# Hidden Bias in Empirical Textualism

MATTHEW JENNEJOHN,\* SAMUEL NELSON\*\* & D. CAROLINA NÚÑEZ\*\*\*

A new interpretive technique called "corpus linguistics" has exploded in use over the past five years from state supreme courts and federal courts of appeals to the U.S. Supreme Court. Corpus linguistics involves searching a large database, or corpus, of text to identify patterns in the way in which a certain term is used in context. Proponents of the method argue that it is a more "empirical" approach than referencing dictionaries to determine a word's public meaning, which is a touchstone in originalist approaches to legal interpretation.

This Article identifies an important concern about the use of corpus linguistics in legal interpretation that courts and scholarship have overlooked: bias. Using new machine learning techniques that analyze bias in text, this Article provides empirical evidence that the thousands of documents in the Corpus of Historical American English (COHA), the leading corpus currently used in judicial opinions, reflect gender bias. Courts and scholars have not considered that the COHA is sexist, raising the possibility that corpus linguistics methods could serve as a vehicle for infecting judicial opinions with longstanding prejudices in U.S. society.

In addition to raising this important new problem, this Article charts a course for dealing with it. It explains how hidden biases can be made transparent and introduces steps for "debiasing" corpora used in legal interpretation. More broadly, it shows how the methods introduced here can be used to study biases in all areas of the law, raising the prospect of a revolution in our understanding of how discriminatory biases affect legal decisionmaking.

# TABLE OF CONTENTS

Intro	ODUC	TION	769
I.	Ем	PIRICAL TEXTUALISM AND CORPUS LINGUISTICS	773
	А.	INTRODUCTION TO CORPUS LINGUISTICS	775

<sup>\*</sup> Professor of Law at BYU Law School. @ 2021, Matthew Jennejohn, Samuel Nelson & D. Carolina Núñez.

<sup>\*\*</sup> Research Fellow at BYU Law School.

<sup>\*\*\*</sup> Associate Dean of Faculty and Curriculum and Professor of Law at BYU Law School.

Authors are listed in alphabetical order. Many thanks to Natalie Packard, Jaimee Crossley, and Ryan Wallentine for excellent research assistance. Finally, many thanks to Brad Bernthal, Sadie Blanchard, Jesse Egbert, Jennifer Fan, Alexis Hoag, Cathy Hwang, Tom Lee, Da Lin, Jeremy McClane, Yaron Nili, Darren Rosenblum, Paolo Saguato, Greg Shill, Andrew Winden, and David Wingate for thoughtful comments and discussions regarding earlier drafts. All errors remain the authors'.

	В.	USI	SING CORPUS LINGUISTICS IN LEGAL INTERPRETATION				
		1.	Accuracy	777			
		2.	Quantifiability and Verifiability	780			
	C.	COF	RPUS LINGUISTICS IN THE COURTS	780			
II.	Тн	e Lin	IITS OF CORPUS LINGUISTICS	782			
	А.	A BRIEF CASE STUDY OF CORPUS LINGUISTICS IN STATUTORY INTERPRETATION					
	B.	PRIOR WORK ON THE LIMITS OF CORPUS LINGUISTICS IN LEGAL INTERPRETATION					
		1.	The Illusion of Accuracy and Objectivity	784			
		2.	Concerns About Corpora	785			
		3.	Judicial Competence	786			
III.	Me	Measuring Hidden Bias					
	A.	DAT	DATA				
	B.	HYPOTHESES.					
	C.	ME	ΓHODS	789			
	D.	RES	RESULTS				
		1.	Both Occupational Words and Adjectives in the COHA Reflect Gender Bias	794			
		2.	The Intensity of Gender Bias Changes over Time	795			
		3.	Both Male and Female Authors Write Biased Text, but Female Authors Less So	800			
	E.	SUN	/MARY	802			
IV.	NE	EXT STEPS FOR EMPIRICAL APPROACHES TO LEGAL INTERPRETATION					
	A.	MAKING HIDDEN BIAS TRANSPARENT					
	B.	DEF	BIASING LEGAL CORPUS LINGUISTICS	805			
		1.	Identifying Potential Instances of Bias in Corpus-Based Legal Interpretation	806			
		2.	Corpus Construction and Methodology	806			

	3.	Judging and Mitigating Bias	807
С	. NE	W QUESTIONS THAT WE SHOULD BE ANSWERING	808
Conclu	SION		809

#### INTRODUCTION

A new method for the interpretation of legal texts, such as constitutions, statutes, and regulations, is spreading through the U.S. judiciary.<sup>1</sup> Proponents of the method, known as "corpus linguistics," argue that it is a more reliable, empirical way of discerning the public meaning of a word than resorting to dictionaries, legislative history, or judges' intuitions.<sup>2</sup> Many federal and state courts have applied the method, including Justice Thomas of the U.S. Supreme Court.<sup>3</sup>

In essence, the corpus linguistics method allows a user to search a large body of text, or corpus, for a particular word to identify patterns in usage that reveal information about a word's meaning. For example, a user may track a word's frequency over time, identify the words that most frequently occur in close vicinity to that search term, or review each instance of a word's usage in context.<sup>4</sup>

3. See, e.g., Carpenter v. United States, 138 S. Ct. 2206, 2238 & n.4 (2018) (Thomas, J., dissenting) (using corpus linguistics as evidence that "search" was not associated with "reasonable expectation of privacy"); Wilson v. Safelite Grp., Inc., 930 F.3d 429, 442 (6th Cir. 2019) (Thapar, J., concurring) (describing corpus linguistics as "an important tool ... in figuring out the meaning of a term"); People v. Harris, 885 N.W.2d 832, 839 (Mich. 2016) (using corpus linguistics to show that "information," used alone, may refer to both true and false information); Fire Ins. Exch. v. Oltmanns, 2018 UT 10, ¶ 57 n.9, 416 P.3d 1148, 1163 n.9 (justifying the use of corpus linguistics due to the weaknesses of human intuition); Neese v. Utah Bd. of Pardons & Parole, 2017 UT 89, ¶ 99, 416 P.3d 663, 691 (describing corpus linguistics' tendency to focus on semantics and pragmatics); Craig v. Provo City, 2016 UT 40, ¶ 26 n.3, 389 P.3d 423, 428 n.3 (commending Provo City for briefing on corpus linguistics); State v. J.M. S., 2011 UT 75, ¶¶ 38–40, 280 P.3d 410, 418–19 (Lee, J., concurring) (using corpus linguistics to show that "abortion procedure" refers to a medical procedure); J.M.W., III v. T.I.Z., 2011 UT 38, ¶ 89 & n.21, 266 P.3d 702, 724 & n.21 (Lee, J., concurring in part and concurring in the judgment) (using corpus linguistics to show that "custody" is more closely associated with "divorce" than "adoption"); Muddy Boys, Inc. v. Dep't of Commerce, 2019 UT App 33, ¶¶ 25–26, 440 P.3d 741, 748–49 (stating that corpus linguistics requires large databases rather than small samplings); O'Hearon v. Hansen, 2017 UT App 214, ¶ 25 n.8, 409 P.3d 85, 93 n.8 (citing State v. Rasabout, 2015 UT 72, ¶ 57, 356 P.3d 1258, 1275 (Lee, Associate C.J., concurring in part and concurring in the judgment)); see also Rasabout, 2015 UT 72, ¶¶ 12–13, ¶ 17, 356 P.3d at 1263–64 (using the dictionary to interpret the term "discharge," and noting Associate Chief Justice Lee's suggestion to use corpus linguistics instead); id. ¶ 36, 356 P.3d at 1269 (Durrant, C.J., concurring in part and concurring in the judgment) (explaining that corpus linguistics could be useful in interpreting "the ordinary meaning of statutory terms"); id. ¶¶ 57–65, 356 P.3d at 1275-77 (Lee, Associate C.J., concurring in part and concurring in the judgment) (proposing the use of corpus linguistics as an additional tool for statutory interpretation).

<sup>1.</sup> See infra note 3 and accompanying text.

<sup>2.</sup> See, e.g., Neal Goldfarb, A Lawyer's Introduction to Meaning in the Framework of Corpus Linguistics, 2017 BYU L. REV. 1359, 1367–68; Thomas R. Lee & James C. Phillips, Data-Driven Originalism, 167 U. PA. L. REV. 261, 289–90, 292 (2019); Thomas R. Lee & Stephen C. Mouritsen, Judging Ordinary Meaning, 127 YALE L.J. 788, 829 (2018); Stephen C. Mouritsen, Hard Cases and Hard Data: Assessing Corpus Linguistics as an Empirical Path to Plain Meaning, 13 COLUM. SCI. & TECH. L. REV. 156, 202 (2011).

