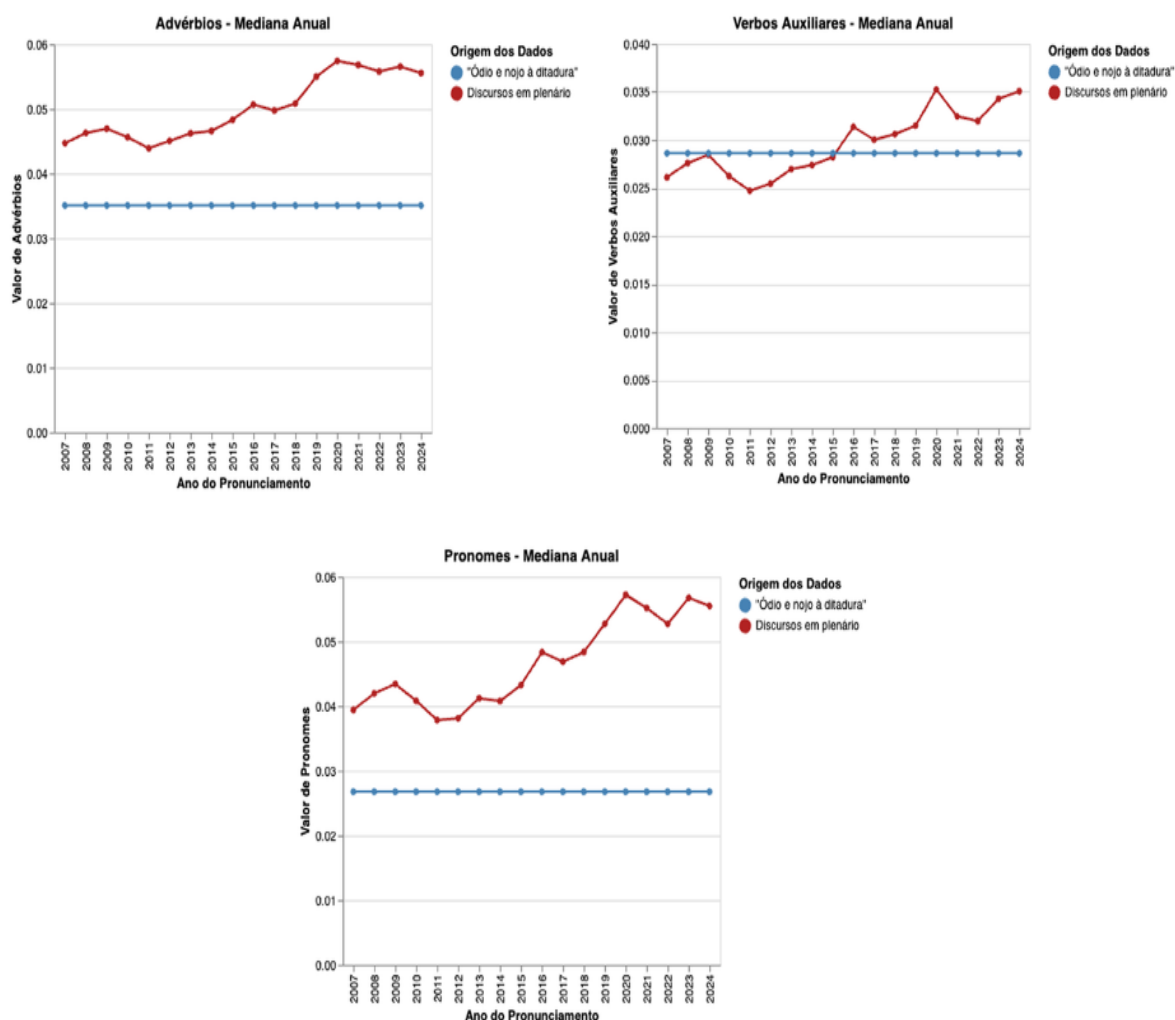
A figura da parte de baixo do painel evidencia as mudanças, pois compara estritamente as medianas anuais de 2007 e 2024, anos inicial e final da pesquisa. Nela se percebe aumento relativo superior a 20% na proporção de classes como pronomes, verbos auxiliares e advérbios. Por sua vez, adjetivos, preposições e nomes próprios registraram queda na mediana anual de pelo menos 10%. A divisão das classes gramaticais nesses dois blocos – classes gramaticais em alta e em baixa – é o princípio que organiza a análise que se segue.

2.2.1. Bloco A: classes cuja proporção cresceu

As classes gramaticais cuja proporção cresceu substancialmente no período foram as de **pronomes**, **advérbios** e **verbos auxiliares**⁷ (“AUX”, [S.d.]; Lopes; Duran; Pardo, 2023). O aumento mais expressivo foi entre os pronomes, cuja proporção cresceu 40%. Em seguida, vêm os verbos auxiliares (34%), e, finalmente, constam os advérbios, cujo incremento foi de 25%. As três métricas seguiram trajetória parecida – com picos locais em 2009 e 2016 e pico global em 2020.

⁷ Verbos auxiliares são uma categoria de verbos funcionais que acompanham o verbo lexical de uma oração, como em formas conjugadas dos verbos “ser”, “estar”, “ir”, “haver”, “ter” e “vir”. Essa classe integra o catálogo do Universal Dependencies, classificação empregada em sistemas de linguística computacional, como `textdescriptives-spaCy`, empregado neste artigo.

De todas essas métricas, a única que cruzou a marca de referência foi a de verbos auxiliares – a partir de 2015, essa categoria de palavras esteve mais presente nos discursos em plenário do que na fala de promulgação da Constituição de 1988. As demais permaneceram muito acima da proporção observada no célebre pronunciamento do deputado Ulysses Guimarães.



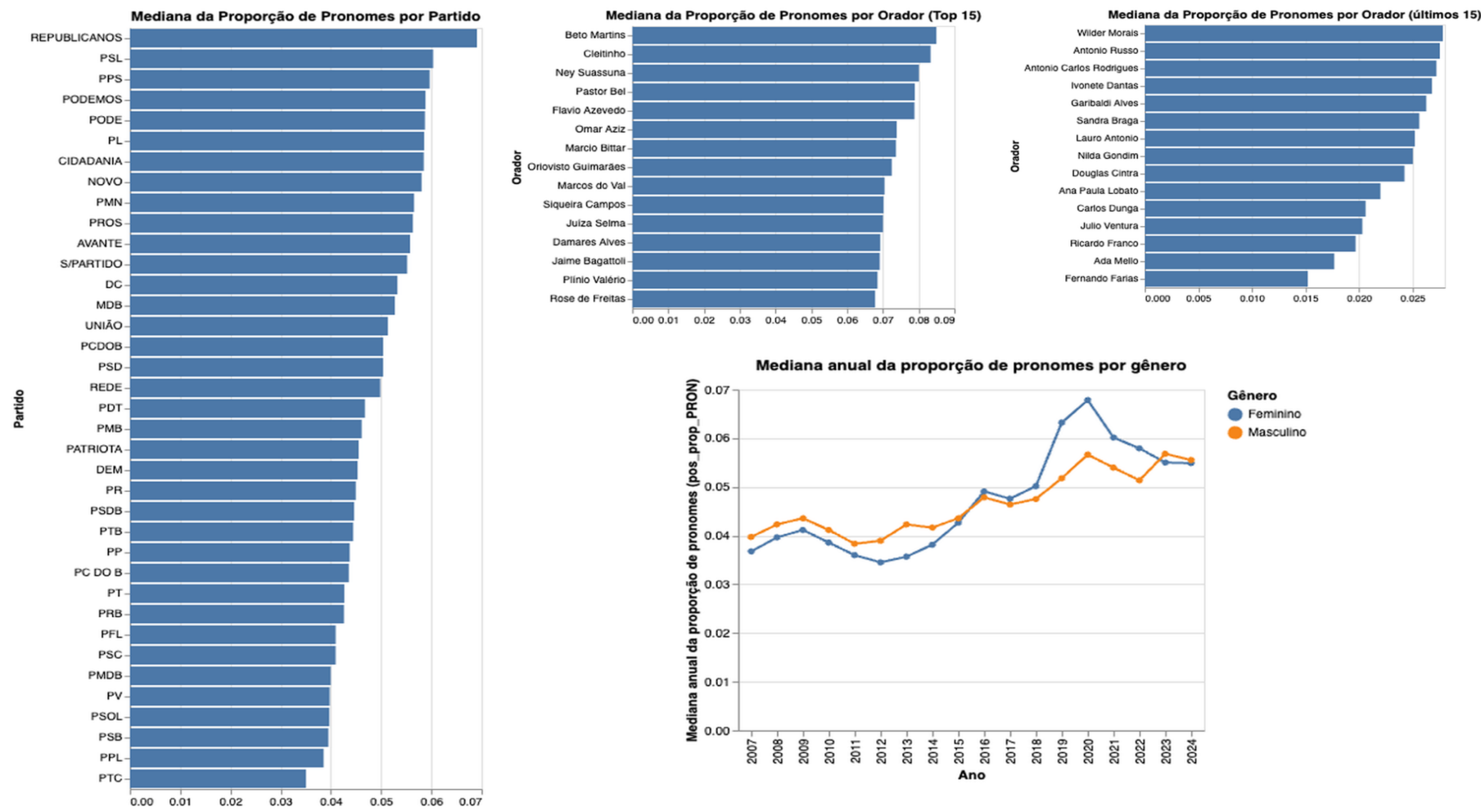
Painel 4 – Proporção de advérbios, verbos auxiliares e pronomes ante o total de palavras. Mediana anual. Fonte: elaboração própria a partir de textdescriptives.

Estudos de caso

Os estudos de caso relacionados a pronomes, verbos auxiliares e advérbios apresentam padrões similares aos observados nos estudos sobre leitura Flesch. Nas agregações, partidos de direita aparecem no topo e partidos de esquerda no fim, padrão que deve ser lido como descritivo e sensível à composição temporal e individual do *corpus*.

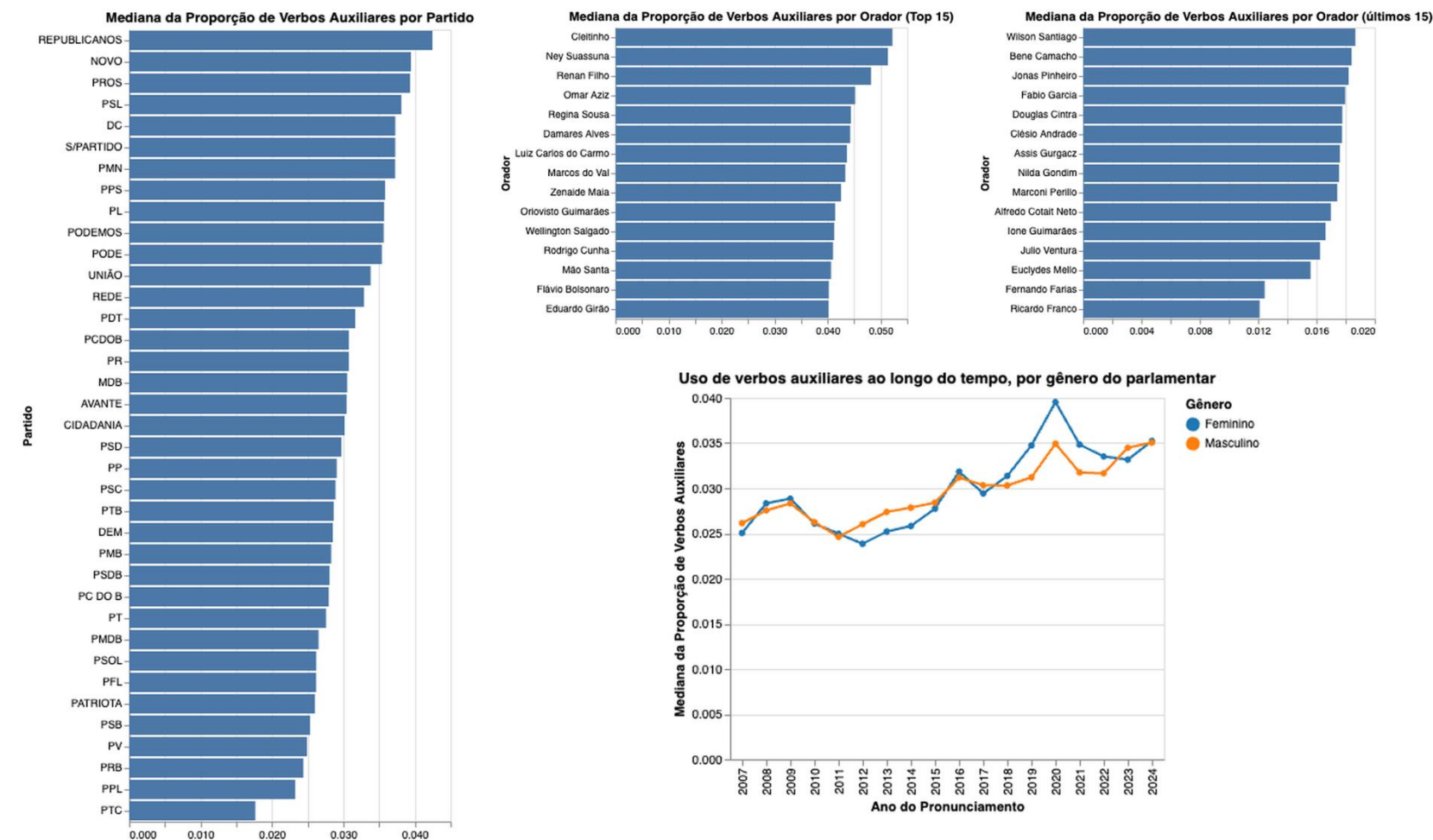
Senadores de legislaturas recentes e com muitos discursos incluem-se entre os oradores com maior proporção; por sua vez, senadores mais antigos e com poucos discursos constam entre os oradores com proporção menor. Finalmente, durante boa parte do período, o crescimento global da proporção dessas classes deu-se em relativa igualdade entre os gêneros, salvo no período 2019-2023 (2016-2023, quanto aos advérbios), quando as senadoras superaram seus colegas homens por ampla margem.

Pronomes



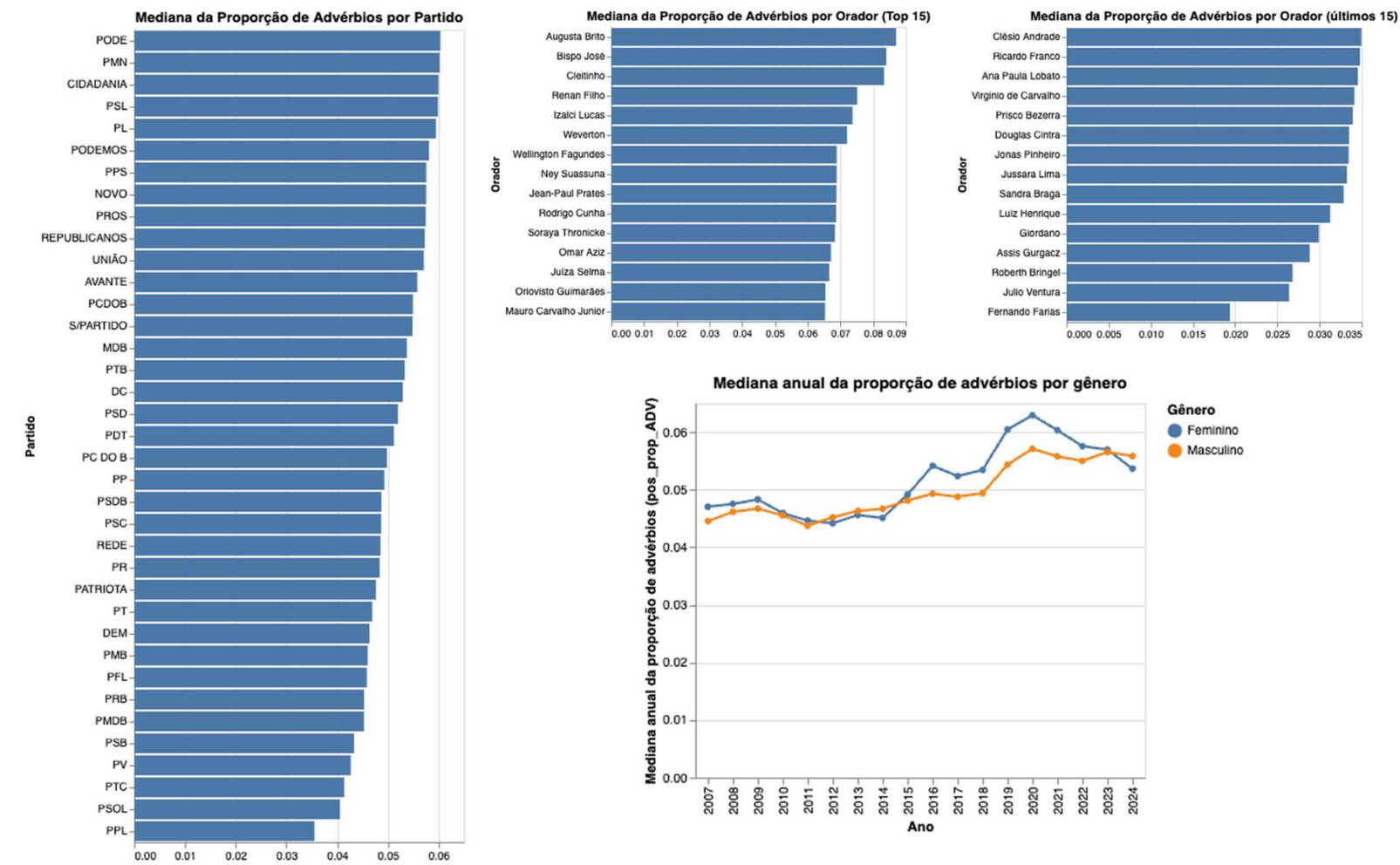
Painel 5 – Pronomes. Mediana agregada por partido, orador e gênero do orador. Fonte: elaboração própria a partir de textdescriptives.