<sup>4.</sup> See infra Section I.A.

Advocates argue that this information tells us something important about the meaning of a term.<sup>5</sup> Corpus linguistics, they argue, reveals the common meaning of a word more reliably than dictionaries, ad hoc Google searches, or intuition.<sup>6</sup>

Proponents often use H.L.A. Hart's classic problem of how to define "vehicles" in the prohibition "no vehicles in the park" as an example of how corpus linguistics can be used to improve legal interpretation.<sup>7</sup> Analysis of the top fifty words occurring alongside the word *vehicle* in the Corpus of Historical American English (COHA), a massive corpus of texts from 1810 to 2009, reveals that the word *vehicle* rarely occurs with the words *bicycle* or *airplane* but frequently with words related to automobiles, such as *motor*, *highways*, and *streets*.<sup>8</sup> Examining the instances of *vehicle* in context corroborates the tendency of the word to refer to automobiles rather than other things that might fall in a dictionary's definition of *vehicle*.<sup>9</sup> Thus, based on that public meaning of the word *vehicle* as revealed by corpus linguistics, the rule "no vehicles in the park" refers to automobiles, not bicycles or airplanes.

Like any new discipline or methodology, corpus linguistics—as applied to legal interpretation—has issues that it must address to realize its potential. Two methodological criticisms stand out. The first challenges the accuracy and objectivity of existing corpus linguistics methodologies as a tool of statutory construction. Critics have argued, for example, that the frequency by which a word appears with other words in a corpus provides little information about what the word of interest means.<sup>10</sup> In situations where usage is varied—the very places where corpus linguistics is supposed to be useful—reliance on word frequencies is an undisciplined process likely to produce indeterminate results.<sup>11</sup> In addition, scholars have recognized that corpus linguistics analysis, despite its ability to quantify incidents of usage, nonetheless relies on the kind of qualitative intuition

7. Lee & Mouritsen, supra note 2, at 800, 836-45.

8. Id. at 837–40.

<sup>5.</sup> See, e.g., Stefan Th. Gries & Brian G. Slocum, Ordinary Meaning and Corpus Linguistics, 2017 BYU L. REV. 1417, 1441; Lee & Mouritsen, supra note 2; Brian G. Slocum, Ordinary Meaning and Empiricism, 40 STATUTE L. REV. 13, 18 (2018).

<sup>6.</sup> See, e.g., Goldfarb, supra note 2; Lee & Mouritsen, supra note 2, at 868 n.296; Lee & Phillips, supra note 2, at 283 (explaining that words draw meaning from other words surrounding them, whereas dictionaries provide the meanings of words in isolation); Friedemann Vogel, Hanjo Hamann & Isabelle Gauer, Computer-Assisted Legal Linguistics: Corpus Analysis as a New Tool for Legal Studies, 43 LAW & SOC. INQUIRY 1340, 1346 (2018); Stephen C. Mouritsen, Comment, The Dictionary Is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning, 2010 BYU L. REV. 1915, 1924 ("A dictionary cannot tell us precisely what meaning a word must bear in a particular context, because the lexicographer cannot know a priori every context in which the term will be found.").

<sup>9.</sup> *See, e.g., Vehicle,* MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/vehicle [https://perma.cc/YE2H-3UFR] (last visited Jan. 6, 2021) (including not only "motor vehicle" but also "investment vehicle," an "inert medium (such as a syrup)," and "a work created especially to display the talents of a particular performer" as examples, among others, of the definition of "vehicle").

<sup>10.</sup> See infra Part II.

<sup>11.</sup> See Stefan Th. Gries, Corpora and Legal Interpretation: Corpus Approaches to Ordinary Meaning in Legal Interpretation, in THE ROUTLEDGE HANDBOOK OF FORENSIC LINGUISTICS 628, 638–39 (Malcolm Coulthard et al. eds., 2d ed. 2020).

and individual judgment that corpus linguistics proponents have criticized in other legal interpretation methods.<sup>12</sup> The second type of criticism questions the constitution of corpora themselves. It is not clear, for instance, that the corpora being used in legal interpretation are from linguistic communities relevant to the usages being litigated.<sup>13</sup> As a result, even if judges can draw reliable inferences from a corpus, an unrepresentative corpus may still produce misleading results.

This Article identifies a new concern about using corpus linguistics for legal interpretation. It provides evidence that the COHA, one of the primary corpora that is used in legal interpretation, reflects structural gender bias. Using recently developed machine learning methods, this Article provides quantitative evidence that men and women are routinely referred to differently in the thousands of texts in the COHA, and those differences are usually negative with respect to women. The COHA is a sexist corpus.<sup>14</sup>

Unfortunately, the possibility of gender bias in corpora such as the COHA has been entirely overlooked in the legal scholarship discussing corpus linguistics. Prior work on the question of how to build representative corpora has focused upon what types of media to use without exploring issues of biases in text relating to, among other things, gender, ethnicity, class, sexual orientation, age, and ability.<sup>15</sup> The problem of bias is a lacuna in the law and scholarship applying empirical approaches to legal interpretation.

This Article fills that gap in the literature. It argues that understanding the effects of such biases when using corpus linguistics methods in legal interpretation should be a top priority of scholars and judges. Importantly, understanding the effects of bias in a corpus is imperative even if that bias is merely a reflection of existing bias in culture and language. To continue ignoring this issue is to risk infecting judicial decisionmaking with the structural biases endemic to American society—and embedded in its linguistic patterns—under the guise of an ostensibly neutral, objective interpretive approach.

The Article also contributes more broadly to doctrine and legal scholarship by introducing new methods for quantitatively measuring biases in text. The

13. Gries, supra note 11, at 628, 637-38.

<sup>12.</sup> See, e.g., Anya Bernstein, *Democratizing Interpretation*, 60 WM. & MARY L. REV. 435, 453–54 (2018); Carissa Byrne Hessick, *Corpus Linguistics and the Criminal Law*, 2017 BYU L. REV. 1503, 1505; Jake Linford, *Datamining the Meaning(s) of Progress*, 2017 BYU L. REV. 1531, 1555; *see also* Mouritsen, *supra* note 2 (arguing that corpus methodology cannot escape confirmation bias because biases may shape how judges interpret data); Lawrence M. Solan, *Can Corpus Linguistics Help Make Originalism Scientific?*, 126 YALE L.J.F. 57, 64 (2016) (acknowledging that "like the lexicographer," an originalist employing corpus linguistics analysis "will have other choices to make about how narrowly or broadly, thinly or thickly, to construe a relevant word"); *id.* ("These choices are not strictly linguistic. They depend upon the commitments of the corpus's user, and these commitments depend upon the user's stance with respect to the language being analyzed.").

<sup>14.</sup> The claim here is not that the COHA was assembled in a discriminatory fashion but rather that, as a representative collection of texts, it reflects the gender bias that has existed and persists today in the broader culture. *See infra* Section III.D.

<sup>15.</sup> See James C. Phillips & Jesse Egbert, Advancing Law and Corpus Linguistics: Importing Principles and Practices from Survey and Content-Analysis Methodologies to Improve Corpus Design and Analysis, 2017 BYU L. REV. 1589, 1608–13.

methods are revolutionary in a text-based discipline such as the law. Biases in a variety of relevant text types, from judicial decisions and legislative history to corporate disclosures and documentary evidence at trial, can be identified using the techniques introduced here.

This Article unfolds as follows. Part I introduces corpus linguistics methods and the case law that applies them in legal interpretation. Part II then discusses criticisms of corpus linguistics and notes that, to date, the law and corpus linguistics literature have not considered the possibility or implications of hidden biases within the corpora of texts that are increasingly used by courts.

Part III introduces methods for quantitatively measuring gender bias in a corpus of text. The methodological discussion is undertaken in plain English to introduce a general audience to the machine learning techniques that are used to measure bias. Those techniques are then used to analyze gender bias in the COHA. The results of that analysis provide evidence that the documents in the COHA reflect gender bias. A search for public meaning to aid the interpretation of a legal text from any time period thus runs the risk of incorporating that gender bias.