Verbos auxiliares



Painel 6 – Verbos auxiliares. Mediana agregada por partido, orador e gênero do orador. Fonte: elaboração própria a partir de textdescriptives.

Advérbios

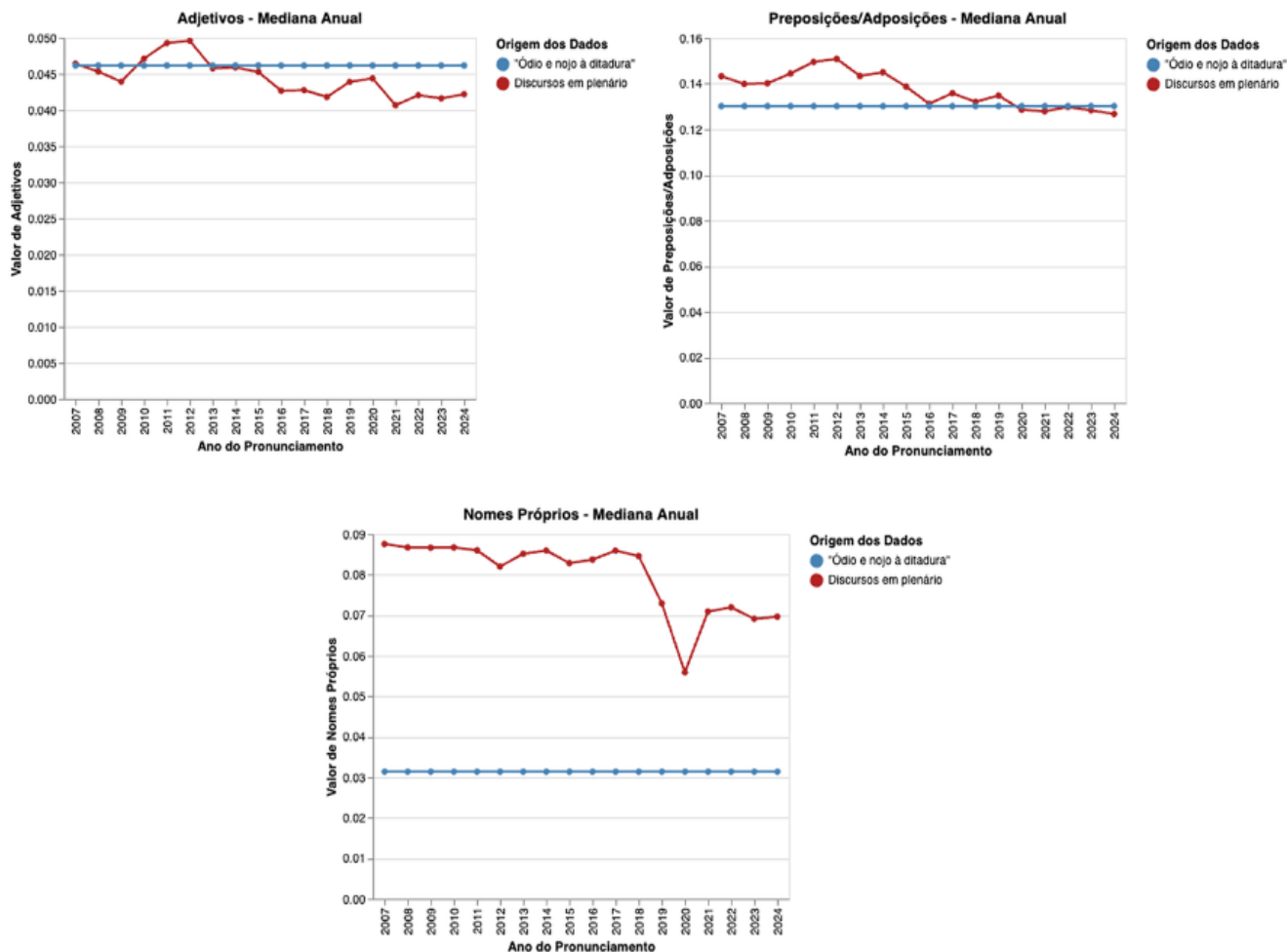


Painel 7 – Advérbios. Mediana agregada por partido, orador e gênero do orador. Fonte: elaboração própria a partir de textdescriptives.

2.2.2. Bloco B: classes cuja proporção diminuiu

As classes cuja proporção se reduziu ao longo do período foram as de **adjetivos, preposições e substantivos próprios**. O declínio das duas primeiras seguiu trajetória parecida, com ondulações e um pico global em 2012. Uma e outra, no entanto, divergem quanto ao momento em que ultrapassaram a linha de referência. O discurso em plenário típico passou a ter menos adjetivos do que a fala de promulgação da Constituição a partir de 2013; quanto às preposições, no entanto, ultrapassou esse limiar apenas em 2020. Finalmente, os nomes próprios seguiram curso *sui generis*: após anos de trajetória estável, essa classe gramatical caiu bastante a partir de 2019, com recuperação apenas parcial depois do fim da pandemia – com a ressalva de um possível viés de registro quanto aos discursos feitos nas sessões deliberativas remotas, cujos vídeos foram transcritos por inteligência artificial (v. 4.1, Considerações metodológicas).

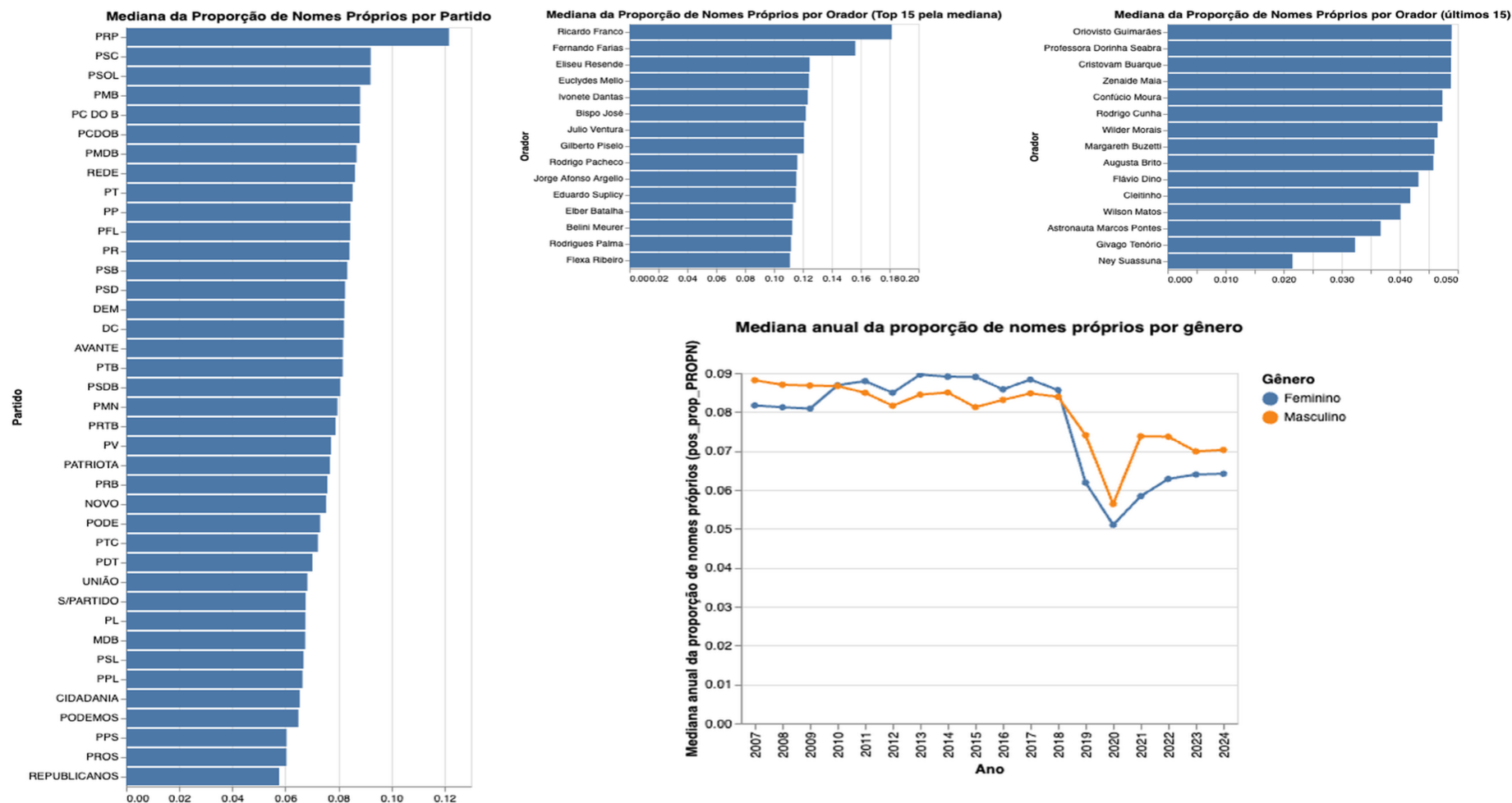
Os estudos de caso do bloco B apresentam tendências curiosamente inversas àquelas observadas nos estudos das classes do bloco A, tratadas na seção anterior. No agregado, legendas de esquerda encontram-se no topo da lista de partidos; as agremiações de direita, ao final. Com algumas exceções, como os senadores Rodrigo Pacheco, Flexa Ribeiro e Eduardo Suplicy, os oradores que usaram nomes próprios de maneira mais consistente em seus discursos foram parlamentares de poucas falas, ou de legislaturas mais antigas. Considerando os oradores que menos lançaram mão dessas classes gramaticais, a lógica se inverte. Quanto ao gênero do orador, vê-se novamente movimento diferenciado das senadoras a partir de 2019, quando passam a usar as classes do bloco B menos frequentemente que seus colegas homens, em um discurso típico.



Painel 8 – Proporção de adjetivos, preposições e nomes próprios ante o total de palavras. Mediana anual. Fonte: elaboração própria a partir de textdescriptives.

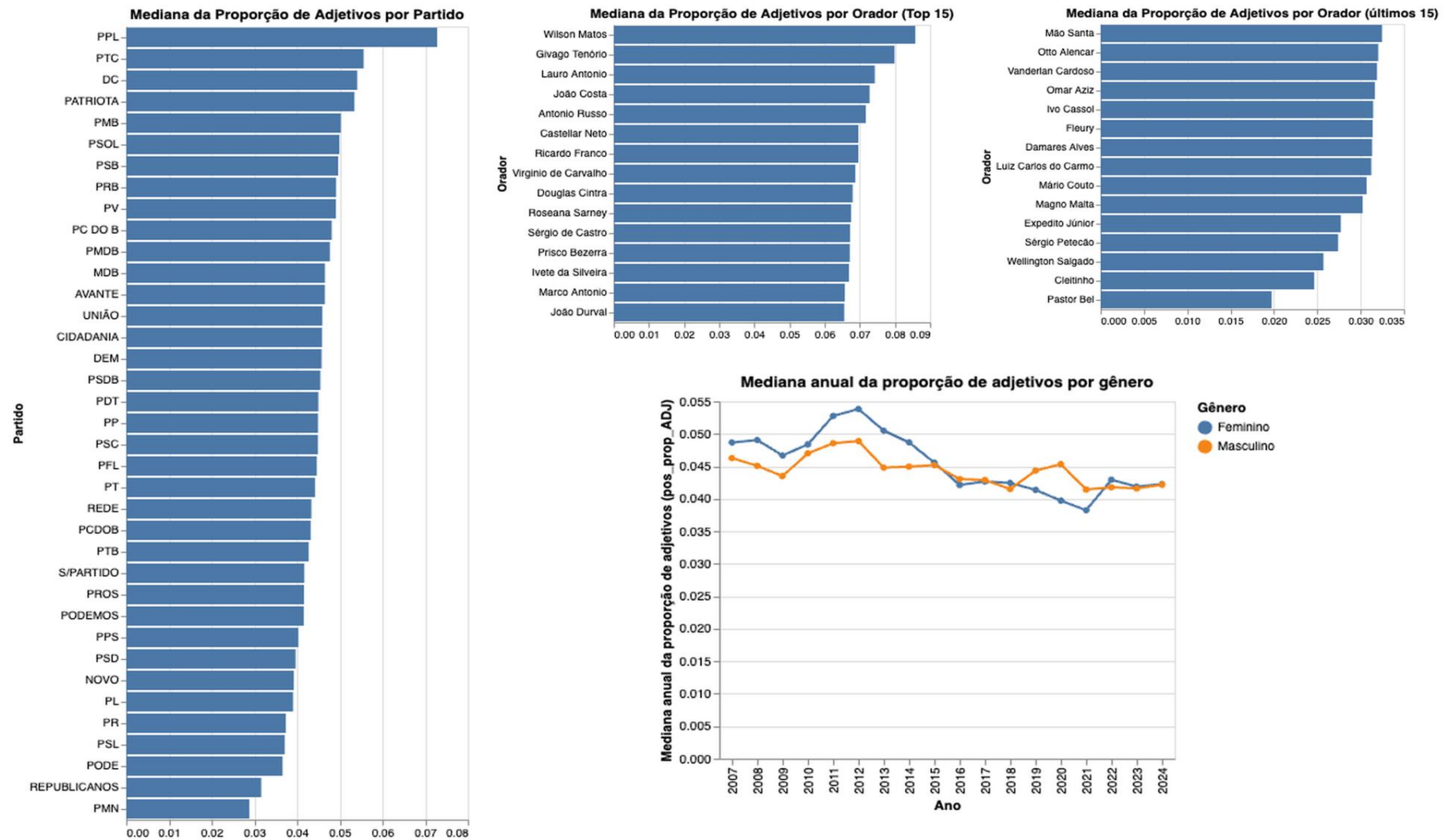
Estudos de caso

Nomes próprios



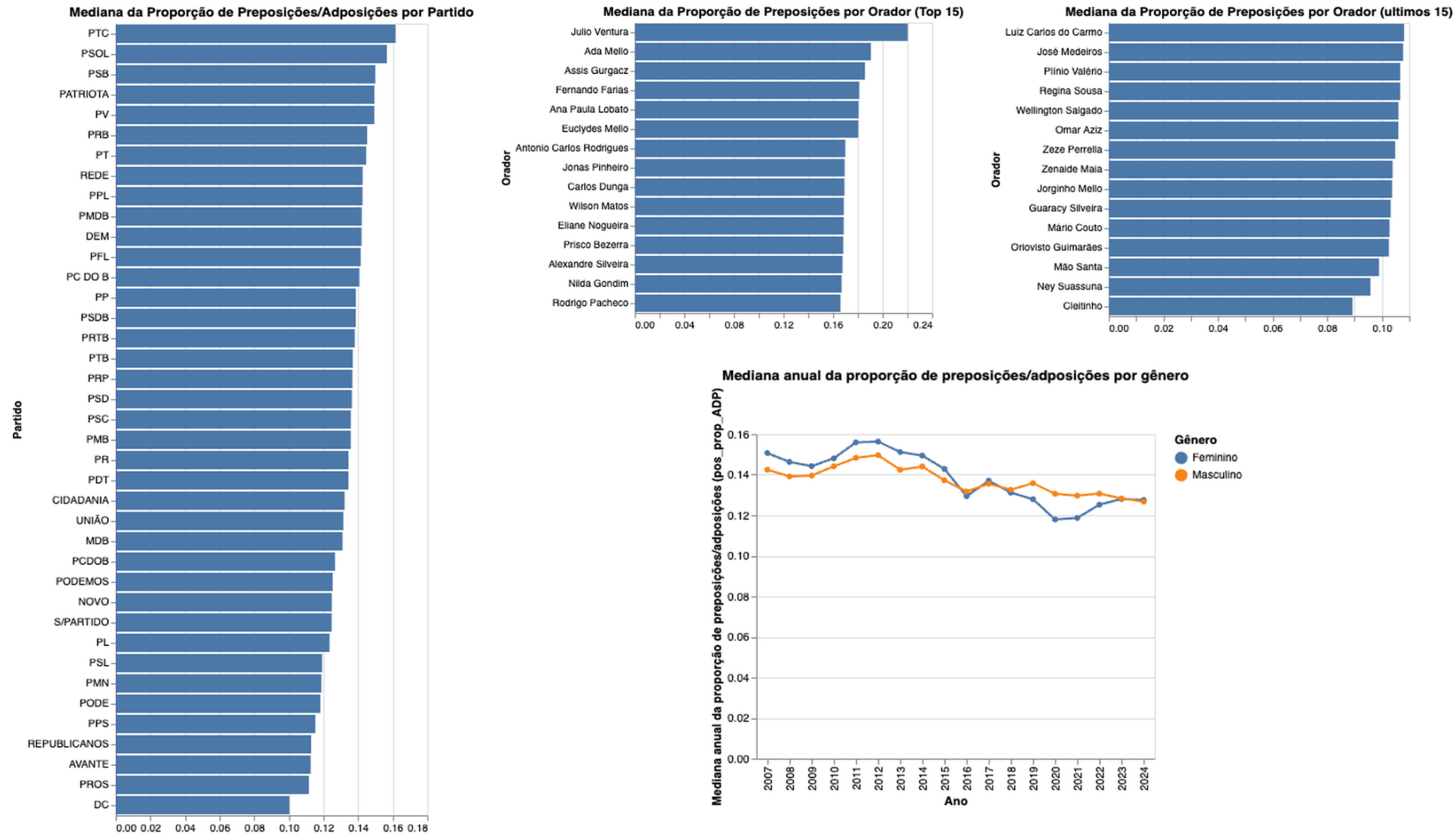
Painel 9 – Nomes próprios. Mediana agregada por partido, orador e gênero do orador. Fonte: elaboração própria a partir de textdescriptives.

Adjetivos



Painel 10 – Adjetivos. Mediana agregada por partido, orador e gênero do orador. Fonte: elaboração própria a partir de textdescriptives.

Preposições



Painel 11 – Preposições. Mediana agregada por partido, orador e gênero do orador. Fonte: elaboração própria a partir de textdescriptives.

2.3. Caso especial: pronomes pessoais retos

De todas as variações observadas neste trabalho, talvez a mais surpreendente seja quanto ao uso de pronomes. Como se trata de uma categoria limitada, do ponto de vista do léxico, vale fazer uma análise mais detalhada a respeito dessa classe gramatical. Para elaborar o próximo gráfico, utilizei dados extraídos por NILC-Metrix, método desenvolvido pelo Núcleo Interinstitucional de Linguística Computacional da Universidade de São Paulo. Esses dados dizem respeito a um tipo de pronome específico: os pronomes pessoais.

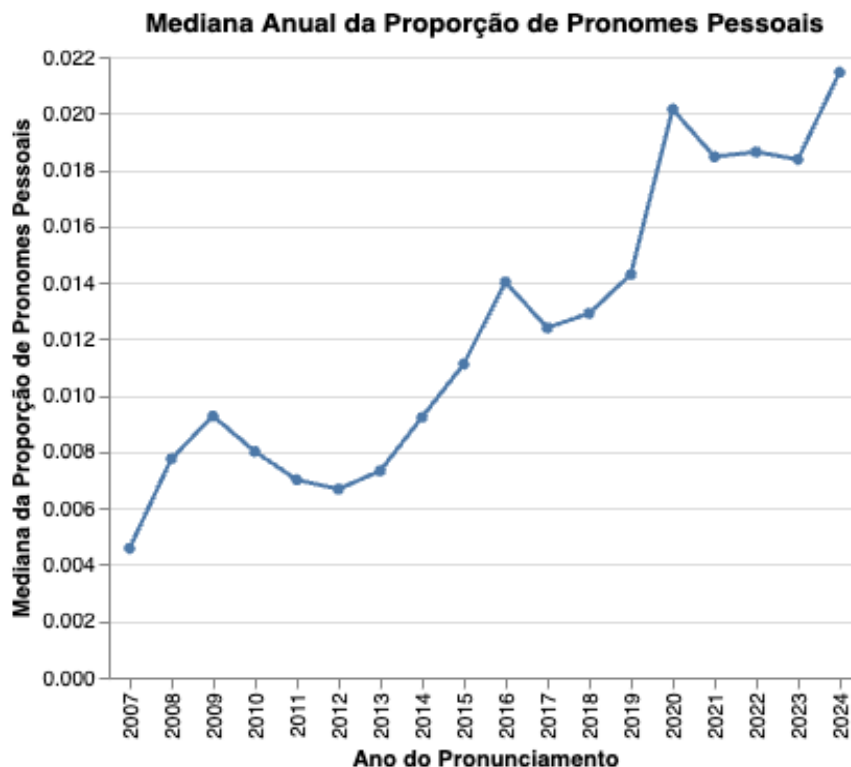


Figura 3 – Pronomes pessoais. Mediana anual. Fonte: elaboração própria a partir de NILC-Metrix⁸.

⁸ É necessário destacar um fato: a decisão por trabalhar apenas com discursos sem apartes foi tomada depois da reprodução dos scripts necessários à leitura por NILC-Metrix, feita por amostragem. Em outras palavras, a amostragem de 5% de discursos anuais com mais de 300 palavras ficou defasada, porque os apartes foram descontados depois da análise. Para fins de transparência, apresento os dados com e sem apartes, em cadernos Jupyter devidamente designados no repositório (Blanco, 2025a). A comparação entre os gráficos NILC-Metrix com e sem apartes a respeito de pronomes pessoais indica que não houve maiores divergências na trajetória temporal que fundamenta a argumentação deste artigo. Por causa dessa questão metodológica, este é o único gráfico deste artigo que apresenta dados obtidos por NILC-Metrix – mas a ferramenta tem vantagens importantes, como discutido em seções posteriores do trabalho.

A proporção desses pronomes em um discurso típico teve resultado surpreendente: quintuplicou ao longo da série histórica. O gráfico abaixo baseia-se na hipótese de que muitos desses pronomes seriam pronomes pessoais retos – “eu”, “tu”, “ele/ela”, “nós”, “vós”, “eles/elas”.

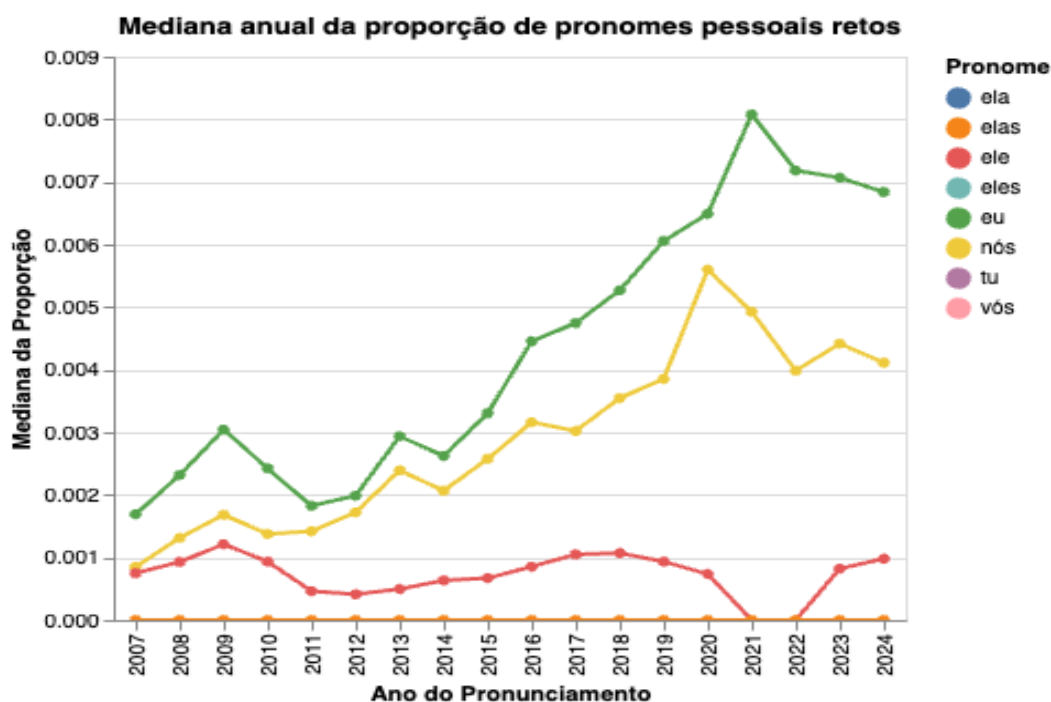
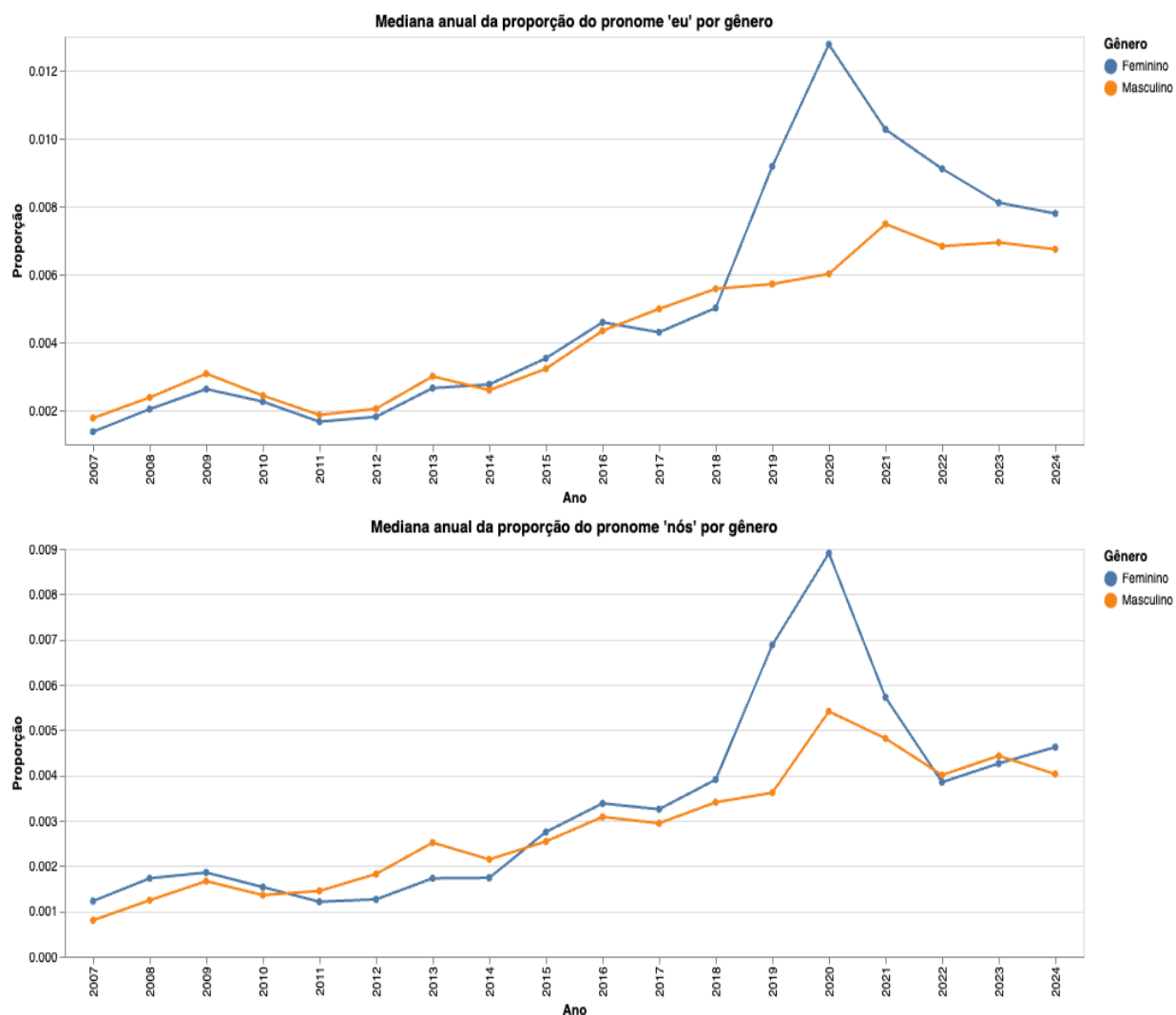


Figura 4 – Pronomes pessoais retos. Mediana anual. Fonte: elaboração própria a partir de algoritmo de pesquisa simples.

Percebe-se aumento expressivo nos pronomes da primeira pessoa – “eu” e “nós” – e alguma variação do pronome “ele” entre 2008 e 2011 e entre 2013 e 2020. Quanto aos demais, a mediana observada foi de zero, provavelmente porque a maior parte dos discursos não têm o pronome⁹.

A análise por gênero do orador esclarece um aspecto importante.

⁹ Embora os cadernos Jupyter que embasam minhas análises (v. repositório Kaggle) tenham gráficos com a média dos pronomes “ela” e “elas”, optei por não os citar neste trabalho por entender que a média, nesse contexto, é antes um retrato da agenda legislativa do que do comportamento discursivo dos parlamentares. Por exemplo: o ano de 2016 foi aquele em que o pronome “ela” atingiu o valor médio mais alto; mas isso não surpreende ninguém, tendo em vista o *impeachment* da ex-presidente Dilma Rousseff naquele ano.



Painel 12 – Incidência das palavras “eu” e “nós”. Mediana anual. Fonte: elaboração própria a partir de algoritmo de pesquisa simples.

Os gráficos acima acompanham as dinâmicas de gênero registradas no bloco A. O aumento do uso desses pronomes pelas senadoras acompanhou o dos senadores; mas, entre 2019 e 2021, o índice teve um salto entre as senadoras. Isso ocorreu com ambos os pronomes, mas a prevalência das senadoras no uso do “eu” foi ainda maior ao longo do tempo, prolongando-se até o final da série histórica.

2.4. Síntese dos resultados

Para simplificar a leitura, elaborei uma tabela que sintetiza os resultados descritos acima.

Dimensão	Métrica / Classe	Resultado agregado	Destaques por orador	Destaques por partido	Destaques por gênero
Leiturabilidade	<i>Flesch Reading Ease.</i>	Aumento da mediana anual a longo do período.	Valores medianos mais altos entre oradores recentes e com alta atividade em plenário.	Partidos de direita e centro-direita concentram maiores medianas.	Mediana masculina superior até 2019; feminina superior entre 2019–2023; convergência em 2024.
	<i>Gunning Fog.</i>	Redução da mediana anual.	—	—	—
	<i>SMOG.</i>	Redução da mediana anual.	—	—	—
	<i>ARI.</i>	Redução da mediana anual.	—	—	—
	<i>Coleman–Liau.</i>	Redução da mediana anual; cruzamento da linha de referência após 2016.	—	—	—
Classes gramaticais – Bloco A	Pronomes.	Aumento da proporção mediana (~40%).	Maiores valores entre oradores recentes e com muitos discursos.	Maiores medianas entre partidos de direita.	Crescimento semelhante até 2018; vantagem feminina entre 2019–2023.
	Verbos auxiliares.	Aumento da proporção mediana (~34%).	Oradores recentes concentram maiores valores.	Partidos de direita concentram maiores medianas.	Senadoras superam senadores entre 2019–2023.
	Advérbios.	Aumento da proporção mediana (~25%).	Oradores recentes figuram entre os maiores usuários.	Partidos de direita apresentam maiores medianas.	Diferença modesta; vantagem feminina entre 2016–2023.
Classes gramaticais – Bloco B	Adjetivos.	Redução da proporção mediana (>10%).	Oradores de legislaturas mais antigas figuram entre os maiores valores.	Partidos de esquerda concentram maiores medianas.	Senadoras lideram até 2015; inversão posterior.
	Preposições.	Redução da proporção mediana (>10%).	Oradores mais antigos concentram maiores valores.	Partidos de esquerda concentram maiores medianas.	Diferenças pequenas; inversão a partir de 2018.
	Nomes próprios.	Redução acentuada a partir de 2019.	Exceções institucionais (Rodrigo Pacheco).	Partidos de esquerda figuram no topo do ranking.	Uso feminino superior até 2018; masculino depois.
Caso especial	Pronomes pessoais (geral).	Aumento expressivo da mediana ($\approx 5\times$).	—	—	—
	Pronomes pessoais retos.	Crescimento concentrado na primeira pessoa.	—	—	Uso feminino superior nos pronomes da primeira pessoa a partir de 2019.

Tabela 2 – Síntese dos resultados. Fonte: elaboração própria com auxílio de inteligência artificial, modelo GPT-5.2.

A análise da tabela permite sistematizar os achados da pesquisa. Percebe-se uma convergência de trajetórias entre métricas de leitura Flesch e classes gramaticais do Bloco A (pronomes, verbos auxiliares e advérbios), tanto no resultado agregado quanto entre os agrupamentos por orador, partido e gênero.

Essas métricas parecem distinguir, em geral, três grupos: parlamentares mais recentes, com grande atividade discursiva; parlamentares de partidos de direita ou de centro-direita; e, no período entre 2019 e 2023, parlamentares mulheres. O caso especial, de pronomes pessoais retos na primeira pessoa, segue no mesmo sentido, ao menos quanto ao resultado por gênero.

Quanto às classes do Bloco B (adjetivos, preposições e nomes próprios), vê-se tendência praticamente oposta. Estas são classes cuja proporção se reduziu ao longo do período estudado; no banco de dados, são mais frequentes em discursos de oradores mais antigos, ou de partidos de esquerda. No início da série histórica, eram mais comuns nos discursos das senadoras; no período entre 2019 e 2023, porém, passaram a ser mais comuns entre parlamentares homens.

Como analisar essas tendências? Na seção seguinte, busco uma interpretação. Antes, no entanto, reforço algumas das prevenções necessárias a um estudo deste tipo.

3 DISCUSSÃO

3.1. Considerações metodológicas

Como observei em outro artigo baseado no mesmo método (Blanco, 2025b), os resultados de uma pesquisa como esta talvez sejam mais bem interpretados como uma fotografia modernista, tirada de um ângulo incomum. Ela revela aspectos da realidade, mas é incapaz de captá-la como um todo. O fato de que a pesquisa é baseada em dados e estatísticas descritivas não deve distrair o leitor dessa realidade; por isso, as questões levantadas por este trabalho devem ser recebidas como hipóteses, a serem testadas com pesquisas específicas.