Finally, Part IV of this Article discusses the normative implications of its study and the new methods that it introduces. This Part first calls for further development of the methods to measure different types of bias, beyond the gender bias analyzed here, to map the scale of the bias problem in the corpora commonly used in legal interpretation. This Part then considers possible avenues for "debiasing" the corpora used in legal interpretation. Finally, this Part discusses the costs and benefits of using more sophisticated quantitative methods for legal interpretation. These more rigorous tools for identifying patterns of word usage respond to a concern that the method that judges and practitioners currently employ leads to indeterminate results.<sup>16</sup> However, even if they provide a better interpretive mouse trap, introducing greater methodological sophistication in the interpretive process presents a real trade-off: It will further remove the analysis of legal text from judges' hands.

Therefore, this Article urges courts and scholars to incorporate concerns about bias into the further development and application of corpus linguistics methods to legal interpretation. The key next step is to clarify what value, if any, empirical textualism can provide when its methods are improved so that we have a clearer sense of the trade-off between judicial accuracy and access to justice for average litigants. Until that clarity is achieved, we question whether further experimentation with corpus linguistics by entrepreneurial litigants and judges is prudent. Rather, we advocate for a period of study and methodological refinement before further application of empirical textualism in the resolution of disputes.

<sup>16.</sup> Gries, *supra* note 11, at 628, 629–31; Ethan J. Herenstein, *The Faulty Frequency Hypothesis: Difficulties in Operationalizing Ordinary Meaning Through Corpus Linguistics*, 70 STAN. L. REV. ONLINE 112, 113 (2017); Mouritsen, *supra* note 6, at 1920–21.

#### I. EMPIRICAL TEXTUALISM AND CORPUS LINGUISTICS

A judge's interpretation of a phrase or word—and the method by which the judge arrives at that interpretation—can mean the difference between, for example, freedom and imprisonment,<sup>17</sup> deportation and continued residence,<sup>18</sup> vindication of a right and dismissal of a claim,<sup>19</sup> or recovery and loss for breach of contract.<sup>20</sup> What a judge believes the word *vehicle* to mean will dictate whether the man who rides his bicycle into a park past a "no vehicles in the park" sign will be fined.<sup>21</sup>

It is no wonder, then, that the scholarly debate over how exactly legal texts should be interpreted remains vibrant and spirited today despite the decades of accumulated writings that already exist on the subject. The classic debate between purposivists and textualists highlights a potential tension between an author's purpose in writing the text and the text's literal meaning.<sup>22</sup> If a lawmaking body's sole purpose in deciding that there should be "no vehicles in the park" is to keep people from driving their cars through the park, then the purposivist judge may decide that walking a bicycle through the park is not a violation of the rule embodied in the sign. But a textualist judge may instead decide that the text of the sign and the underlying rule require a different outcome. In a strict textualist approach, the man's bicycle might be deemed a vehicle that should not have been "in the park."

Despite the tension between purposivism and textualism, few would argue that the text of a statute, regulation, constitution, or other legal text does not matter.<sup>23</sup>

<sup>17.</sup> *See, e.g.*, Muscarello v. United States, 524 U.S. 125, 126–27 (1998) (deliberating whether "carries a firearm" includes having a firearm in a locked glove compartment of car); United States v. Costello, 666 F.3d 1040, 1043 (7th Cir. 2012) (discussing whether the verb "to harbor" requires proof of concealment).

<sup>18.</sup> *See*, *e.g.*, Costello v. Immigration & Naturalization Serv., 376 U.S. 120, 122 (1964) (discussing whether "*is* convicted" means that one must be an alien at the time of conviction to be deported under the Immigration and Nationality Act); Barber v. Gonzales, 347 U.S. 637, 638–39 (1954) (discussing whether a U.S. national born in a former U.S. territory made an "entry" into the United States when traveling from the territory to the continental United States).

<sup>19.</sup> See, e.g., Carpenter v. United States, 138 S. Ct. 2206, 2213–14 (2018) (demonstrating how Fourth Amendment rights hinge on the interpretation of "searches").

<sup>20.</sup> *See, e.g.*, Hegel v. First Liberty Ins. Corp., 778 F.3d 1214, 1216 (11th Cir. 2015) (showing that recovery of damages from an insurance company depends on the interpretation of "structural damage"); Cyze v. Banta Corp., No. 07 C 2357, 2009 WL 2905595, at \*3 (N.D. Ill. Sept. 8, 2009) (explaining that severance benefits for a terminated employee depend on the interpretation of "in anticipation of").

<sup>21.</sup> See H. L. A. Hart, Positivism and the Separation of Law and Morals, 71 HARV. L. REV. 593, 606–07 (1958).

<sup>22.</sup> See Michael Herz, Purposivism and Institutional Competence in Statutory Interpretation, 2009 MICH. ST. L. REV. 89, 93–94; see also William Baude & Stephen E. Sachs, The Law of Interpretation, 130 HARV. L. REV. 1079, 1115–16 (2017) (comparing the two approaches and discussing the benefits and weaknesses of both); Note, *Textualism as Fair Notice*, 123 HARV. L. REV. 542, 551–53 (2009) (discussing the public choice argument for textualism). But see Jill C. Anderson, Misreading like a Lawyer: Cognitive Bias in Statutory Interpretation, 127 HARV. L. REV. 1521, 1536–37 (2014) (arguing that, in many cases, the distinction between a textualist approach and a purposivist approach obscures the reality of multiple and unambiguous readings of a phrase).

<sup>23.</sup> See Lee & Mouritsen, supra note 2, at 793; Mouritsen, supra note 2, at 161.

Nowhere is this more true than in the courts, where judges often start any exercise in legal interpretation by looking for the "ordinary meaning" of a text, often with reference to a dictionary, past experience with a term, or even gut instinct.<sup>24</sup> A judge may ultimately treat the ordinary meaning of the text as merely a piece of evidence of a lawmaking body's intent or as the only analysis that matters, or anywhere in between, depending on how "plain,"<sup>25</sup> "clear,"<sup>26</sup> or "obvious"<sup>27</sup> the judge deems the text to be.<sup>28</sup> Even purposivist scholarship recognizes the value of text,<sup>29</sup> and some might say that the classic camps distinguish themselves by degree rather than by kind. Purposivists are more willing to explore the context in which a law was created to understand the meaning of the text,<sup>30</sup> while textualists are more likely to begin and end with the text in spite of potentially contradicting evidence of legislative intent.<sup>31</sup>

The argument, then, is not simply about whether text matters. That much seems quite clear. The argument is much more nuanced than that. It is about *when* the text matters, *how* to figure out *what* the text means, *why* it matters, and *how much* it matters. But, more importantly, the argument is also fundamentally about whether any of these questions can actually be answered in any meaningful way and whether they are worth trying to answer at all.<sup>32</sup> Textualists argue that the ordinary meaning of a text, at least in theory, is more objectively discoverable than is the intent of its authors or its purpose.<sup>33</sup> This objectivity, in turn, furthers consistency in application, allows more accurate predictions of a law's

29. HENRY M. HART, JR. & ALBERT M. SACKS, THE LEGAL PROCESS: BASIC PROBLEMS IN THE MAKING AND APPLICATION OF THE LAW 1375 (William N. Eskridge, Jr. & Philip P. Frickey eds., 1994).

30. STEPHEN BREYER, ACTIVE LIBERTY: INTERPRETING OUR DEMOCRATIC CONSTITUTION 85 (2005); HART JR., & SACKS, *supra* note 29.

<sup>24.</sup> See, e.g., Muscarello v. United States, 524 U.S. 125, 128, 130 (1998) (relying on the Oxford English Dictionary, Webster's Third New International Dictionary, Random House Dictionary of the English Language Unabridged, and Black's Law Dictionary to interpret the meaning of carry in "carries a firearm"); United States v. Costello, 666 F.3d 1040, 1043 (7th Cir. 2012) (noting that ""[t]o shelter' has an aura of protectiveness, as in taking 'shelter' from a storm" and that ""[s]heltering' doesn't seem the right word for letting your boyfriend live with you" in interpreting "harboring"); Cabell v. Markham, 148 F.2d 737, 739 (2d Cir. 1945) (describing the enterprise of statutory interpretation as requiring "sympathetic and imaginative discovery" of the statute's purpose).

<sup>25.</sup> Abbe R. Gluck, The States as Laboratories of Statutory Interpretation: Methodological Consensus and the New Modified Textualism, 119 YALE L.J. 1750, 1784 (2010).

<sup>26.</sup> JOHN F. MANNING & MATTHEW C. STEPHENSON, LEGISLATION AND REGULATION: CASES AND MATERIALS 60 (2d ed. 2013) ("When the Court finds the text to be clear in context, it now routinely enforces the statute as written.").

<sup>27.</sup> William Baude & Ryan D. Doerfler, *The (Not So) Plain Meaning Rule*, 84 U. CHI. L. REV. 539, 545 (2017).

<sup>28.</sup> See MANNING & STEPHENSON, supra note 26; Baude & Doerfler, supra note 27; Gluck, supra note 25, at 1756–58.

<sup>31.</sup> John F. Manning, *What Divides Textualists from Purposivists*?, 106 COLUM. L. REV. 70, 92–93 (2006).

<sup>32.</sup> See Baude & Sachs, supra note 22, at 1088 ("[T]here just may not be a single right way to read a legal text."); Richard A. Posner, Statutory Interpretation—In the Classroom and in the Courtroom, 50 U. CHI. L. REV. 800, 808 (1983) (observing about the "start with the words' canon" that "[i]t is ironic that a principle designed to clarify should be so ambiguous").

<sup>33.</sup> Mouritsen, supra note 2, at 172-73.

application, and fosters surer reliance on the law by the governed.<sup>34</sup> Though the nobility of these goals is virtually uncontested, some scholars contend that searching for ordinary meaning does little to achieve these goals.<sup>35</sup> They argue that there is no clear way to define ordinary meaning and that, in any event, ordinary meaning cannot be objectively determined.<sup>36</sup> Textualists have admitted to these challenges but offer ordinary meaning as the method that best achieves these goals, even if it does not achieve them perfectly every time.<sup>37</sup>

It is in the unsatisfying gap between "it's the best we've got" and "it's not good enough" that some scholars are turning to corpus linguistics as a way to address the uncertainty in ordinary meaning and to rehabilitate textualism.<sup>38</sup> Corpus linguistics has been hailed as "an evolution"<sup>39</sup> and "an important tool"<sup>40</sup> in legal interpretation. That sentiment has caught on among scholars of legal interpretation and increasingly among judges.<sup>41</sup> With so much riding on a judge's interpretation of even a single word, corpus linguistics deserves close scrutiny.

In this Part, we describe the current discussion and use of corpus linguistics as a method of legal interpretation. Section I.A provides a basic description of corpus linguistics methodology and tools. Section I.B catalogues the rise of corpus linguistics in U.S. courts and illustrates how corpus linguistics is used by courts. Section I.C describes the current academic literature discussing the potential use of corpus linguistics in legal interpretation. This Part also prepares the reader for a discussion of corpus linguistics' limits in Part II.

### A. INTRODUCTION TO CORPUS LINGUISTICS

As computing power has rapidly increased, linguists have created larger and larger digitized databases of real-world natural language, called

<sup>34.</sup> Id. at 172.

<sup>35.</sup> WILLIAM N. ESKRIDGE, JR., PHILIP P. FRICKEY & ELIZABETH GARRETT, CASES AND MATERIALS ON LEGISLATION: STATUTES AND THE CREATION OF PUBLIC POLICY 792–93 (4th ed. 2007) (""[P]lain meaning' is . . . a deeply ambiguous term."); Richard H. Fallon Jr, *The Meaning of Legal "Meaning" and Its Implications for Theories of Legal Interpretation*, 82 U. CHI. L. REV. 1235, 1268 (2015) (asserting that "[u]ncertainty and division seem inevitable" in statutory interpretation).

<sup>36.</sup> Fallon, *supra* note 35, at 1272 (arguing that there is "no single, linguistic fact of the matter concerning what statutory or constitutional provisions mean").

<sup>37.</sup> *See* Lee & Mouritsen, *supra* note 2, at 795 ("While the search for ordinary meaning is hard, the premises of this inquiry are too deeply embedded in our law and too clearly rooted in important policy considerations to give up at the first sight of difficulty or indeterminacy ....").

<sup>38.</sup> See id. at 828.

<sup>39.</sup> Students Present New Insights on Original Meaning of Constitution to Judges Using "Big Data" of Corpus Linguistics, GA. ST. U.: GA. ST. NEWS HUB (May 21, 2018), https://news.gsu.edu/corpus-linguistics [https://perma.cc/B78M-LRZA].

<sup>40.</sup> BYU Law Launches First Legal Corpus Linguistics Technology Platform, BYU L., https://law. byu.edu/news/byu-law-launches-first-legal-corpus-linguistics-technology-platform [https://perma.cc/ G2PM-8HNP] (last visited Jan. 7, 2021) (quoting Lawrence Solum, a professor of law and an expert in constitutional theory).

<sup>41.</sup> *See, e.g.*, Carpenter v. United States, 138 S. Ct. 2206, 2238 & n.4 (2018) (Thomas, J., dissenting) (using corpus linguistics as evidence that "search" was not associated with "reasonable expectation of privacy"); People v. Harris, 885 N.W.2d 832, 839 (Mich. 2016) (using corpus linguistics to show that "information," used alone, may refer to both true and false information).

"corpora."<sup>42</sup> A single corpus might contain millions of words collected from books, magazines, newspapers, television transcripts, scholarly articles, blog posts, or other sources.<sup>43</sup> For linguists, natural text is an ideal subject of study precisely because its authors are unaware that their words are the subject of study.<sup>44</sup> The text, if representative of the speech community at issue, depicts how a speech community actually speaks or writes.<sup>45</sup> With a large enough sample of natural language, linguists claim to be able to identify and describe subtle linguistic patterns of which the speech community may be wholly unaware.<sup>46</sup>

Linguists analyze a corpus through corpus linguistics tools and functions that return data about specific words in that corpus. Most commonly, linguists use tools that allow them to (1) identify the frequency of a word's occurrence in the corpus;<sup>47</sup> (2) understand a word's collocation, or its tendency to appear close to other words in the corpus;<sup>48</sup> and (3) review autogenerated lists of every instance of a word's usage—including its context—throughout the corpus.<sup>49</sup> A corpus user may refine each of these functions by restricting the genre, time period, source, part of speech, or any other feature or field recognized by the corpus.

As a basic illustration of each of these types of tools, we examine the term *vehicle* in the Corpus of Contemporary American English (COCA), a corpus that, at the time of this writing, contains over one billion words evenly divided among "spoken, fiction, popular magazines, newspapers, academic texts, and ... TV and Movies subtitles, blogs, and other web pages."<sup>50</sup> The "frequency" function for

48. Id. at 20-21.

<sup>42.</sup> See Douglas Biber, Corpus-Based and Corpus-Driven Analyses of Language Variation and Use, in THE OXFORD HANDBOOK OF LINGUISTIC ANALYSIS 193, 193 (Bernd Heine & Heiko Narrog eds., 2d ed. 2015).

<sup>43.</sup> See, e.g., Mark Davies, *The* 385+ *Million Word* Corpus of Contemporary American English (1990–2008+): Design, Architecture, and Linguistic Insights, 14 INT'L J. CORPUS LINGUISTICS 159, 161–62 (2009).

<sup>44.</sup> *Cf.* TONY MCENERY, RICHARD XIAO & YUKIO TONO, CORPUS-BASED LANGUAGE STUDIES: AN ADVANCED RESOURCE BOOK 6 (2006) (explaining that natural language texts demonstrate "what speakers believe to be acceptable utterances" but that invented language examples "may not represent typical language use").

<sup>45.</sup> See SUSAN HUNSTON, CORPORA IN APPLIED LINGUISTICS 20–22 (Michael H. Long & Jack C. Richards eds., 2002).

<sup>46.</sup> Biber, *supra* note 42, at 193, 197. For examples of how linguists use corpora, see Max Boholm, *Risk and Quantification: A Linguistic Study*, 39 RISK ANALYSIS 1243, 1247, 1250–53 (2019) (using corpora to explore quantification in various linguistic practices rather than to fully explain identified differences, and isolating how "risk" was used in different kinds of sources within corpora); Fanny Meunier, *Corpus Linguistics and Second/Foreign Language Learning: Exploring Multiple Paths*, 11 REVISTA BRASILEIRA LINGUÍSTICA APLICADA 459, 465–72 (2011) (suggesting the use of corpora to aid in second language acquisition); and Juan Lorente Sánchez, '*Give It Him and Then I'll Give You Money for It'*. *The Dative Alternation in Contemporary British English*, 6 RES. CORPUS LINGUISTICS 15, 21–22 (2018) (breaking down the usage of dative grammatical structures between speech and writing and between written books and periodicals, written-to-be-spoken, and written miscellaneous categories).