Na verdade, a fotografia que este artigo apresenta tem sabidos problemas: questões de perspectiva, resolução, focos embaçados, sujeira na lente. Do ponto de vista da pesquisa, essas questões representam vieses – condições que, se não invalidam o trabalho, precisam ser antecipadas e reconhecidas. Gostaria de antecipar algumas delas.

O primeiro viés que merece destaque é o de coleta e registro dos dados. As requisições *http* que resultaram no banco de dados foram, no geral, eficazes, mas há pelo menos duas questões relevantes que precisam ser observadas.

A primeira é que a requisição foi complementada com a transcrição dos discursos em vídeo, comuns durante as sessões remotas da pandemia. Isso significa que há duas fontes distintas para os discursos do banco de dados, cada qual com seu modo de registro. Em contraste com as notas taquigráficas, as transcrições contêm apenas texto, sem observações de registro (como indicações do nome e partido do orador, sinalizações de contexto *etc.*). Além disso, embora a rapidez dos modelos de linguagem seja um fato relevante, é pouco provável que o resultado das transcrições supere, em precisão, o trabalho de um taquígrafo treinado.

A segunda questão diz respeito aos apartes. No início, considerei que ignorar os apartes implicaria negligenciar uma parte importante de como as pessoas falam no plenário do Senado Federal. Mas, ao analisar o assunto, percebi que os apartes determinariam leiturabilidade, proporções de pronomes, verbos auxiliares e, principalmente, nomes próprios (porque todos os apartes têm as respectivas marcas taquigráficas de autoria, que incluem o nome do aparteante), o que terminaria alterando demais as estatísticas. Além disso, observei que incorporar os apartes traria seu próprio viés. Como essas intervenções se reduziram em 90% ao longo do período estudado (Blanco, 2025b), a comparação estilística ao longo do tempo ficaria fatalmente distorcida; a existência de um aparte, afinal, introduz forçosamente elementos de personalidade nos discursos – justamente uma das premissas que compõem meu argumento, como explico mais adiante nesta seção.

A seleção de discursos sem apartes deu-se por meio de consulta aos metadados fornecidos pelo Portal de Dados Abertos do Senado Federal (o *script* consultou os discursos que tiveram apartes registrados e os subtraiu do total de falas do banco de dados). Este é um banco de dados consistente, mas algumas de suas convenções poderiam ter sido mais bem tratadas por mim, ao programar os *scripts*. Por exemplo, intervenções do presidente da sessão não são consideradas apartes, mas nem por isso deixam de constituir um diálogo. Isso fez com que, em alguns (poucos) casos, longos diálogos entre orador e presidente fossem considerados pelo sistema como uma fala monológica. Em retrospectiva, esse era um problema a ser abordado com uma técnica de pesquisa mais sofisticada,

como, por exemplo, filtros com expressões regulares (*regex*). Fica a lição para o próximo trabalho¹⁰.

Os vieses não se restringem aos problemas de coleta e registro. Sob o ponto de vista da análise, outras questões metodológicas podem vir à tona. Em primeiro lugar, vale refletir sobre a adequação das métricas de leitura a discursos. Estes são textos que, afinal, foram feitos para ser lidos em voz alta, cuja compreensão depende, em grande parte, do desempenho do orador: tom de voz, ênfase, prosódia *etc.* A esse respeito, observo que, para os fins deste trabalho, as métricas servem apenas como *proxies* de estilo – e, como tais, devem ser corroboradas por outras categorias de análise. Foi esse raciocínio que inspirou linguistas como Degani (2016) a analisarem discursos políticos com as métricas de *Flesch*, e é nesse contexto que busquei situar esta pesquisa.

Apesar de consolidadas, as métricas de leitura adotadas neste artigo não estão isentas de críticas. Allen e McNamara (2011), por exemplo, discordam do principal critério empregado por essas métricas, a extensão de palavras e/ou sentenças – seja em caracteres, seja em sílabas. A ideia central desses autores é a de que o critério da extensão é superficial e não alcança a compreensão do texto propriamente dita, que dependeria de outros fatores, como coesão e sintaxe (Allen; McNamara, 2011, p. 97). Esse raciocínio parece ainda mais adequado quando consideramos aspectos da língua portuguesa, cujas palavras costumam ter mais sílabas do que a inglesa. Tal fato motivou o Núcleo Interinstitucional de Linguística Computacional da USP a desenvolver NILC-Matrix, que adapta métricas ao português.

Do ponto de vista técnico, destacam-se problemas inerentes ao uso de *textdescriptives* para analisar notas taquigráficas. O fato de que o método analisa a morfossintaxe dos textos com base em *tokens* (unidades básicas de texto: palavras, pontuação *etc.*) faz com que, em termos de locuções, a contagem esteja subestimada. Além disso, algumas das abreviaturas empregadas pelos taquígrafos não são reconhecidas como tais. Por exemplo: em alguns casos, a pontuação que delimita a palavra abreviada é confundida com um ponto final, o que provavelmente influenciou a contagem de sentenças – e, portanto, as métricas de leitura. Em outros casos, o prenome do senador Omar Aziz foi

¹⁰ Outro problema consiste no fato de que a decisão final de excluir os apartes prejudicou a amostragem feita para a análise com NILC-Matrix, como explicado na nota de rodapé nº 5. Mas isso, como diz a nota, não mina o argumento central deste estudo.

confundido com um advérbio; o verbo “foi” foi lematizado na sua forma conjugada, e não como verbo “ser”; *etc.*

Meu principal passo para abordar essas questões foi escolher a mediana como medida básica – trata-se, afinal, de estatística resistente aos “pontos fora da curva”, os chamados *outliers* (por exemplo, um discurso específico com muitos nomes próprios). Mas isso não resolve tudo: mesmo em termos de mediana, a inclusão de apartes determinou a estatística de nomes próprios (porque os apartes são citados nominalmente). Em outras palavras, a mediana não elimina os vieses, apenas reduz seu impacto sobre as tendências gerais, o verdadeiro objeto desta análise. Além disso, busquei usar sempre as métricas para comparar o plenário do Senado consigo próprio, sem um referente externo. Salvo um trabalho mais aprofundado de limpeza e padronização de textos, não vejo como prudente usar o resultado dessas métricas para comparações, em termos absolutos, com outros bancos de dados.

Tampouco fiz teste formal de robustez, o que suporia recalcular as séries históricas com exclusão dos anos de pandemia, separação das fontes de transcrição ou inclusão dos discursos com apartes. Por essa razão, os resultados da pesquisa devem ser lidos como tendências descritivas dos discursos em Plenário ao longo do tempo. Ainda assim, a convergência entre diferentes métricas e classes gramaticais sugere que os padrões observados não se reduzem a um único artefato metodológico.

Feitas essas prevenções, os métodos empregados neste trabalho se complementam; restringindo-se o objetivo do estudo a um panorama geral, as aplicações são suficientes para subsidiar pesquisas futuras. Vamos a algumas delas.

3.2. Leiturabilidade e complexidade textual

Minha reflexão parte de um paradoxo. Tomando as métricas em conjunto, vemos uma associação entre as métricas do Bloco A – em termos de trajetória global e nos estudos de caso de gênero, oradores e partidos –, e entre estas e o índice de leiturabilidade *Flesch Reading Ease*. Em outras palavras, pronomes, advérbios, verbos auxiliares e leiturabilidade parecem caminhar juntos na evolução dos pronunciamentos feitos no plenário do Senado Federal ao longo do período estudado. Algumas avaliações específicas demonstram isso.

A clusterização¹¹ dos discursos corrobora o esquema dos blocos A e B. O critério adotado como subsídio ao algoritmo foi a proporção de pronomes, preposições e verbos auxiliares¹². O resultado dividiu os pronunciamentos em dois clusters, o e 1, sendo o cluster 0 um grupo com alta proporção de preposições, e o cluster 1, um conjunto com alta proporção de pronomes e verbos auxiliares (além de baixa proporção de preposições).

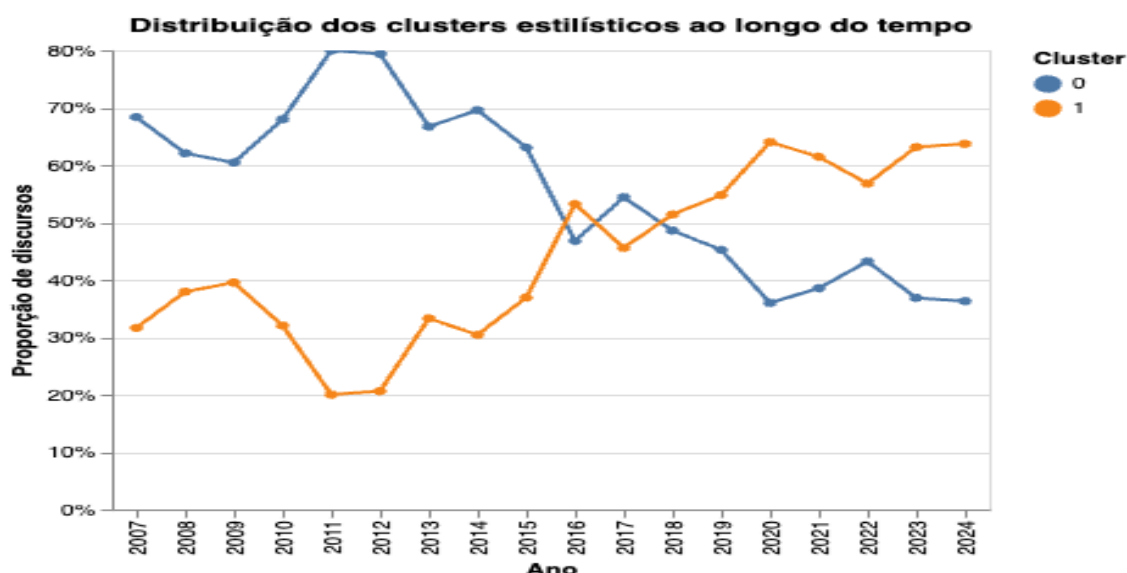


Figura 5 – Clusters estilísticos baseados na proporção de pronomes, preposições e verbos auxiliares. Fonte: elaboração própria, com base em textdescriptives.

Grosso modo, o **cluster 0** corresponde ao **Bloco B** (alta proporção de adjetivos, preposições e nomes próprios), enquanto o **cluster 1** corresponde ao **Bloco A** (alta proporção de pronomes, verbos auxiliares e advérbios). O cluster 1, antes uma parcela minoritária dos discursos, cresceu ao longo da série histórica, alcançando 65% das falas de 2024; por sua vez, o cluster 0 fez o movimento contrário, chegando a 35% desse mesmo ano. O gráfico abaixo dá outra medida da evolução dos clusters, comparando a variação global dos elementos analisados neste trabalho entre 2007 e 2024:

¹¹ Procedimento de *machine learning* que divide um conjunto de dados em grupos de características semelhantes, de forma não supervisionada (isto é, sem a definição prévia das categorias). Sobre o tema: JAMES, Gareth *et al.* **An introduction to statistical learning: with applications in Python**. Cham: Springer, 2023, p. 520.

¹² Esse critério foi adotado por meio de um algoritmo que ordenou os clusters possíveis segundo métricas de coesão dos grupos, como Silhouette Score, Davies-Bouldin, Calinski-Harabazs e inércia. Ver mais em: DUARTE BLANCO, Pedro. **Repositório do projeto “falando_nela”**. Kaggle, 2025. Disponível em: <<https://www.kaggle.com/datasets/pedro-duarteblanco/falando-nela>>. Acesso em: 19 nov. 2025.

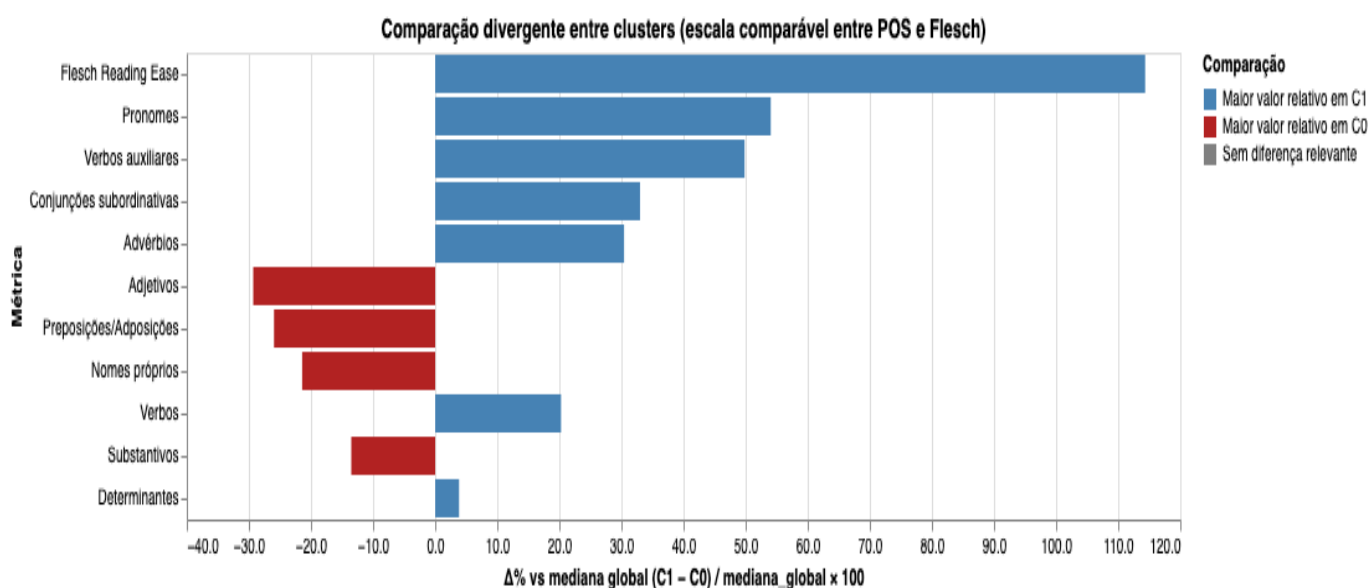


Figura 6 – Comparação divergente entre os clusters. Fonte: elaboração própria, com base em textdescriptives.

Há semelhança marcante entre esse gráfico, que apresenta o contraste entre os clusters estatísticos, e o gráfico do painel 3 (v. 3.2, Classes gramaticais), que faz a mesma comparação quanto à proporção de classes gramaticais. Em outras palavras, quando dividi as classes em blocos, procedi de maneira intuitiva; mas o método estatístico dos clusters deu reforço estatístico a essa percepção.

O problema é que, quanto à complexidade textual, esses blocos não formam grupos coerentes. Um exemplo disso é a relação entre leitura e pronomes. Por um lado, a documentação de NILC-Matrix indica forte associação entre leitura e complexidade, o que é bastante intuitivo: quanto mais fácil de ler, mais simples é o texto. Por outro, o mesmo arquivo refere que a interpretação básica de outros componentes trabalhados neste estudo, como a proporção de pronomes, correlaciona-os positivamente com a complexidade textual¹³. No entanto, há manifesta correlação entre pronomes e leitura *Flesch*; é o que se vê na figura abaixo:

¹³ A respeito da métrica de leitura *Flesch Reading Ease*, a documentação indica que “quanto maior o resultado da métrica, menor a complexidade textual”. Quanto aos pronomes, estes “exigem um esforço para descobrir seu referente e, portanto, quanto maior o resultado da métrica, maior a complexidade textual”, segundo a referência (“NILC-Matrix Doc”, [S.d.]).

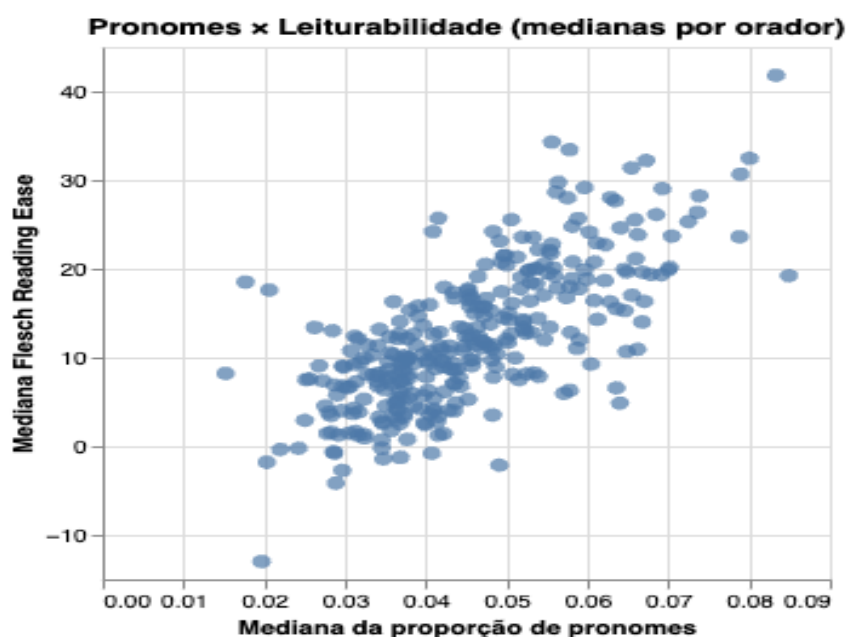


Figura 7 – Leiturabilidade versus proporção pronominal, por orador.
Fonte: elaboração própria a partir de textdescriptives.