<sup>47.</sup> HUNSTON, *supra* note 45, at 21.

<sup>49.</sup> *Id.* Many scholars refer to this tool as KWIC, or Keyword in Context. *See, e.g.*, Davies, *supra* note 43, at 167–68.

<sup>50.</sup> *Search*, CORPUS CONTEMP. AM. ENG., https://www.english-corpora.org/coca [https://perma.cc/DZ6F-48PV] (last visited Jan. 7, 2021).

*vehicle* returns 42,305 hits.<sup>51</sup> The "context" function results in a list of every instance in which *vehicle* appears in the corpus, along with the surrounding language, including when it is used to refer to a means of transportation, such as "the vehicle ran onto the sidewalk and injured a pedestrian," and when it is used as a figurative means of transportation, as in "[t]his tradition is older than § 1983, the statutory vehicle for bringing claims against state and local actors," along with the date, genre, and source title for each instance.<sup>52</sup> The COCA also allows users to identify which words tend to appear with *vehicle* within a user-defined distance from the search term. Limiting the distance to four words before and after *vehicle* returns a list of collocates—words that tend to appear with *vehicle* in the corpus—including *motor*, *stolen*, and *electric*, listed by part of speech.<sup>53</sup>

# B. USING CORPUS LINGUISTICS IN LEGAL INTERPRETATION

The primary arguments in favor of using corpus linguistics as a tool for textual legal interpretation fall into two related categories: (1) accuracy and (2) quantifiability and replicability. These arguments are largely framed in terms of comparison: corpus linguistics vastly improves judges' existing legal interpretation practices and methods.<sup>54</sup> The arguments in favor of corpus linguistics, then, are as much a criticism of current methods of legal interpretation as they are a case for corpus linguistics. Given proponents' pointed criticisms of courts' often-used methods of legal interpretation, it is no wonder that many judges have rushed to adopt corpus linguistics. Corpus linguistics proponents paint a portrait of a haphazard and unprincipled legal interpretation practice in need of reform.

# 1. Accuracy

Many scholars have catalogued the increased use of dictionaries in legal interpretation over the last fifty years.<sup>55</sup> And although dictionaries may be apt at

<sup>51.</sup> Frequency Results for "Vehicle," CORPUS CONTEMP. AM. ENG., https://www.english-corpora. org/coca/?c=coca&q=91230344 [https://perma.cc/762M-U6KP] (last visited Jan. 7, 2021) (figures accurate as of the time of this writing).

<sup>52.</sup> See id.

<sup>53.</sup> Frequency Results for Collocates Within Four Terms of "Vehicle," CORPUS CONTEMP. AM. ENG., https://www.english-corpora.org/coca [https://perma.cc/MAL7-9HW3] (click on the plus sign next to "Browse" on the left side of the page; select "Collocates"; enter "vehicle" for "Word/phrase"; do not enter anything in the "Collocates" box; select the number four in both the positive and negative directions; and select "Find collocates") (last visited Jan. 7, 2021) (results accurate as of the time of this writing).

<sup>54.</sup> See, e.g., Lee & Mouritsen, supra note 2, at 866.

<sup>55.</sup> Mouritsen, *supra* note 2, at 160, 164 (explaining that textualist judges rely on intuition and extrinsic aids like dictionaries); Samuel A. Thumma & Jeffrey L. Kirchmeier, *The Lexicon Has Become a Fortress: The United States Supreme Court's Use of Dictionaries*, 47 BUFF. L. REV. 227, 231 (1999) (explaining that over time the Supreme Court has increased its reliance on dictionaries, and describing that, during the entire 1950s, the Court relied on a dictionary definition to define eighteen terms in eleven opinions but that in the 1997–1998 term alone, the Court relied on a dictionary definition to define twenty-seven terms in seventeen opinions); *see also* Stephen C. Mouritsen, *Contract Interpretation with Corpus Linguistics*, 94 WASH. L. REV. 1337, 1348–49 (2019) (highlighting how the failure to use dictionaries can result in reversal on appeal); James C. Phillips, Daniel M. Ortner & Thomas R. Lee, *Corpus Linguistics & Original Public Meaning: A New Tool to Make Originalism More* 

describing the historically known, possible definitions and usages of a word,<sup>56</sup> they cannot tell the reader which of those definitions or usages is relevant in any context.<sup>57</sup> Neither can dictionaries identify which meaning of a word is its ordinary meaning, nor even assume a shared understanding of the term's ordinary meaning.<sup>58</sup> Corpus linguistics proponents have noted that judges nonetheless use dictionaries as authoritative evidence of what legislative text must mean.<sup>59</sup> In doing so, judges rely on a flawed understanding of how a dictionary works or, more concerningly, on intuition to cherry-pick among definitions.<sup>60</sup>

Even where judges recognize the incongruence between a dictionary's purpose and the task of determining ordinary meaning, they have been unable to identify a method that does not suffer from similar problems. Stephen Mouritsen, Thomas

56. Mouritsen, *supra* note 6, at 1921–23 (describing dictionaries' role in defining unknown terms and instantiating contested meanings).

57. See Goldfarb, supra note 2, at 1367; Lee & Phillips, supra note 2, at 283 (stating that words draw meaning from other words surrounding them, but dictionaries provide the meanings of words in isolation); *id*. ("[T]he communicative content of a phrase isn't always the sum of its parts."); Vogel et al., supra note 6; Mouritsen, supra note 6 ("A dictionary cannot tell us precisely what meaning a word must bear in a particular context, because the lexicographer cannot know a priori every context in which the term will be found.").

58. See John D. Ramer, Note, Corpus Linguistics: Misfire or More Ammo for the Ordinary-Meaning Canon?, 116 MICH. L. REV. 303, 307–308 (2017).

59. See Mouritsen, supra note 6, at 1924–25 ("[J]udges have increasingly sought to employ dictionaries for persuasive ends. They have done so by arguing that because multiple dictionaries define a term in a given way, a particular definition ought somehow to be controlling in a given case. Judges have also maintained that because a definition has been placed in a particular position in what the judge perceives as the dictionary's structural hierarchy, or because the derivation of a term reveals that its original use was similar to the meaning the judge favors, the judge's particular meaning should be preferred. These conclusions are erroneous, not simply because they are at variance with the descriptive purpose for which most contemporary dictionaries are created, but because they rely upon deeply flawed assumptions about the structure and content of the information presented in dictionaries."); see also Jacob Crump, Comment, *Corpus Linguistics in the* Chevron *Two-Step*, 2018 BYU L. REV. 399, 401 ("[T]he temptation is for judges to reflexively turn to dictionaries to marshal support for their own intuitions about linguistic ambiguity and the reasonableness of various interpretations. But the problem is, this type of reasoning allows judges to look out over the crowd of dictionary definitions and pick out their friends.").

60. See Ramer, supra note 58 (explaining that, when two judges find support in different dictionaries, "the dispute is . . . based on the judges' differing intuitions about the word's ordinary meaning"); see also Phillips et al., supra note 55, at 29 (referring to the use of Google as a "stumble," though a "stumble in the right direction"); Mouritsen, supra note 6, at 1969 (giving an example of a disagreement between two judges over which dictionary definition to use being resolved by differing intuitions).

*Empirical*, 126 YALE L.J.F. 21, 23 (2016) ("[O]riginalism often relies heavily on an imperfect tool contemporaneous dictionaries . . . ."); Mouritsen, *supra* note 6, at 1915, 1924–25 (describing how judges refer to dictionaries when faced with hard cases). For instances of judicial use of dictionaries, see, for example, Fire Ins. Exch. v. Oltmanns, 2018 UT 10, ¶ 10 & n.1, 416 P.3d 1148, 1151 & n.1 (using multiple dictionaries to define "jet ski"); and see also People v. Harris, 885 N.W.2d 832, 838 (Mich. 2016) (using dictionaries to interpret "information"); State v. Thonesavanh, 904 N.W.2d 432, 436 (Minn. 2017) (turning first to the dictionary when the definition of "takes" was not included in the statute); State v. Rasabout, 2015 UT 72, ¶ 12, 356 P.3d 1258, 1263 (using dictionary entries to determine the "clearest reading" of "discharge"); State v. J.M.S., 2011 UT 75, ¶ 14, 280 P.3d 410, 413 (using dictionaries to show that "procedure" has "multiple interpretations"); and O'Hearon v. Hansen, 2017 UT App 214, ¶ 25, 409 P.3d 85, 93 ("A 'starting point' for a court's 'assessment of ordinary meaning is the dictionary.").