No gráfico de dispersão acima, cada ponto é um senador. Vê-se que há poucos oradores para quem o aumento na proporção de pronomes não está acompanhado de um incremento na leiturabilidade *Flesch* dos respectivos discursos. **Enfim: os resultados da pesquisa mostram que os discursos se tornaram mais simples, sob o aspecto da leiturabilidade, mas mais complexos sob o aspecto da proporção pronominal.** Isso significa que a ideia de complexidade textual, apesar de importante, não é suficiente para descrever o que definiu a evolução discursiva em plenário ao longo da série histórica. É possível que isso se deva à própria fluidez do conceito.

No contexto dos discursos, talvez as métricas de leiturabilidade sejam mais bem interpretadas não só como uma medida de complexidade textual, mas também como uma técnica retórica de ênfase. O exemplo abaixo, fragmento de discurso da senadora Damares Alves, ilustra esse caso – principalmente em relação ao comprimento das frases:

O intérprete está aqui do meu lado, senhores. Ele está interpretando para a comunidade surda que está lá em cima. E o guia-intérprete é aquele que faz o trabalho para o surdo-cego. Quando eu era criança, Senador Omar, uma criança surdo-cega não ia para a escola. Uma criança surda já era difícil ir. Imagine uma criança surdo-cega. Graças a esses profissionais, elas estão nas escolas hoje. Mas elas não estão

só nas escolas, elas estão nas universidades. Elas estão fazendo pós-graduação. Nós já temos alunos aqui em Brasília fazendo mestrado, doutorado, sendo surdo-cego. E o guia-intérprete surdo-cego não trabalha desse jeito. O cego tem que ficar ali juntinho com a mão dele. E é libras tátil que a gente fala. Não é fácil. (Pronunciamento de Damares Alves, 2023).

Percebe-se, na passagem acima, o emprego combinado de frases curtas com repetições anafóricas (“elas estão”, “uma criança”) e gradação (“elas não estão só nas escolas, elas estão nas universidades”, “mestrado”, “doutorado”), produzindo uma homenagem expressiva aos intérpretes da Linguagem Brasileira de Sinais Tátil. A trajetória das métricas de leiturabilidade reflete o aumento desse tipo de discurso.

3.3. Pessoalização e democratização

3.3.1. Pessoalização

De certa forma, o achado da seção anterior era previsível. A complexidade de um texto tem muitas facetas; não se pode reduzi-la ao número de sílabas ou ao tamanho das sentenças, como dizem os críticos das métricas de leiturabilidade citados acima (Allen; McNamara, 2011).

Nesse contexto, a proporção de pronomes, dado cuja variação se destacou entre os resultados deste artigo, dá o que pensar. Os pronomes têm uma alta taxa de correlação com várias outras classes – positiva com aquelas que caracterizam o bloco A, mencionado anteriormente, e negativa em relação às do bloco B. As tabelas abaixo mostram a correlação de Pearson¹⁴ entre as trajetórias anuais agregadas quanto às proporções de pronomes e outras classes gramaticais (isto é, o co-movimento dessas variáveis no tempo), sendo que as classes em itálico foram aquelas examinadas com mais detalhes no trabalho:

¹⁴ A técnica de Pearson mede a intensidade e a relevância de uma correlação entre dados, a partir da divisão da covariância entre estes pelos respectivos desvios-padrão. O “r” é um número de -1 a 1, que indica correlação negativa ou positiva (0 significa ausência de correlação). O “p” mostra a probabilidade de dados não relacionados atingirem o “r” indicado; assim, significa a relevância dessa correlação medida (“Pearson Correlation – an overview ScienceDirect Topics”, [S.d.]).

Variável	r de Pearson	p-valor
Advérbios	0,990	< 0,001
Verbos auxiliares	0,981	< 0,001
Verbos	0,926	< 0,001
Conjunções subordinativas	0,906	< 0,001
Determinantes	0,851	< 0,001

Tabela 3 — Correlações de Pearson positivas com a proporção de pronomes.
Fonte: elaboração própria com base em textdescriptives.

Variável	r de Pearson	p-valor
Preposições	-0,947	< 0,001
Nomes próprios	-0,864	< 0,001
Adjetivos	-0,819	< 0,001

Tabela 4 — Correlações de Pearson negativas com a proporção de pronomes.
Fonte: elaboração própria com base em textdescriptives.

As tabelas indicam que a divisão dos discursos em dois blocos (um, ascendente, e outro, descendente) tem respaldo estatístico, como mencionado. Também mostram que a prevalência de pronomes pode ser adotada como um dos critérios dessa divisão, sendo um dos principais aspectos do bloco que se expandiu nos últimos anos da série histórica. Mas é possível ir além.

Talvez um conceito que explique melhor a dinâmica do plenário seja o de registros. Estes são, nas palavras dos linguistas Douglas Biber e Susan Conrad, “variedades (comunicativas) definidas em termos de parâmetros situacionais gerais”, que “podem ser definidas em termos linguísticos” (Biber; Conrad, 2019, p. 2). A relação entre os componentes linguísticos e paralinguísticos de um registro é, essencialmente, funcional (Biber; Conrad, 2019, p. 7) – em outras palavras, o padrão linguístico adotado decorre da circunstância e do propósito da situação comunicativa. Há registros que dizem respeito a conversas informais, prosa acadêmica, texto jornalístico *etc.*

Uma taxa alta de pronomes é uma das marcas do registro conversacional (Biber; Conrad, 2019, p. 94-95); o fato de que essa taxa teve um incremento em meio à queda nos apartes em plenário (Blanco, 2025b) reforça a ideia de que o discurso se destina ao “auditório implícito” das redes sociais, por exemplo. Com quem mais os oradores estariam conversando?¹⁵

O resultado de pesquisas por expressões específicas nos textos dos discursos ajuda a lançar luz a essa questão. Por exemplo, é comum que os senadores façam em seus discursos uma apóstrofe ao público que acompanha a sessão virtualmente, pela TV ou pela Rádio Senado (no início do período estudado) ou, ainda, pelas redes sociais. Em regra, a apóstrofe, bastante formulaica, incorpora a expressão “que nos acompanham”, como nesta passagem do senador Astronauta Marcos Pontes:

Sr. Presidente, Sras. e Srs. Senadores e todos aqueles que nos acompanham através da TV Senado e das redes sociais, esta discussão é extremamente importante para o país, para o futuro do país como um todo. (Pronunciamento de Astronauta Marcos Pontes em 09/04/2024, 2024).

Como mostra o gráfico abaixo, esse tipo de interpelação passou a crescer a partir de 2007, com referências aos veículos de comunicação do Senado – e teve um salto em 2013. Este foi o ano em que a política nacional entrou de vez na rede social Facebook, na esteira das mobilizações conhecidas como “Jornadas de Junho”, coordenadas primordialmente pelas redes. Após algumas oscilações conjunturais, a frequência dessa fórmula no texto dos discursos estabilizou-se, provavelmente porque as redes deixaram de ser novidade na vida pública, e voltou a crescer em 2023, ano em que tomaram posse os vencedores das eleições do ano anterior.

¹⁵ Há, obviamente, uma chance de os oradores estarem conversando com o presidente da sessão, cuja manifestação não é considerada aparte pelo Portal de Dados Abertos do Senado Federal. Mas parece uma hipótese remota, em termos de tendência.

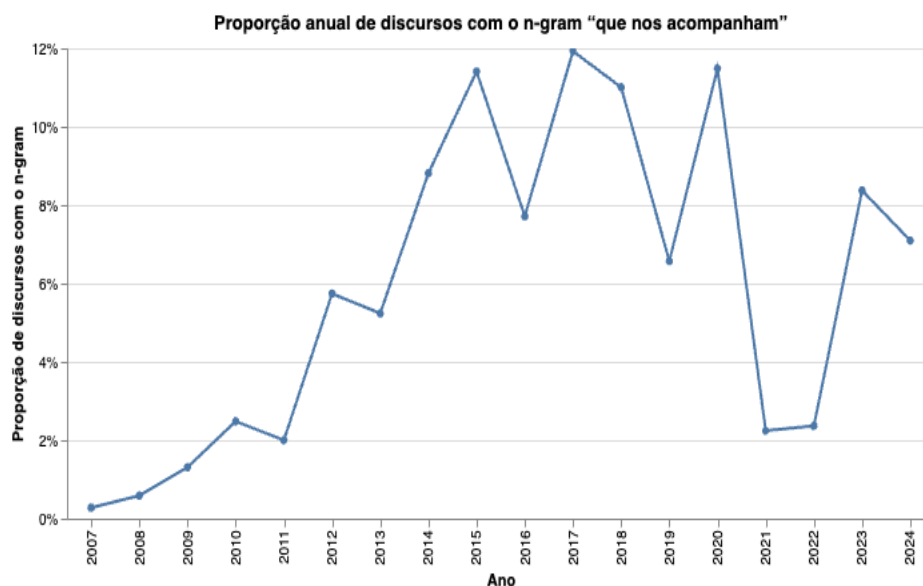


Figura 8 – Proporção de discursos com a expressão “que nos acompanham”.
Fonte: elaboração própria.

Em outro exemplo, a percentagem de discursos com pelo menos um ponto de interrogação saltou de pouco mais de 40% do total anual de discursos para 60%.

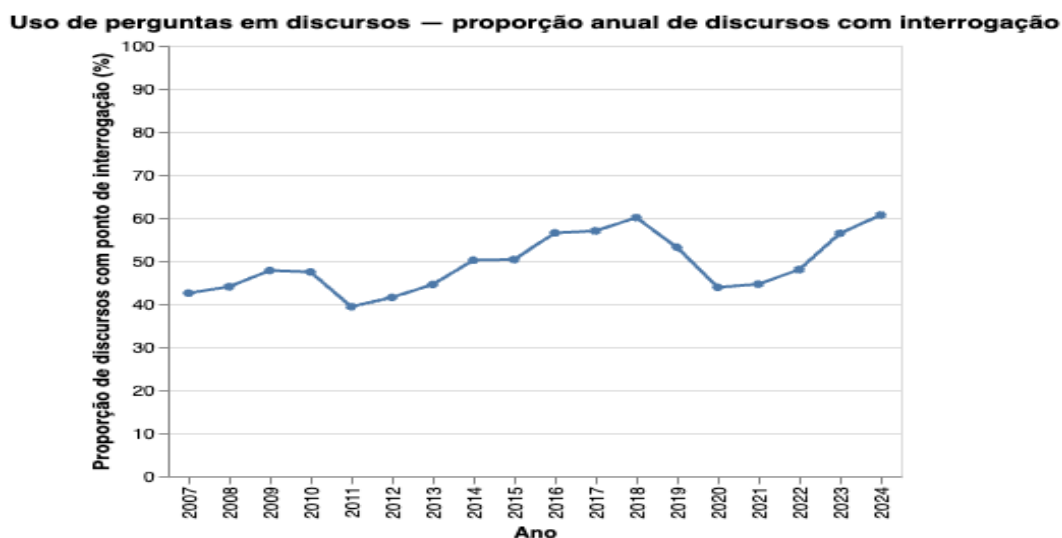


Figura 9 – Discursos com pelo menos um ponto de interrogação. Percentagem.
Fonte: elaboração própria.

Tendo em vista que o espaço amostral se constitui de discursos sem apartes, as interrogações encerram ou perguntas retóricas, ou questionamentos dirigidos ao presidente da sessão (cuja resposta, como vimos, não é considerada aparte pela taxonomia do Senado Federal). A primeira hipótese parece mais provável. Nesse caso, a pergunta retórica seria uma figura de linguagem destinada a provocar efeitos de páthos no público *online*.

O crescimento expressivo da proporção de pronomes pessoais retos (v. Figura 3, supra) permite desenvolver a hipótese conversacional. O aumento no uso dos pronomes da primeira pessoa, um dos achados mais singulares desta pesquisa, enseja reflexão. Na forma singular (“eu”), o pronome reto de primeira pessoa pode denotar autopropaganda, ou reivindicação de protagonismo (*credit claiming*) no jogo político (Mayhew; Arnold, 2004, pp. 75–77). Na forma plural (“nós”), pode servir para reforçar vínculos ou mobilizar bases. Em uma terceira explicação, os pronomes de primeira pessoa podem decorrer de uma questão geracional, sendo mais comuns entre parlamentares recentes – situação analógica à dos fóruns *online* referenciados por Biber e Conrad (2019, p. 198)¹⁶; *etc.*

Todas essas interpretações se relacionam com um fato básico a respeito dos membros do Congresso Nacional – o desafio da representação popular em um ambiente competitivo, caracterizado, em grande parte, pelo alto grau de incerteza (Mayhew; Arnold, 2004, p. 47). Nesse contexto, a trajetória dos pronomes retos indica uma versão parlamentar do processo de “pessoalização sintética” teorizado pelo linguista Norman Fairclough: “a simulação de conversas privadas, ‘cara-a-cara’, no discurso voltado ao público de massa” (Fairclough, 1996, p. 98). Esse processo implicaria, na visão do autor, “a subordinação de todos os demais aspectos de semântica e prática discursiva a fins instrumentais” – em outras palavras, a comunicação estratégica por excelência. Não é necessário grande esforço para adaptar esse raciocínio à conjuntura observada neste estudo. As redes sociais, ao possibilitarem a reprodução de trechos dos discursos no *smartphone* pessoal dos cidadãos, levariam essa lógica ao seu auge: os parlamentares competem – com seus pares e com outros atores da chamada “economia da atenção” –, e por isso lançariam mão de métodos vistos na televisão comercial ou no entretenimento, por exemplo. O estilo “clipável” a que me referi em outro estudo (Blanco, 2025b) trata justamente disso.

¹⁶ O argumento é o de que usuários novatos desses fóruns usariam mais pronomes, porque seu conhecimento do tema discutido se limita à própria experiência. Em contraste, usuários mais experientes teriam mais conhecimento e usariam registro mais informativo. Expandindo esse raciocínio, pode ser que parlamentares novatos sintam mais necessidade de apresentar suas realizações como agente político, por exemplo.

3.3.2. Democratização

Outra das teorias de Fairclough (1996, pp. 201-207) estabelece que o discurso social passa por uma “democratização”, caracterizada em parte por um uso casual e horizontalizado (isto é, sem estabelecer hierarquia explícita entre falantes e ouvintes) da língua. Esse parece ter sido um dos movimentos observados nesta pesquisa.

Advérbios

Como vimos, o processo acima descrito não se compõe apenas de aumento na proporção de pronomes. Além da leituraabilidade, já discutida, outras métricas convergem para produzir o efeito teorizado por Fairclough. Os advérbios também compõem o registro conversacional: na visão dos linguistas Biber e Conrad (2019, p. 97), palavras dessa classe são comuns em um discurso baseado em frases curtas (ou seja, com leituraabilidade alta), o que caracteriza um texto com esse registro.¹⁷ Vale destacar aqueles que mais variaram, em uma comparação entre os anos inicial e final do período:

¹⁷ É possível que os autores tenham concebido um diálogo propriamente dito, com as partes respondendo umas às outras – por isso a referência a frases curtas. Minha hipótese é a de que os advérbios terminados em “-mente” desempenham função enfática e apelativa, e, ao fazê-lo, podem compor um registro mais interativo. Ao evocar linguagem técnica e burocrática, essas palavras constituiriam uma forma “fácil” de reivindicar competência ou legitimidade institucional. Seria, em outras palavras, uma espécie de hipercorreção estilística, como no uso contemporâneo de “particularmente”, palavra que se encontra entre os advérbios cuja frequência mais cresceu entre 2007-2024. Esse raciocínio contraria a ideia de “democratização” do discurso nos termos formulados por Fairclough; mas é bom lembrar que a conceituação desse autor, bastante crítica, tem lá suas ironias, e o próprio Fairclough concede que a “democratização”, aqui, pode ser apenas cosmética (Fairclough, 1996, p. 201). De qualquer maneira, é bom lembrar que há diferentes estratégias para se buscar persuasão e popularidade, e nem todas passam pelo ideal da linguagem simples – sobretudo em um país desigual como o Brasil, em que a educação bacharelesca tem um peso cultural forte. Embora a escolha seja, em muitos casos, questionável do ponto de vista estilístico, um advérbio terminado em “-mente” pode dar “brilho” (superficial) ao orador sem comprometer o entendimento da frase.

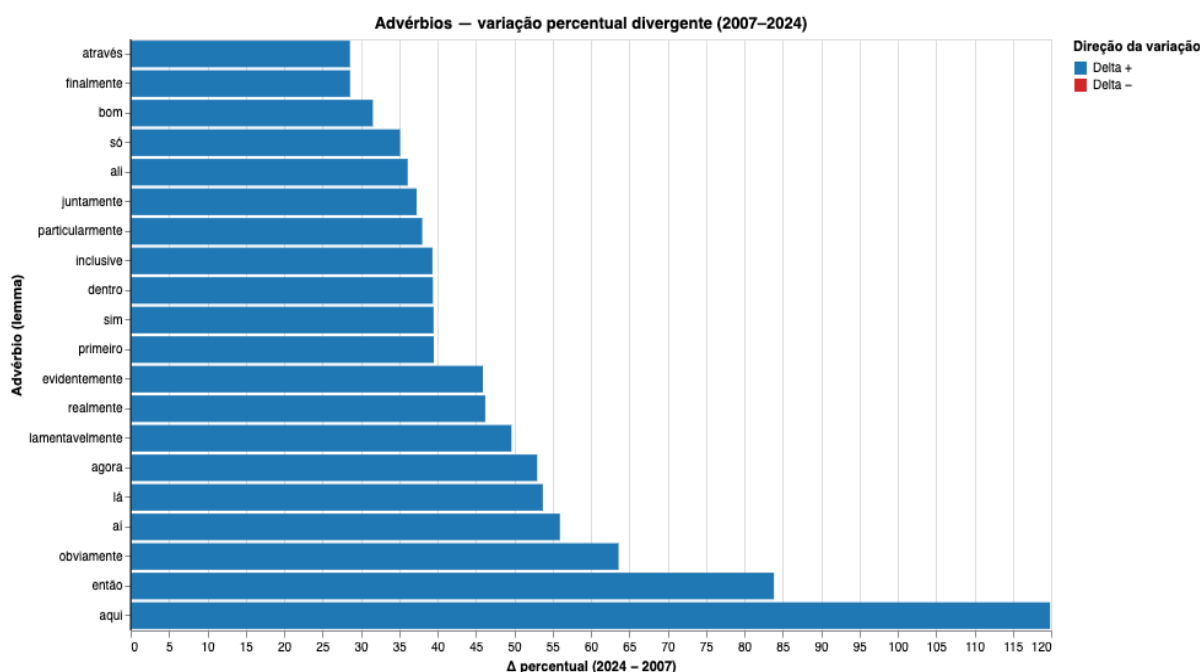


Figura 10 – Advérbios. Principais variações 2007-2024, em por cento. Fonte: elaboração própria.

Percebe-se uma série de marcadores de oralidade – “só”, “aí”, “ali” –, palavras do registro informal brasileiro contemporâneo (e.g., “particularmente”) e operadores argumentativos de uso comum (“evidentemente”, “obviamente”). Estas são palavras que reforçam a ideia de democratização da linguagem – são persuasivas para o grande público, mas não deixam de ser compreensíveis.

Verbos auxiliares

Outra das classes cuja proporção aumentou consideravelmente, os verbos auxiliares – “ter”, “haver”, “ser”, “ir” *etc.* – influenciam o processo de democratização porque introduzem elementos de informalidade no discurso. Por exemplo, o verbo “ir” tem sido cada vez mais usado em construções de futuro perifrástico (e.g., “vou fazer”), como neste trecho do senador Fabiano Contarato:

*Esse PL nº 2.564 **vai alcançar** 2,4 milhões de enfermeiros, técnicos de enfermagem, auxiliares de enfermagem e parteiros.*
(Pronunciamento de Fabiano Contarato em 12/05/2021, 2021).

O gráfico abaixo mostra a evolução temporal desse uso do verbo “ir”. Percebe-se que, enquanto o uso perifrástico mais do que dobrou, os demais usos tiveram um aumento pequeno durante o período estudado:

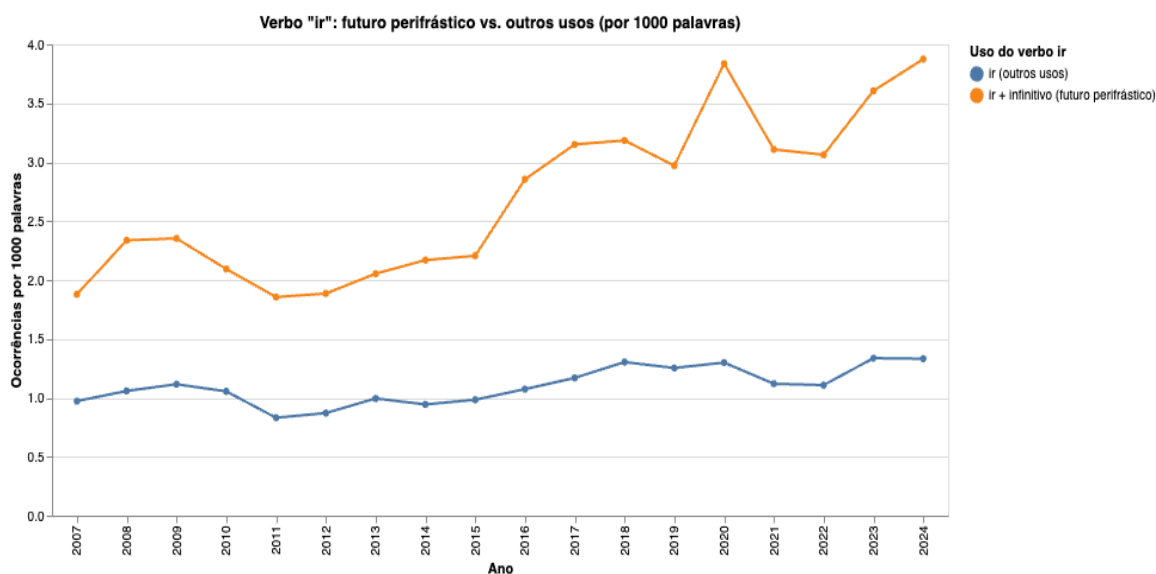


Figura 11 – Verbo “ir”. Usos por mil palavras. Fonte: elaboração própria.

Por sua vez, a variação no uso do verbo “ser” segue trajetória ascendente, embora menos intensa do que no exemplo anterior:

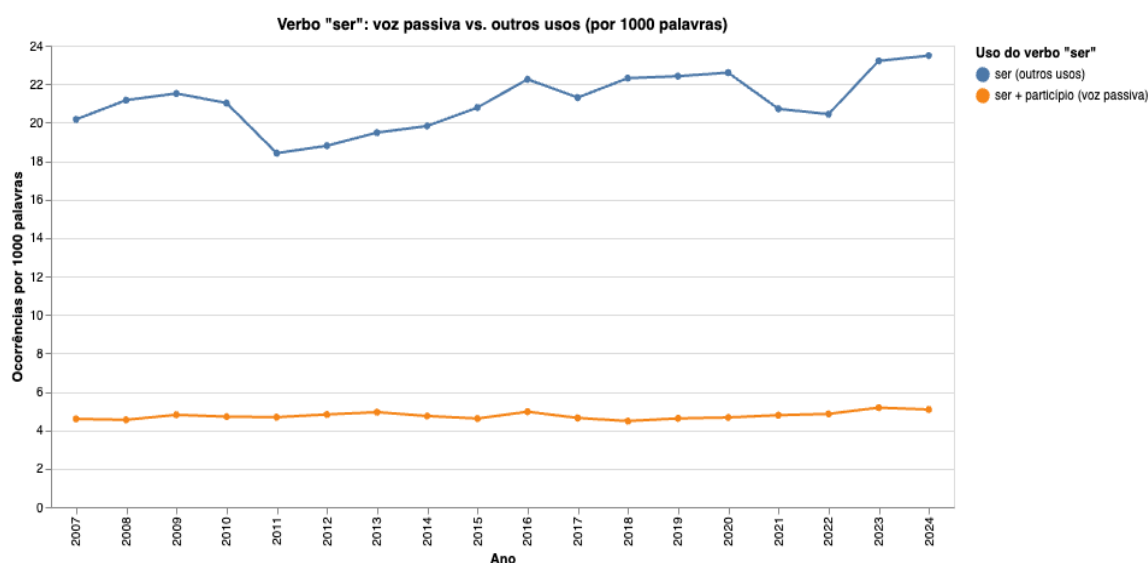


Figura 12 – Verbo “ser”. Usos por mil palavras. Fonte: elaboração própria.

Percebe-se que o emprego do verbo “ser” em construções de voz passiva permaneceu estável ao longo da série histórica; mas os demais usos tiveram alta. Possivelmente, trata-se de uso avaliativo do verbo, complementado por adjetivo ou substantivo. A passagem abaixo, da senadora Rose de Freitas, ilustra esse tipo de uso:

*Esse é um assunto que interessa ao país, às mulheres, que são mais de metade da população. Como é que... **Vetado** (sic) **já é um absurdo*** (Pronunciamento de Rose de Freitas, 2022).

Uma busca pelos adjetivos e substantivos que complementam essa forma verbal sugere algumas tendências. Os adjetivos, listados no gráfico abaixo, sugerem um padrão hiperbólico, constituído por palavras como “inegável”, “absurdo”, “inadmissível”, “óbvio” e “inaceitável”.

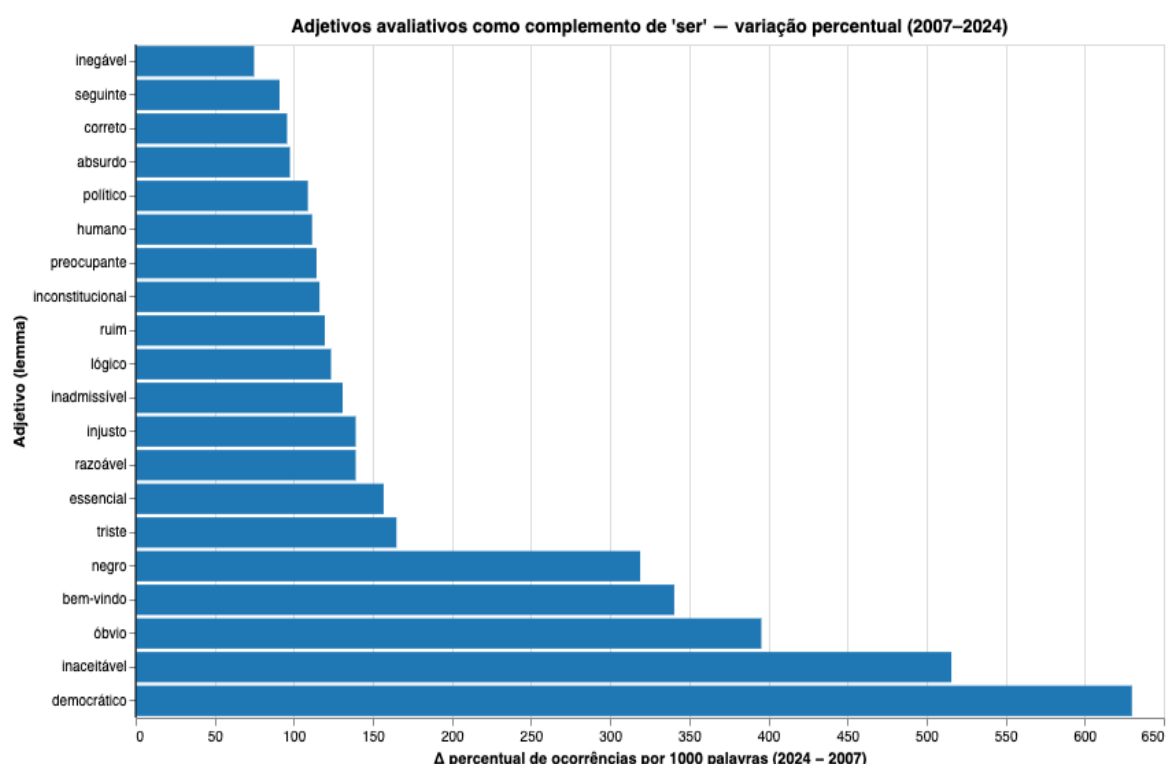


Figura 13 – Adjetivos que complementam o verbo “ser”. Variação 2007-2024.
Fonte: elaboração própria.

Por sua vez, os substantivos parecem guardar relação mais estreita com o ambiente político-legislativo; mesmo assim, referências a “golpe”, “brincadeira”, “tragédia” e “crime” podem reforçar o padrão anteriormente descrito¹⁸.

¹⁸ É claro que, em muitos casos, essas palavras descrevem fatos veiculados no noticiário ou apurados pela justiça, o que reflete o acirramento da conjuntura nos últimos anos. Em outras palavras, a linguagem teria ficado mais hiperbólica porque a vida se tornou mais hiperbólica.

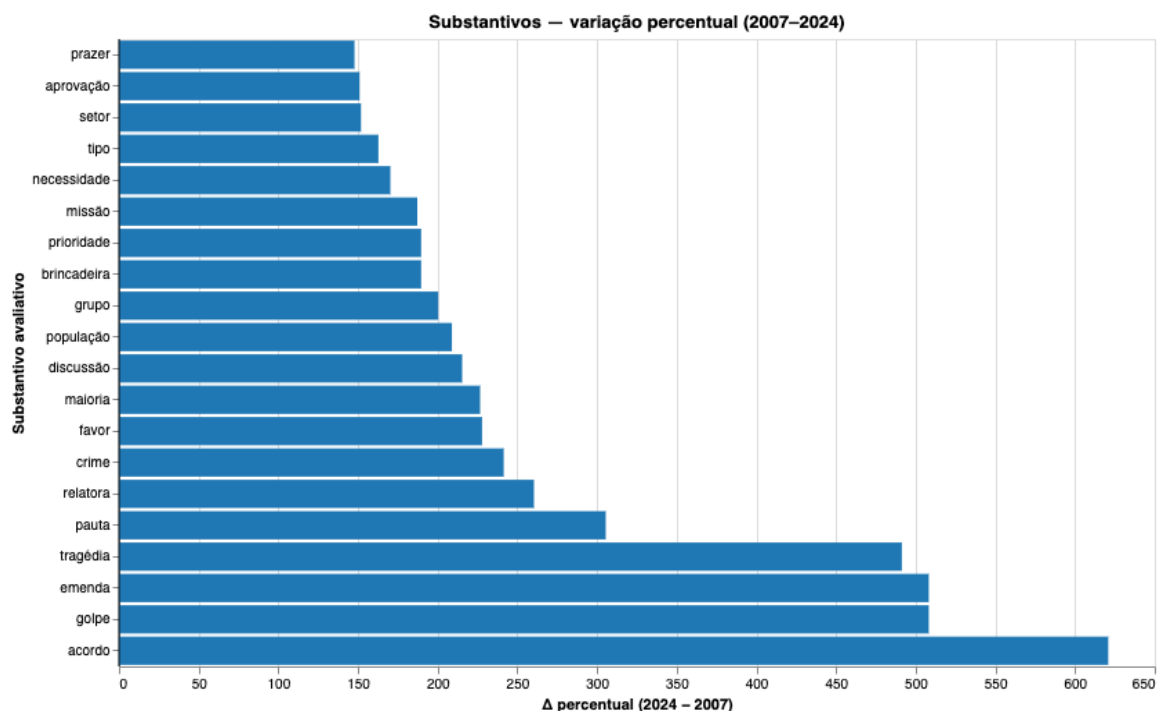


Figura 14 – Substantivos que complementam o verbo “ser”. Variação 2007-2024.
Fonte: elaboração própria.

Declínio do bloco informativo-referencial

Todas as trajetórias descritas como democratização do discurso compõem um movimento mais amplo, de expansão de falas “conversacionais” – isto é, pessoais, informais, apelativas. Nesse contexto, é importante voltar a atenção para as formas cuja proporção caiu ao longo da série histórica, pois isso ajuda a compreender a evolução do período como um todo. Trata-se das classes do bloco B (ou, no arranjo estatístico, cluster o): adjetivos, preposições e nomes próprios. Biber e Conrad associam essas classes a registros de discurso escrito, ou – no caso dos adjetivos e nomes próprios – informativo (Biber; Conrad, 2019, p. 272).

Como visto nos resultados da pesquisa (v. 3.2, Classes gramaticais e suas frequências), adjetivos e preposições seguiram trajetória parecida, registrando tendência de queda a partir de 2012. Por sua vez, os nomes próprios seguiram estáveis até os anos da pandemia de covid-19, quando tiveram declínio vertiginoso. Este pode ter resultado do viés de registro descrito acima, tendo o algoritmo de transcrição dos vídeos “confundido” um substantivo próprio com um substantivo comum, por exemplo.

Após esse período, no entanto, a métrica de nomes próprios permaneceu dois pontos percentuais abaixo do nível original, de 9%. Há várias causas

possíveis para esse fenômeno. Uma retórica mais pessoalizada – e, portanto, menos institucional – dá menos oportunidades para agradecimentos e medidas. Isso pode ter sido reforçado pela redução global no tempo de fala disponível aos senadores (Blanco, 2025b), ou pelo formato de clipes voltados às redes sociais, comuns durante a pandemia.