Lee, and James Phillips have criticized Judge Posner's rudimentary use of Google to determine the meaning of *harboring* as used in a statute that prohibited "conceal[ing], harbor[ing] or shield[ing] from detection" an "alien in any place, including any building or any means of transportation."<sup>61</sup> Judge Posner's search of *harbor* in Google did not—and could not—limit results to the correct part of speech, could not account for the "black box of the Google algorithm," ignored "what biases [were] being introduced," and depended on Judge Posner's intuition in deciding what strings of words to search.<sup>62</sup>

Proponents offer corpus linguistics as the antidote to legal interpretation based on "scattershot, impressionistic evidence."<sup>63</sup> "A big part of the motivation behind introducing corpus linguistics into legal interpretation is to increase the sophistication and quality of interpretive analysis."<sup>64</sup> Corpus linguistics, proponents argue, takes a word's context into account in a way that dictionaries cannot.<sup>65</sup> Corpus linguistics claims to allow "meanings of words to be investigated in light of other words in which they co-occur" through the use of collocation searches.<sup>66</sup> In addition, corpus linguistics allows judges to take into account other relevant context by either using a specialized corpus that is specific to a particular industry, time period, or geographic region, or limiting search results to particular genres or time periods, for example.<sup>67</sup> In this way, corpus linguistics claims to be more precisely calibrated to the task of interpreting legal text.

65. *See, e.g.*, Goldfarb, *supra* note 2, at 1363, 1377 (using corpus linguistics to reveal details about the word *carry* not recorded in dictionaries); Vogel et al., *supra* note 6.

66. Gries & Slocum, *supra* note 5; *see also* Slocum, *supra* note 5 ("Contextual considerations are such an integral aspect of meaning that even determinants of meaning that are based on generalized intent and systematicities of language usage may require consideration of the particularized context of the statute.").

67. Some scholars have suggested that corpora for particular industries could help reveal the meaning of industry terms of art. *See, e.g.,* James A. Heilpern, *Dialects of Art: A Corpus-Based Approach to Technical Term of Art Determinations in Statutes,* 58 JURIMETRICS 377, 379 (2018) ("The emerging discipline of law and corpus linguistics now provides practitioners, expert witnesses, and judges with new tools to directly analyze the ordinary meaning of a word *within* an industry . . . without first having to determine whether that industry uses a phrase in a distinct manner.").

Many originalist constitutional scholars have used the Corpus of Founding Era American English (COFEA) to analyze the meaning of constitutional terms. *See* John W. Welch & James A. Heilpern, *Recovering Our Forgotten Preamble*, 91 S. CAL. L. REV. 1021, 1065–67 (2018) (stating that "corpus linguistics can help mitigate the problems associated with linguistic drift"—"the notion that language usage and meaning shifts over time"—by creating quantifiable data from a body of text written in the Founding Era). The COFEA is a corpus of documents that were created during the time of the drafting of the Constitution. *See* Solan, *supra* note 12, at 58 (describing the COFEA as a corpus with "at least 100

<sup>61.</sup> United States v. Costello, 666 F.3d 1040, 1041 (7th Cir. 2012) (quoting 8 U.S.C. § 1324(a)(1)(A) (iii) (2012)); *see* Lee & Mouritsen, *supra* note 2, at 812–13; Phillips et al., *supra* note 55, at 29 (stating that the Google approach ultimately fails because "the attempt misuses a flawed tool").

<sup>62.</sup> Lee & Mouritsen, supra note 2, at 812–13.

<sup>63.</sup> Ben Zimmer, *The Corpus in the Court: 'Like Lexis on Steroids,*' ATLANTIC (Mar. 4, 2011), https://www.theatlantic.com/national/archive/2011/03/the-corpus-in-the-court-like-lexis-on-steroids/72054 [https://perma.cc/3RUH-RMQN].

<sup>64.</sup> Neal Goldfarb, Corpus Linguistics in Legal Interpretation: When Is It (In)appropriate? 2 (Sept. 5, 2019) (presented in February 2019 at the Fourth Annual Law & Corpus Linguistics Conference at BYU Law School) (available at https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3333512 [https://perma. cc/2CH8-ZQKC]).

# 2. Quantifiability and Verifiability

The other primary argument propelling the corpus linguistics movement is that corpus linguistics is "replicable and falsifiable"<sup>68</sup> and "quantifiable and verifiable."<sup>69</sup> This argument, like the argument that corpus linguistics is more accurate, is often based on a comparison to the use of dictionaries. When judges select a definition from a dictionary as the relevant meaning of legal text, they have little basis to do so. This practice not only results in the inaccurate results described above but also leaves few traces of a judge's process for selecting a particular dictionary or precedent with which to check or verify those results.

Proponents argue that corpus linguistics, however, provides a way to determine the ordinary meaning of a term based on frequencies of particular linguistic features and patterns within a large corpus.<sup>70</sup> With corpus linguistics, a judge may rely on numbers rather than intuition to determine the ordinary meaning of a word. This approach, proponents argue, is "objective" and "data-driven."<sup>71</sup> For example, judges can record actual figures for the frequency with which a word is used and the collocation of a word. By providing the queries that they used, judges allow others to recreate the same queries and verify those results. A verifiable result, in turn, is one that is accurate, reliable, and protects against bias.<sup>72</sup>

#### C. CORPUS LINGUISTICS IN THE COURTS

Corpus linguistics has experienced a meteoric rise as a method of legal interpretation. The first identifiable introduction of corpus linguistics to the world of legal interpretation appeared in a *Brigham Young University Law Review* 

million words of text written between 1760 and 1799"); see also James Cleith Phillips & Sara White, The Meaning of the Three Emoluments Clauses in the U.S. Constitution: A Corpus Linguistics Analysis of American English from 1760–1799, 59 S. TEX. L. REV. 181, 202 (2017) (introducing the COFEA); Lee J. Strang, The Original Meaning of "Religion" in the First Amendment: A Test Case of Originalism's Utilization of Corpus Linguistics, 2017 BYU L. REV. 1683, 1694 (discussing the advantages of the COFEA).

<sup>68.</sup> Lee & Mouritsen, *supra* note 2; *see also* Lee & Phillips, *supra* note 2 (describing replicability as a goal of corpus linguistics); Mouritsen, *supra* note 2 (explaining that the corpus linguistics approach is quantifiable, reliable, and repeatable); Omri Ben-Shahar, *Data Driven Contract Interpretation: Discovering "Plain Meaning" Through Quantitative Methods*, JOTWELL: CONTRACTS (June 13, 2018), https://contracts.jotwell. com/data-driven-contract-interpretation-discovering-plain-meaning-through-quantitative-methods [https:// perma.cc/RY95-YDXA] (reviewing Mouritsen, *supra* note 55).

<sup>69.</sup> Mouritsen, *supra* note 2; *see also* State v. Rasabout, 2015 UT 72,  $\P\P$  84–85, 356 P.3d 1258, 1281 (Lee, Associate C.J., concurring in part and concurring in the judgment) (describing corpus linguistics as an easy-to-use, transparent, replicable, and reliable tool); Lee & Mouritsen, *supra* note 2 (highlighting the "observable and quantifiable data" used); Crump, *supra* note 59, at 405 (describing corpus linguistics as scientific).

<sup>70.</sup> See Elizabeth S. Parks & Kristen Barta, Are You My Mother? Perpetuating Gender Inequality Through Listening Expectations and Relational Roles, 8 J. Res. GENDER STUD. 28, 33–34 (2018).

<sup>71.</sup> *See* Crump, *supra* note 59, at 402, 404; Mouritsen, *supra* note 6, at 1970 (arguing that the debate becomes "empirical" rather than "metaphysical").