Em relação às preposições, vale ressaltar que o movimento de queda não alcançou todas as palavras dessa classe gramatical. Entre as mais comuns, o declínio foi expressivo apenas na preposição “de”, como mostra o gráfico abaixo:

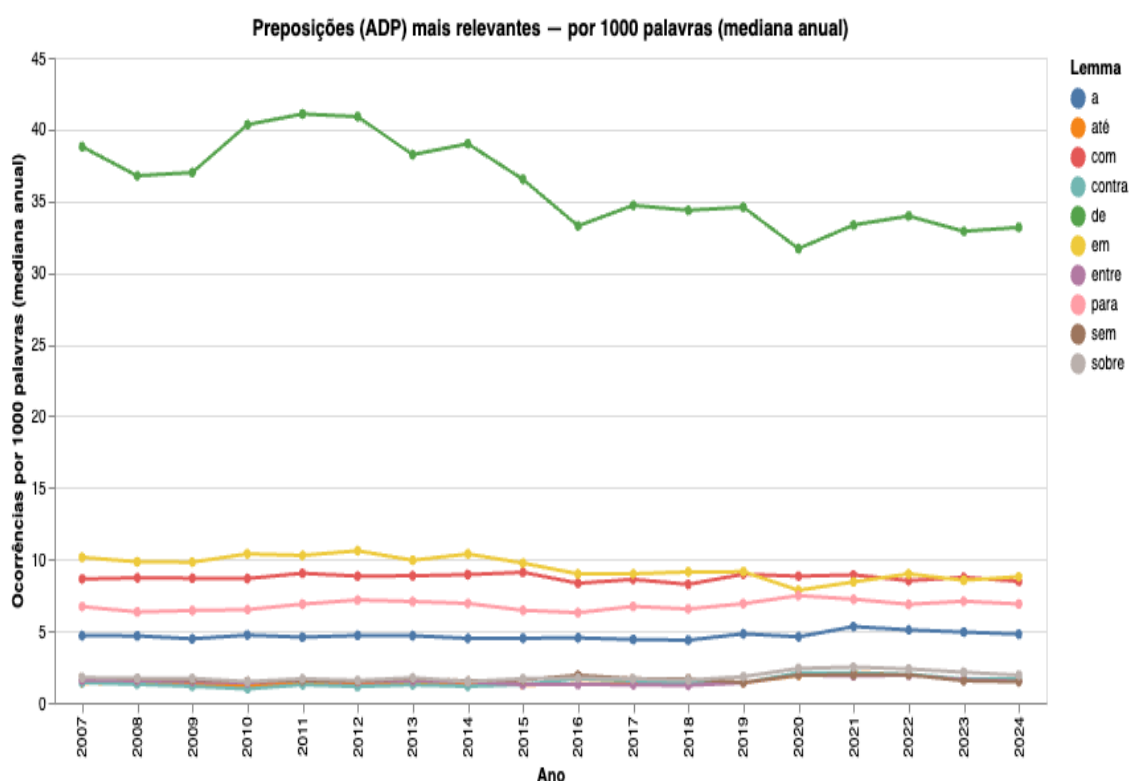


Figura 15 – Principais preposições. Série histórica. Fonte: elaboração própria.

Esse resultado não é trivial. A palavra “de” costuma ser associada a nominalizações¹⁹ e cadeias sintáticas complexas de linguagem escrita (Castilho, 2010, p. 221). A redução na sua ocorrência pode significar um esforço por um discurso mais concreto, direto e fluido.

¹⁹ Nominalização é a “[t]ransformação de um verbo em substantivo, mediante sufixações (...) Os nomes nominalizados projetam argumentos, preservando a estrutura gramatical do verbo de base”. (Castilho, 2010, p. 685). Em geral, considera-se que essas palavras, comuns em linguagem burocrática, tornam os textos mais prolixos e abstratos.

“Era o que (eu) tinha a dizer”

Um caso simples ilustra alguns dos desenvolvimentos mencionados nesta seção. Trata-se da expressão “era o que tinha a dizer”²⁰, cuja ocorrência se reduziu drasticamente ao longo do período estudado, como mostra o gráfico abaixo:

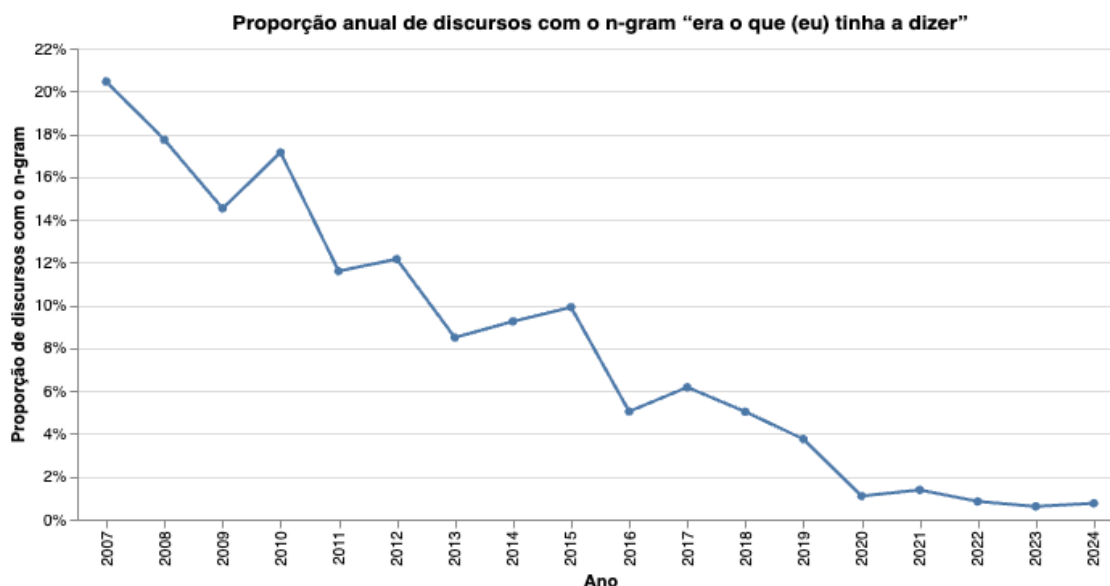


Figura 16 – “Era o que (eu) tinha a dizer” (soma das incidências com e sem o pronome pessoal). Por mil palavras. Fonte: elaboração própria.

Uma das fórmulas mais clássicas do plenário, inspirada nas figuras de modéstia da retórica romana, a frase – que chegou a surpreendentes 20% dos discursos de 2007 –, é historicamente usada para concluir discurso lido em forma de monólogo – isto é, sem apartes²¹. Muitas vezes, a peça tem alto conteúdo informativo, como neste exemplo do senador Romero Jucá:

*O SR ROMERO JUCÁ (PSDB – RR) – Sr. Presidente, Sr^{as}. e Srs. Senadores, ninguém duvida que nossa vida está cada vez mais imersa em tecnologia. **Do computador e da impressora, de onde saíram as páginas que tenho em mãos**, ao painel eletrônico, que registra a nossa presença e totaliza os votos nesta Casa, passando pelo equipamento da TV Senado, que leva os debates aos lares dos brasileiros, estamos cercados de tecnologia. (...)*

Pelos últimos números tabulados pelo MEC, estamos falando de mais de 124 mil pessoas capacitadas para operar laboratórios de informática, espalhados em mais de 4.600 escolas de 1.700

²⁰ Também formulada com o pronome pessoal reto, “era o que eu tinha a dizer”.

²¹ Menos de 10% dos discursos com essa expressão tiveram apartes (Blanco, 2025a).

*municípios de todas as regiões do País. São resultados preliminares de um programa em fase de implantação. O **atingimento** do objetivo final, **qual seja**, a diminuição da “exclusão digital” dos alunos da rede pública, só vai ser constatado no atual governo.*

Vitorioso o Presidente Fernando Henrique, vitorioso também o meu partido, PSDB, que alcançou resultados significativos semelhantes em programas desenvolvidos em âmbito estadual ou municipal. É o caso das “Infovias para o Desenvolvimento”, a interligação das escolas técnicas estaduais do Ceará em uma rede de teleconferências, financiada numa parceria entre Governo do Estado e Governo Federal. É outra experiência que vale a pena ser conhecida e vem servindo de modelo para outros Estados da Federação.

Era o que tinha a dizer, Sr. Presidente.

Muito obrigado. (Pronunciamento de Romero Jucá, 2003).

O trecho foi extraído de pronunciamento anterior à série histórica abordada neste trabalho, mas se assemelha a inúmeros casos do início do período. Nele, abundam marcas de registro escrito (a começar pela própria admissão do orador de que tinha “páginas em mãos”): frases longas, expressões formais (“qual seja”), numerais sem contexto analítico (e.g., comparação com figuras concretas), nominalizações (e.g., “atingimento”) *etc.* O fim abrupto, em tom eufemístico, parece pouco afeto à retórica do nosso tempo, em que a influência da televisão reforça a importância de uma conclusão expressiva (resumida no jargão em inglês “*clinger*”) (Lehrman, 2010). Parece, enfim, uma leitura voltada para registro nos anais do Senado, talvez dirigidas a assessorias de imprensa – e não para um auditório.

3.4. Discurso, redes sociais, populismo

Como conciliar as interpretações acima com o declínio do diálogo em plenário, constatado em trabalho anterior (Blanco, 2025b)? Na verdade, as ideias se complementam: por meio de pronunciamentos mais compreensíveis, pessoais e democráticos (e menos informativos), os parlamentares buscariam não um debate com seus pares, mas um vínculo direto com a população, por meio das redes sociais. Segundo essa hipótese, existe, sim, interatividade – mas não dentro do plenário. Esse é um aspecto importante para entendermos melhor a evolução dos discursos no período.

Discurso e populismo

Uma das hipóteses que buscam explicar esse fenômeno é a do ressurgimento do populismo. Em texto bastante crítico, Tzvetan Todorov lista alguns dos traços observados no discurso populista:

*O formato de apresentação do populismo é a demagogia; seu conteúdo se reduz a algumas constantes. Acima de tudo, o populista se recusa a afastar-se do aqui e agora, de indivíduos concretos; foge das abstrações, das distâncias, da duração, favorecendo, em vez disso, o concreto, o próximo e o imediato. Enquanto o democrata ideal procura inspiração naquilo que Rousseau chamou de vontade geral – uma construção hipotética do que, em qualquer momento, seria melhor para o conjunto do povo –, o populista se dirige à multidão com a qual está em contato: um comício em um espaço público, as muitas pessoas que escutam uma transmissão pela televisão ou pelo rádio. Exige-se do democrata a defesa de valores impopulares, a defesa de sacrifícios, porque ele também se preocupa com as gerações futuras; já o populista explora as emoções necessariamente efêmeras do momento. Em nome do interesse geral, o democrata está disposto a intervir em favor das minorias no país, enquanto o populista prefere ater-se às certezas da maioria.*²²

A descrição de Todorov menciona alguns dos resultados deste trabalho – notadamente, referenciais de circunstância concreta, como “aqui” e “agora” (v. 3.2.1 e 4.4, em relação a advérbios), e a suposição de um auditório implícito composto do grande público. Em contraste, a inclusão de “indivíduos” contraria, em tese, a redução observada nos nomes próprios (v. 3.2.2, quanto a nomes próprios). No geral, porém, o autor acerta ao sublinhar o papel das emoções na retórica populista; o despertar destas na população pode ser um dos efeitos buscados pela política com as mudanças relatadas neste trabalho, algo sugerido por padrões hiperbólicos observados, por exemplo, na seção 4.4 (v. adjetivos avaliativos associados ao verbo “ser”).

A hipótese populista ajuda a explicar algumas das dinâmicas observadas nos estudos de caso. Uma das principais características atribuídas ao populismo é o conservadorismo essencial, associado à negação da existência de uma luta

²² Citação do livro original em inglês. A tradução coube ao autor, com auxílio do modelo de inteligência artificial GPT 5.2.

de classes (Bobbio; Matteucci; Pasquino, 2022, p. 981). No recorte analisado, os partidos e os parlamentares situados à direita no espectro político apareceram entre aqueles com valores mais altos nas métricas associadas à democratização do discurso (v. 3.2.1 e 3.2.2, Estudos de caso)²³. Outra é o personalismo, muitas vezes cultivado à margem dos canais de comunicação tradicionais (Feres Júnior; Gagliardi, 2021). As tendências de pessoalização e de simplificação da linguagem convergem com essa característica.

No entanto, a identificação dessas tendências com o populismo esbarra em alguns obstáculos. Em primeiro lugar, o conceito de “populismo” não descreve algumas facetas do novo uso da linguagem (como demonstra o caso dos nomes próprios, referido em parágrafo anterior). Segundo, o termo não veicula informação sobre o elemento funcional da nova forma de falar, a saber, o uso instrumental das redes sociais. Terceiro, o adjetivo “populista”, politicamente carregado, leva à confusão histórica, pois evoca movimentos políticos distantes no tempo – varguismo, peronismo *etc.*

Talvez a principal razão para descartar a equivalência entre as práticas populistas e a forma emergente de discursar remeta aos estudos de caso sobre gênero. Em todos os resultados da pesquisa (v. 3.2.1 e 3.2.2), o período entre 2018 e 2023 registrou uma proporção mais alta de senadoras (ou mais baixa, no caso do bloco B, que agrupa discursos com registro informativo) nas métricas adotadas neste trabalho. Esse período coincide com aquele em que as senadoras buscaram entre si uma interação prioritária, por meio de apartes, e parece ter tido como centro a institucionalização da bancada feminina no Senado Federal (Blanco, 2025b, p. 24). O alinhamento desses dois processos sugere a correlação entre o uso da língua e a estratégia de plenário, possivelmente com a finalidade de demonstrar atuação concertada ao público externo. A adoção da nova forma de discursar pelas senadoras sugere uso desse tipo de fala em favor do pluralismo – ao contrário do que sustenta Todorov, na passagem citada acima. Enfim, o ressurgimento do populismo, embora relevante como contexto histórico, não é capaz de definir, por si só, a evolução formal dos discursos em plenário.

²³ Uma possível explicação para isso é a aproximação entre acadêmicos e especialistas (em outras palavras, pessoas que usam linguagem complexa por ofício) e os partidos de esquerda (ou o distanciamento entre esses atores e os partidos de direita) – fato constatado por intelectuais militantes como Chantal Mouffe (Mouffe; Colonna d’Istria, 2018, p. 13).

Registro versus estilo

Outra discussão pertinente consiste no debate sobre se a nova tendência constitui de fato um registro ou um estilo, em termos linguísticos. Segundo Biber e Conrad (2019, p. 15), a diferença entre essas duas classificações consiste no elemento funcional, isto é, o caráter necessário da variação linguística adotada; enquanto o registro teria características essenciais à situação comunicativa, o estilo responderia antes a maneirismos e preferências estéticas.

É difícil abordar essa questão sem reconhecer limites importantes deste artigo – notadamente, o fato de que o trabalho mostra apenas correlações e hipóteses, sem comprovar causas e consequências das trajetórias observadas na pesquisa. É possível que os parlamentares tenham, com apoio de assessorias de comunicação, aproximado seus discursos ao que é visto nas redes sociais; mas é igualmente possível que os senadores empossados na segunda metade da série histórica tenham vencido as respectivas eleições justamente por (entre outros fatores) falarem da maneira como falam, isto é, adaptando suas palavras às novas formas de circulação do discurso político.

Nesse contexto, a mudança qualitativa no funcionamento do plenário (com a extensão do auditório implícito às redes sociais) leva à ideia de que o modo de falar emergente seria, em alguma medida, condição de sobrevivência política, sobretudo para políticos em início de carreira. Em outras palavras, em um “momento populista” (Mouffe; Colonna d’Istria, 2018, p. 8), em que a “telegenia” dos candidatos consiste em um fator importante de sucesso (Taguieff, [S.d.]), a forma emergente de discursar comporia um registro.

Democratização e democracia

Uma terceira discussão que vem à tona concerne à relação entre democracia e o processo descrito neste trabalho. Quão democrática é a democratização do discurso? Não é pergunta fácil. Como mencionado na nota de rodapé nº 16, os conceitos de “pessoalização” e “democratização” parecem conter ironia – como destaca o próprio autor, Norman Fairclough, ao observar que a pessoalização do discurso é “sintética”, isto é, um artifício.

Para responder a essa pergunta, é preciso refletir sobre o caráter democrático das novas formas de comunicação. O papel das mídias sociais no regime é, no mínimo, ambivalente. Por um lado, as redes são operadas por algoritmos obscuros, cujo retorno muitas vezes é manipulado ou influenciado por transações financeiras. Além disso, têm um efeito polarizante sobre a agenda

(Lanier *apud* Blanco, 2025b), com impactos sobre o dia a dia da vida pública. A hipótese confirmada de que os discursos em plenário têm sido substituídos por vídeos destinados às redes (Blanco, 2025b) é um exemplo eloquente disso.

Por outro lado, as redes dão mais alcance ao discurso político: melhoram o acesso da população aos debates e permitem o surgimento de novas pautas e novos candidatos. O caso das pautas femininas, reconhecidas e impulsionadas no Senado Federal a partir de fins dos anos 2010, ilustra esses fatos. Além disso, algumas das tendências observadas encontram eco em iniciativas institucionais de caráter democratizante. É o caso da Política de Linguagem Simples, definida por lei de 2025 e adotada internamente pelo próprio Senado. Embora a lei seja posterior à série histórica analisada neste trabalho, é possível que a adesão às novas formas de se expressar tenha se expandido sob o mesmo impulso político que determinou a aprovação das novas regras para a administração pública (que incluem orientações como “usar frases curtas, em ordem direta e na voz ativa”; “usar palavras comuns, de fácil compreensão”; “evitar o uso de substantivos no lugar de verbos” *etc.*) (Agência Senado, 2025).

Discurso e ideologia

Outro ponto concerne à orientação ideológica predominante, entre os pronunciamentos que adotam o novo registro linguístico. O fato de que parlamentares conservadores se destacaram nos estudos de caso indica que os partidos de direita e centro-direita foram pioneiros na adaptação do seu discurso aos novos tempos, em tendência que parece ter sido mundial.

Um dos lugares-comuns do colunismo político de hoje em dia, a ideia de que a esquerda se encontra presa ao jargão acadêmico (Carville, 2025) trata de tema que pode ter influenciado os resultados deste trabalho. O argumento, obviamente, traz consigo a sua própria réplica: Degani (2016) vê no uso do registro por políticos de direita um sinal de anti-intelectualismo. De qualquer maneira, não há nada de necessário nesse achado. A ascensão de políticos de grande apelo popular, como Zohran Mamdani, atual prefeito de Nova Iorque, sinaliza que a esquerda mundial busca novas formas de expressão (Pilkington, 2025)²⁴.

²⁴ Nesse contexto, destaca-se o comentário do novo prefeito de que seu grupo político devia “ouvir mais e ensinar menos”.

Uma tipologia

Em busca de uma tipologia que contemplasse todas essas considerações – registro *versus* estilo, populismo *versus* democracia –, foi priorizada uma definição de caráter descritivo. Segundo essa definição, o discurso emergente, correspondente ao bloco A/cluster 1 (verbos auxiliares, pronomes, advérbios, leiturabilidade Flesch), pertence ao **registro conversacional mediado**. Em linhas gerais, um discurso elaborado sob esse registro é informal, pessoal, apelativo. Por meio dele, os oradores priorizam a interação potencial com as redes sociais à interação real com seus pares. O principal condicionante formal seria o videoclipe como meio de circulação. O discurso se reveste de um potencial simbólico relevante, embora seja pouco aproveitado no debate parlamentar propriamente dito, como destacado no TD nº 355 (Blanco, 2025b).

Em contraste, o discurso antigo, correspondente ao bloco B/cluster 0 (adjetivos, nomes próprios, preposições), pertence ao **registro referencial-informativo**. Por meio dele, os oradores priorizam algumas tradições da Casa: são mais formais, cerimoniosos, e destinam seus discursos ora aos colegas, ora à imprensa – ou mesmo aos anais do Senado, como era comum no início da série histórica (lembramos que os discursos remetidos aos anais sem pronunciamento em plenário foram excluídos da análise). Nesses casos, o pronunciamento prioriza exposição de conteúdo; há pouco cuidado com linguagem simples, leiturabilidade ou performance. O alcance do discurso depende, em grande parte, da repercussão deste ante a mídia tradicional. Enfim, essa tipologia tem a vantagem de representar a transição verificada durante a série histórica – independentemente se os oradores são populistas ou democratas, progressistas ou conservadores.

Uma discussão final diz respeito aos impactos decorrentes da transição descrita neste artigo. As novas tendências podem ser vistas como sinais de mudanças importantes no funcionamento do Senado Federal. A ascensão do registro conversacional mediado rompe com tradições importantes do Senado: trata-se, afinal, da Casa revisora do processo legislativo, espaço de reflexão e ponderação política, reduto de eloquência onde peroraram figuras como Francisco Sá, o “rouxinol de Diamantina”, e Rui Barbosa, a “águia de Haia” *etc.* (Carneiro *apud* Rosa, 2021, p. 102; Vido, 1960).

A discussão, no entanto, transcende questões de linguagem. É possível que a trajetória dos discursos signifique uma mudança fundamental não só de forma,

mas também do conteúdo. Especificamente, a ideia de que o parlamentar deve discursar “falando a língua do povo” parece contrária à representação política idealizada por, entre outros, Edmund Burke:

O Parlamento não é um congresso de embaixadores de interesses diferentes e hostis, cujos interesses cada um deve assegurar, como um agente e um defensor, contra outros agentes e defensores; mas o Parlamento é uma assembleia deliberativa de uma nação, com um interesse, o da totalidade – em que nenhum propósito local, nenhum preconceito local, deveria guiar, exceto o bem comum, resultante da razão geral da totalidade. (Burke, 2012).

Segundo esse raciocínio, um registro conversacional, essencialmente imitativo, não traduziria a representação como delegação de confiança enfatizada por Burke, que pressupõe o parlamentar como agente de deliberação racional em favor do bem comum, mas uma representação à maneira de espelho sociológico (Cotta *apud* Bobbio; Matteucci; Pasquino, 2022) – isto é, identitária. É possível que esta seja uma resposta discursiva ao desgaste da confiança no Congresso Nacional observado nos últimos anos, principalmente na segunda metade dos anos 2010 (Corporación Latinobarómetro, 2024). Mas é também possível que isso decorra de estratégias deliberadas de diferenciação, mais ou menos populistas – a crítica “aos políticos tradicionais”, “a tudo isso que está aí” *etc.*

4 CONCLUSÃO

Ao longo do estudo, foi desenvolvido o seguinte argumento: ao longo da série histórica, os discursos no plenário do Senado passaram a ter cada vez mais traços de um registro linguístico que chamei de “conversacional mediado”. Tornaram-se mais simples, pessoais e democratizados, no sentido pretendido por Fairclough, isto é, sem marcas de hierarquia entre falantes e ouvintes. Essas transformações necessárias sugerem uma adaptação funcional do discurso às novas formas de circulação deste – notadamente, vídeos nas redes sociais –, cuja influência sobre o ambiente do plenário foi constatada no Texto para Discussão nº 355 (Blanco, 2025b).

Constata-se que esse ponto de vista foi sustentado com os dados obtidos e com análises à luz de teorias linguísticas. Ao final da pesquisa, porém, restaram algumas dúvidas. De que forma essas mudanças se traduzem em democratização efetiva? É preciso falar como o povo para representar o povo? Qual o impacto

desse processo sobre a qualidade das deliberações, ou sobre o ambiente político como um todo?

Em 2010, o senador José Sarney declarou, em seu discurso de posse como presidente do Congresso Nacional, que os parlamentos tinham perdido “*aquele charme romântico que os acompanhou durante todo o século XIX até metade do século XX (...) [quando] os plenários podiam ser tocados pela palavra, pelo delírio e pelo encantamento dos oradores* (Sarney, 2010).” A observação, um tanto melancólica, reflete a má imagem que a oratória política assumiu com o passar dos anos no Brasil – possivelmente relacionada a palavrório e promessas vazias. “*Infelizmente*”, acrescentou o então decano do Senado, “*ainda somos dominados pela visão do parlamento do discurso, (...) e passamos essa imagem à sociedade, que vincula o Congresso ao plenário.*”

A citação mostra o desgaste do pronunciamento em plenário, já no início da série histórica deste trabalho. Mas a evolução descrita neste Texto para Discussão mostra que o discurso buscou, e ainda busca, sua transformação. Nesse contexto, a ascensão do registro conversacional mediado indica não um declínio da retórica política, mas uma mudança em suas condições de possibilidade – daí o termo “registro”. Enfim, as alterações representam, acima de tudo, desejo renovado, por parte dos parlamentares, de serem ouvidos e compreendidos. Os efeitos dessas transformações sobre a representação parlamentar permanecem em aberto.

5 REFERÊNCIAS BIBLIOGRÁFICAS

AGÊNCIA SENADO. Senado aprova uso da linguagem simples em órgãos públicos. **Senado Notícias**, 12 mar. 2025.

ALLEN, David; MCNAMARA, Danielle. Text readability and intuitive simplification: A comparison of readability formulas. **Reading in a Foreign Language**, v. 23, p. 84-101, 1 abr., 2011.

BIBER, Douglas; CONRAD, Susan. **Register, genre, and style**. Second edition ed. Cambridge New York: Cambridge University Press, 2019.

BLANCO, Pedro Duarte. **Repositório do projeto “falando_nela”**. Kaggle, 2025a. Disponível em: <<https://www.kaggle.com/datasets/pedroduarteblanco/falando-nela>>. Acesso em: 19 nov. 2025.

BLANCO, Pedro Duarte. **Plenário, Palanque, Estúdio: discursos no Plenário do Senado Federal entre 2007 e 2024**. Texto para Discussão nº 355, dez. 2025b.

BOBBIO, Norberto; MATTEUCCI, Nicola; PASQUINO, Gianfranco. **Dicionário de Política**. 11. ed. Brasília: Editora Universidade de Brasília, 2022.

BURKE, Edmund. Discurso aos eleitores de Bristol. **Revista de Sociologia e Política**, v. 20, p. 97-101, 2012.

CARVILLE, James. Opinion James Carville: Out With Woke. In With Rage. **The New York Times**, 24 nov. 2025.

CASTILHO, Ataliba Teixeira de. **Nova gramática do português brasileiro**. São Paulo, SP: Editora Contexto, 2010.

CHERIAN, Golda; JHA, Dr Sunil Kumar. **A Comparative Study of Readability Tools for EFL Assessments**. v. 12, n. 6, 2024.

CORPORACIÓN LATINOBARÓMETRO. **Confianza en el Congreso/Parlamento — Brasil (1995-2024)**. Disponível em: <<https://www.latinobarometro.org/odajds/>>. Acesso em: 2 fev. 2026.

DALE, Edgar; CHALL, Jeanne S. The Concept of Readability. **Elementary English**, v. 26, n. 1, p. 19-26, 1949.

DEGANI, Marta. Endangered Intellect: A Case Study of Clinton Vs Trump Campaign Discourse. **Iperstoria**, n. 8, 1 dez. 2016.

FAIRCLOUGH, Norman. **Discourse and social change**. Cambridge: Polity, 1996.

FERES JÚNIOR, João; GAGLIARDI, Juliana. Populism and the Media: Introduction to Part II. In: **The Politics of Authenticity and Populist Discourses**. Global Political Sociology. [S.l.]: Palgrave Macmillan, 2021.

GARCIA, Othon M. **Comunicação em prosa moderna**. [S.l.]: Fgv Editora, 2010.

GUIMARÃES, Ulysses. **Discurso do presidente da Assembleia Nacional Constituinte, Deputado Ulysses Guimarães**, 4 out. 2023. Disponível em: <https://www.youtube.com/watch?v=ZBQOWSiUO_g>. Acesso em: 13 jan. 2026.

HANSEN, Lasse; OLSEN, Ludvig Renbo; ENEVOLDSEN, Kenneth. TextDescriptives: A Python package for calculating a large variety of metrics from text. **Journal of Open Source Software**, v. 8, n. 84, p. 5153, 25 abr. 2023.

LEHRMAN, Robert. **The political speechwriter's companion: a guide for writers and speakers**. Washington, D.C: CQ Press, 2010.

LOPES, Lucelene; DURAN, Magali Sanches; PARDO, Thiago Alexandre Salgueiro. Atribuição de lemas e atributos morfológicos seguindo as decisões

adotadas na anotação do corpus Portinari-base dentro das diretrizes da Universal Dependencies (UD). **Relatório técnico**, v. 445, 2023.

MAYHEW, David R.; ARNOLD, R. Douglas. **Congress: the electoral connection**. Second edition ed. New Haven London: Yale University Press, 2004.

MOUFFE, Chantal; COLONNA D'ISTRIA, Pauline. **Pour un populisme de gauche**. Paris: Albin Michel, 2018.

NILC-Matrix Doc. Disponível em: <http://fw.nilc.icmc.usp.br:23380/metrixdoc#indices_leiturabilidade>. Acesso em: 26 jan. 2026.

Pearson Correlation – an overview ScienceDirect Topics. Disponível em: <<https://www.sciencedirect.com/topics/computer-science/pearson-correlation>>. Acesso em: 28 jan. 2026.

PILKINGTON, Ed. How Mamdani built an ‘unstoppable force’ that won over New York. **The Guardian**, 6 nov. 2025.

SENADO FEDERAL. Pronunciamento de Astronauta Marcos Pontes em 09/04/2024. Disponível em: <<https://www25.senado.leg.br/web/atividade/pronunciamentos/-/p/pronunciamento/505357>>. Acesso em: 29 jan. 2026.

SENADO FEDERAL. Pronunciamento de Damares Alves. Disponível em: <<https://www25.senado.leg.br/web/atividade/pronunciamentos/-/p/pronunciamento/500691>>. Acesso em: 30 jan. 2026.

SENADO FEDERAL. Pronunciamento de Fabiano Contarato em 12/05/2021. Disponível em: <<https://www25.senado.leg.br/web/atividade/pronunciamentos/-/p/pronunciamento/479150>>. Acesso em: 30 jan. 2026.

SENADO FEDERAL. Pronunciamento de Romero Jucá. Disponível em: <<https://www25.senado.leg.br/web/atividade/pronunciamentos/-/p/pronunciamento/330463>>. Acesso em: 31 jan. 2026.

SENADO FEDERAL. Pronunciamento de Rose de Freitas. Disponível em: <<https://www25.senado.leg.br/web/atividade/pronunciamentos/-/p/pronunciamento/489416>>. Acesso em: 30 jan. 2026.

Readability. Disponível em: <<https://hlasse.github.io/TextDescriptives/readability.html>>. Acesso em: 16 jan. 2026.

ROSA, Ana Silva. **Do direito à política: liberalismo e judicialismo no pensamento de Rui Barbosa**. Rio de Janeiro: Universidade do Estado do Rio de Janeiro, 2021.

SARNEY, Senador José. **Íntegra do discurso do senador José Sarney – Notícias**. Disponível em: <<https://www.camara.leg.br/noticias/137737-integra-do-discurso-do-senador-jose-sarney/>>. Acesso em: 3 fev. 2026.

SMITH, Allison Jane. Donald Trump speaks like a sixth-grader. All politicians should. **The Washington Post**, 3 maio 2016.

SWORD, Helen. **Stylish academic writing**. Cambridge, Mass: Harvard University Press, 2012.

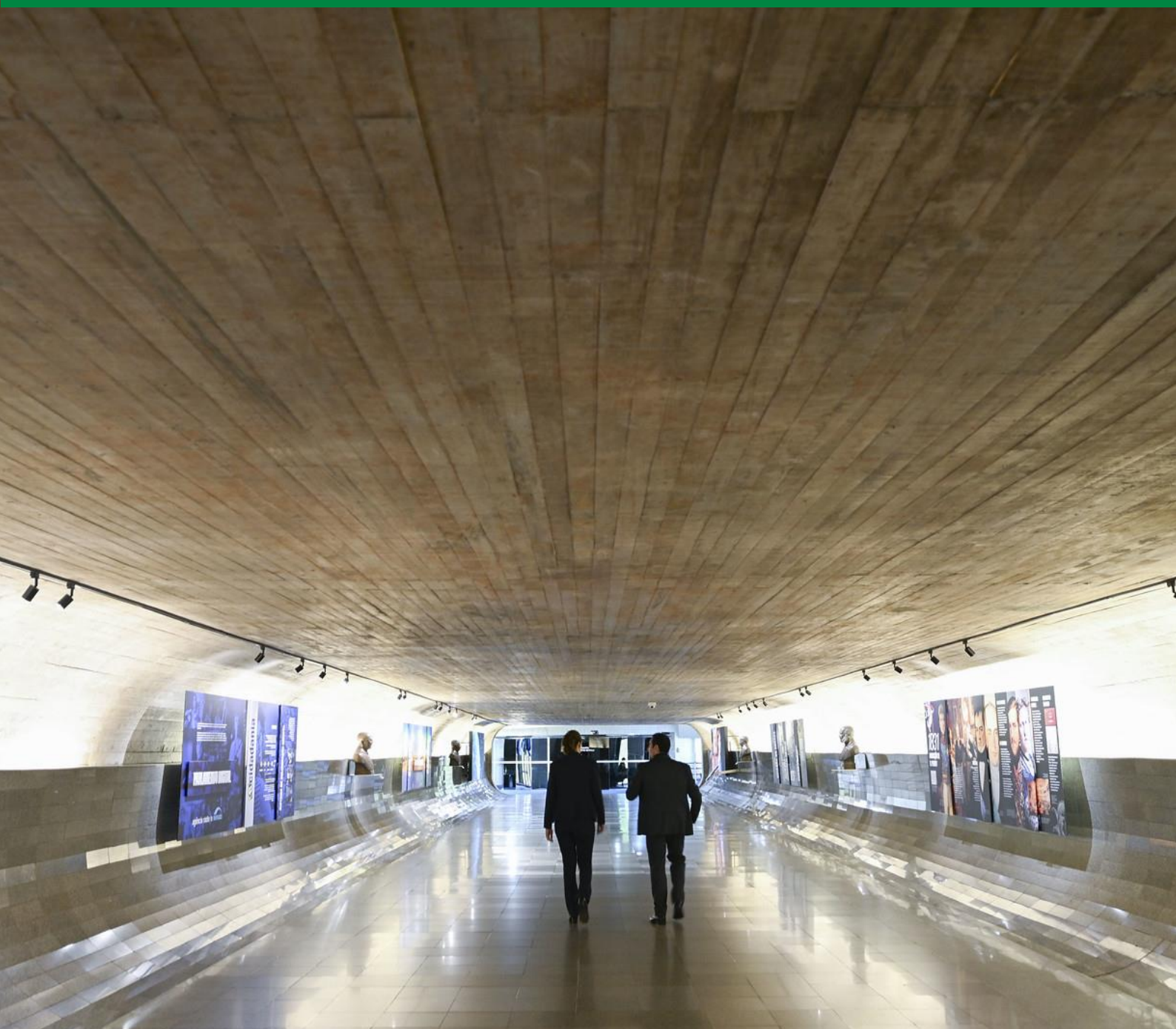
TAGUIEFF, Pierre-André. **Populisme**. Paris: Universalis.fr. [S.d.]. (Nota técnica).

UNIVERSAL DEPENDENCIES. **AUX**: Auxiliary. [S. l.], 2026. Disponível em: https://universaldependencies.org/u/pos/AUX_.html. Acesso em: 30 jan. 2026.

VIDO, Maria. A Oratória Parlamentar na Primeira República. **Revista do Serviço Público**, v. 88, n. 1, p. 103-106, 1960.

Missão da Consultoria Legislativa

Prestar consultoria e assessoramento especializados ao Senado Federal e ao Congresso Nacional, com o objetivo de contribuir com o aprimoramento da atividade legislativa e parlamentar, em benefício da sociedade brasileira.



Núcleo de Estudos e
Pesquisas

Consultoria
Legislativa

SENADO
FEDERAL



ISSN 1983-0645