<sup>72.</sup> See Lee & Mouritsen, *supra* note 2, at 867 ("The potential for subjectivity and arbitrariness is not heightened but reduced by the use of corpus linguistics. Without this tool, judges will tap into their linguistic memory to make assessments about the frequency or prototypicality of a given sense of a statutory term. Such recourse to memory and judicial intuition is neither transparent nor replicable. Nothing is statistically worse than one data point—especially a biased one." (footnote omitted)).

comment in 2010. There, Stephen Mouritsen argued that corpus linguistics offered an approach to interpretation that allowed for empirical rather than metaphysical debate.<sup>73</sup> Less than a year later, Ben Zimmer of the *Atlantic* memorialized corpus linguistics' appearance in U.S. courts: "Nowadays, corpus analysis is no longer an esoteric art for linguists and lexicographers only."<sup>74</sup> Zimmer's article took note of an amicus brief filed in *FCC v. AT&T Inc.*<sup>75</sup> that Chief Justice Roberts may have relied on in considering whether corporations are entitled to "personal privacy."<sup>76</sup> The brief, filed by Neal Goldfarb, provided a corpus linguistics analysis to support the argument that *personal* pertains to an individual, not a corporation.<sup>77</sup> This brief was followed by at least twelve similar briefs filed in the U.S. Supreme Court across multiple cases over the last decade.<sup>78</sup>

In 2018, Justice Thomas cited three corpora in his *Carpenter v. United States* dissent.<sup>79</sup> He argued that a search under the Fourth Amendment could not be interpreted to include a government intrusion on a "reasonable expectation of privacy," as long-held by the Court.<sup>80</sup> "The phrase 'expectation(s) of privacy," he wrote, did not appear in Founding Era documents.<sup>81</sup> Presumably, Justice Thomas relied on a frequency search of the cited corpora to make that statement. That footnote garnered significant attention from commentators,<sup>82</sup> both because it was

76. Id. at 400; Zimmer, supra note 63; see also Robert Ambrogi, In His Carpenter Dissent, Thomas Gives Nod to Emerging Legal Technology: What Do Words Really Mean?, ABOVE L.: EVOLVE L. (June 25, 2018, 1:30 PM), https://abovethelaw.com/legal-innovation-center/2018/06/25/ in-his-carpenter-dissent-thomas-gives-nod-to-emerging-legal-technology [https://perma.cc/F6V1-EGN8] (describing Carpenter v. United States, 138 S. Ct. 2206 (2018), where Justice Thomas cited three corpora, and—while pointing to AT&T—noting that "Carpenter may not have been the first use of corpus linguistics in a Supreme Court opinion").

77. Brief for the Project on Government Oversight et al. as Amici Curiae in Support of Petitioners, *AT&T*, 562 U.S. 397 (No. 09-1279).

78. See, e.g., Brief for Amici Curiae Corpus-Linguistics Scholars Professors Brian Slocum, Stefan Th. Gries, and Lawrence Solan in Support of Employees, Bostock v. Clayton Cty., 140 S. Ct. 1731 (2020) (No. 17-1618) (using corpus linguistics to interpret sex and gender); Brief of Neal Goldfarb as Amicus Curiae in Support of Respondents, Arguing that as to the Second Amendment Issue, the Petition Should Be Dismissed as Improvidently Granted, N.Y. State Rifle & Pistol Ass'n v. City of New York, 140 S. Ct. 1525 (2020) (No. 18-280) (using corpus linguistics to interpret the Second Amendment); Brief Amicus Curiae of the Becket Fund for Religious Liberty in Support of Petitioners at 5, Am. Legion v. Am. Humanist Ass'n, 139 S. Ct. 2067 (2019) (No. 17-1717) (referencing corpus linguistic analysis of establishment); Brief for the Cato Institute as Amicus Curiae Supporting Petitioners at 3, Am. Legion, 139 S. Ct. 2067 (No. 17-1717) (referencing corpus linguistics analysis of Establishment Clause); Brief of Amici Curiae Scholars of Corpus Linguistics, Lucia v. SEC, 138 S. Ct. 2044 (2018) (No. 17-130) (using corpus linguistics to determine the meaning of officer in the Appointments Clause).

79. 138 S. Ct. at 2238 n.4 (Thomas, J., dissenting).

<sup>73.</sup> Mouritsen, supra note 6, at 1970.

<sup>74.</sup> Zimmer, supra note 63.

<sup>75. 562</sup> U.S. 397 (2011).

<sup>80.</sup> Id. at 2238.

<sup>81.</sup> Id.

<sup>82.</sup> See, e.g., Ambrogi, supra note 76; Josh Blackman & James C. Phillips, Corpus Linguistics and the Second Amendment, HARV. L. REV. BLOG (Aug. 7, 2018), https://blog.harvardlawreview.org/corpus-linguistics-and-the-second-amendment [https://perma.cc/BQ9Q-AACB]; Karen Breda, Law & Corpus Linguistics—Free Online Resource, B.C. LEGAL EAGLE (Nov. 2, 2018), bclegaleagle.blogspot.com/2018/11/ law-corpus-linguistics-free-online.html [https://perma.cc/PY5X-SEFP]; James A. Heilpern, Corpus Linguistics

the first time a corpus had been cited by a Supreme Court Justice and because it identified a newly inaugurated corpus likely to be of significant interest to originalists: the Corpus of Founding Era American English (COFEA). This corpus includes a wide range of documents—such as letters, nonfiction books, and legal documents—written between 1760 and 1799 and covers at least 100 million words.<sup>83</sup>

Although Supreme Court Justices might be only beginning to dabble in corpus linguistics, several state and federal circuit judges are adopting the method. To date, corpus linguistics has appeared in at least ten state court and federal circuit court majority, concurring, or dissenting opinions,<sup>84</sup> with one of the most recent appearances in a Sixth Circuit concurrence in which Judge Thapar asked for future briefs to the court to include corpus linguistics: "[A]dversarial briefing on corpus linguistics can help courts as they roll up their sleeves and grapple with a term's ordinary meaning."<sup>85</sup> Litigants and amici outside of the Sixth Circuit are also taking note and including corpus linguistics analyses in their briefs.<sup>86</sup>

#### II. THE LIMITS OF CORPUS LINGUISTICS

Despite the enthusiasm for corpus linguistics among some judges and scholars, others have remained skeptical. In this Part, we catalog the objections that scholars have raised about the increasing use of corpus linguistics. To provide context, however, we first provide an illustration of how corpus linguistics analysis is used in statutory interpretation in Section II.A. In Section II.B, we follow this case study with a summary of the criticisms of corpus linguistics.

A. A BRIEF CASE STUDY OF CORPUS LINGUISTICS IN STATUTORY INTERPRETATION

So how does corpus linguistics actually work in a live case? In *People v. Harris*, the Michigan Supreme Court became the first state court of last resort to use corpus linguistics in a majority opinion.<sup>87</sup> At issue was the meaning of *information* as used in the Disclosures by Law Enforcement Officers Act (DLEOA).<sup>88</sup> The DLEOA prohibits the use in a subsequent criminal proceeding of all "information" provided by a law enforcement officer under threat of an employment sanction.<sup>89</sup> Three police officers who had been criminally charged argued that the

87. 885 N.W.2d 832, 839 (Mich. 2016).

88. See id. at 838.

89. Id. at 834.

Arrives at the Supreme Court, ORIGINALISM BLOG (June 27, 2018, 8:00 AM), https://originalismblog. typepad.com/the-originalism-blog/2018/06/corpus-linguistics-arrives-at-the-supreme-courtjames-a-heilpern.html [https://perma.cc/CC4Z-XMHY].

<sup>83.</sup> Phillips et al., *supra* note 55, at 31.

<sup>84.</sup> See supra note 3.

<sup>85.</sup> Wilson v. Safelite Grp., Inc., 930 F.3d 429, 445 (6th Cir. 2019) (Thapar, J., concurring in part and concurring in the judgment).

<sup>86.</sup> *See* Brief of *Amici Curiae* Scholars of Corpus Linguistics Supporting Petitioners, Rimini St., Inc. v. Oracle USA, Inc., 139 S. Ct. 873 (2019) (No. 17-1625); *see also* Brief of *Amici Curiae* by Certain Legal Historians on Behalf of Plaintiffs at 22 n.64, Blumenthal v. Trump, 373 F. Supp. 3d 191 (D.D.C. 2019) (No. 17-1154) (citing the corpus linguistic discussion of the meaning of the Emoluments Clause).

statements that they had made during an internal police department investigation could not be used in the criminal proceeding because they were covered under the DLEOA.<sup>90</sup> Because the police officers' statements were false, the question was whether *information* was limited to true statements or included false statements. The court turned to corpus linguistics for an answer: "The Corpus of Contemporary American English (COCA) allows users to 'analyze[] ordinary meaning through a method that is quantifiable and verifiable."<sup>91</sup>

The majority explained that it searched the corpus for the term *information* to identify its collocates—words that tend to appear with *information*: "[A]ccurate' is the most common adjective collocated with 'information' . . . . The words 'false' and 'inaccurate' are also commonly collocated with 'information."<sup>92</sup> Based on the collocate list including terms that refer to accuracy as well as terms that refer to inaccuracy, the majority concluded that *information* "can describe *either* true or false statements," and as a result, the officers' statements were covered by the DLEOA.<sup>93</sup>

Looking at the same collocate search results, Judge Stephen Markman arrived at a different conclusion in his partial dissent. He claimed that the COCA supported the interpretation of *information* to refer only to truthful statements:

The term "information" is found within the COCA 168,187 times and yet it is only modified by the term "truthful" 28 times, "true" 18 times, "accurate" 508 times, "inaccurate" 112 times, and "false" 271 times. In other words, the term "information" is modified by one of these adjectives 937 times. The other 167,250 times that the word "information" is used it is unmodified by one of these adjectives. That is, 99.44% of the time "information" in the COCA is unmodified by any of these adjectives related to veracity. Therefore, I disagree with the majority's contention that the COCA affords support for the proposition that the term "information" is "regularly" or "commonly" modified by one of these adjectives.<sup>94</sup>

Beyond illustrating how corpus linguistics might actually work in a real-world case, the dueling opinions also highlight the tension between the primary arguments in favor of corpus linguistics as a method of legal interpretation—that it is accurate and "quantifiable and verifiable"—and its most widely recognized weakness—that, ultimately, intuition and instinct continue to play a role in the corpus linguistics analysis. We further explore that tension in Section II.B below.

<sup>90.</sup> Id. at 833-35.

<sup>91.</sup> *Id.* at 839 (alteration in original) (footnote omitted) (quoting Mouritsen, *supra* note 2). For a description of COCA's content and tools, see *supra* Section I.A.

<sup>92.</sup> Harris, 885 N.W.2d at 839 n.33.

<sup>93.</sup> Id. at 833, 839.

<sup>94.</sup> *Id.* at 850 n.14 (Markman, J., concurring in part and dissenting in part). For an evaluation of the majority's and partial dissent's use of corpus linguistics, see Goldfarb, *supra* note 64, at 48–53.

### B. PRIOR WORK ON THE LIMITS OF CORPUS LINGUISTICS IN LEGAL INTERPRETATION

Proponents of corpus linguistics have recognized the need for caution as they introduce the methodology to courts.<sup>95</sup> Some critics, however, have wondered whether corpus linguistics belongs in the courts at all. Critics raise three relevant concerns. The first is methodological. Corpus linguistics, critics argue, is as fraught with inaccuracy and subjectivity as the traditional interpretation methods corpus linguistics is supposed to replace. Second, critics have raised concerns about corpora themselves. How can we be sure that the corpora being used in legal interpretation are appropriate for that task? Is the ordinary meaning of a word best found in certain kinds of text? Should judges use corpora made specifically for legal interpretation as opposed to corpora created for the broader study of language? Third, some critics have argued that corpus linguistics is outside judicial competency because it employs sophisticated empirical methods and tools with which judges are unfamiliar, and hiring experts to perform these analyses threatens to further increase the costs of litigation for parties.

# 1. The Illusion of Accuracy and Objectivity

Many critics are skeptical that corpus linguistics offers increased accuracy and objectivity when compared to traditional legal interpretation methods. Some have raised concerns with particular aspects of the methodology,<sup>96</sup> and corpus linguistics proponents have responded by recognizing the need for further refinements. But a more fundamental criticism remains: bias and intuition lurk behind corpus linguistics' cloak of empiricism. Even strong proponents of using corpus linguistics in legal interpretation admit that the methodology inevitably involves "a certain degree of subjective intuition"<sup>97</sup> and is subject to bias.<sup>98</sup> Critics argue that this renders corpus linguistics proponents have heavily criticized.<sup>99</sup> As one critic wrote: "The corpus may contain objective, empirical data, but any *analysis* of that data requires people to make decisions."<sup>100</sup> For example, a corpus user

<sup>95.</sup> Lee & Mouritsen, *supra* note 2, at 868 ("The path forward is for judges and lawyers to identify the corpus analysis that we can perform sufficiently and reliably to supplement the tools we are now using . . . . Until then we should proceed cautiously and carefully . . . ."). *See generally* Lawrence M. Solan & Tammy Gales, *Corpus Linguistics as a Tool in Legal Interpretation*, 2017 BYU L. REV. 1311 (identifying ideal preconditions for the use of corpus linguistics in legal interpretation).

<sup>96.</sup> *See, e.g.*, Herenstein, *supra* note 16, at 114 (challenging the premise that a high frequency of usage necessarily indicates a more common or more widely used meaning for a particular word); Jennifer L. Mascott, *The Dictionary as a Specialized Corpus*, 2017 BYU L. REV. 1557, 1563 (positing that, rather than using corpus linguistics to search for a particular meaning, "perhaps *all* permissible meanings . . . consistent with the statutory context should be seen as within a statute's scope").

<sup>97.</sup> Gries & Slocum, supra note 5, at 1447.

<sup>98.</sup> See, e.g., Mouritsen, supra note 2, at 203.

<sup>99.</sup> Evan C. Zoldan, *Corpus Linguistics and the Dream of Objectivity*, 50 SETON HALL L. REV. 401, 419 (2019) ("Despite the emphasis that users of corpus linguistics place on its subjectivity-reducing capabilities, corpus linguistics techniques involve significant subjective interpretive choices. These choices disrupt the dream of objectivity....").

<sup>100.</sup> Bernstein, *supra* note 12, at 455. Lawrence Solan remains cautious, though optimistic, about corpus linguistics for this very reason:

must select a corpus, create a search query, and identify which results are relevant to the legal interpretation question at hand.<sup>101</sup> Each decision can significantly affect the outcome of the analysis,<sup>102</sup> and ultimately, the corpus user will view the results through the lens of personal intuitions and biases.<sup>103</sup> This is, however, the very result that the adoption of corpus linguistics was meant to prevent.

# 2. Concerns About Corpora

Critics have also questioned whether the corpora that judges and scholars use to analyze constitutional and statutory provisions are the appropriate samples for analyzing ordinary meaning. The corpus—or the section of the corpus—that is used for a legal interpretation question will be important to the ultimate results that the corpus provides.<sup>104</sup> Corpora can draw text from

#### Solan, supra note 12.

101. Zoldan, *supra* note 99; *see also* Kevin P. Tobia, *Testing Ordinary Meaning*, 134 HARV. L. REV. 726, 734–35 (2020) (presenting experimental evidence that groups of subjects using legal corpus linguistics do not reliably reach the same, single "ordinary meaning" of a term).

102. See Bernstein, supra note 12, at 449–50 (illustrating how searching for carry versus carry combined with vehicle returns different results).

103. See Herenstein, supra note 16, at 122 (discussing ways to account for a methodological flaw the "faulty frequency hypothesis"—in many corpus linguistics analyses, but ultimately concluding that "the corpus analysis might reflect the linguistic intuitions of those engaged in the corpus analysiswhich, of course, is precisely what corpus linguistics is intended to avoid"); see also Mark C. Suchman, The Power of Words: A Comment on Hamann and Vogel's Evidence-Based Jurisprudence Meets Legal Linguistics-Unlikely Blends Made in Germany, 2017 BYU L. REV. 1751, 1767 ("Given that legal persuasion is heavily verbal, corpus linguistics could, in fact, prove to be a useful legal tool; but in this more pragmatic vision, it would serve not the scholarly agendas of truth and justice but the lawyerly agendas of advocacy and effectuation. By determining empirically which word combinations elicit which legal outcomes, practitioners could argue more compellingly (or more performatively) toward any given ends whether or not those ends were morally just, socially beneficent, or logically coherent."); Carissa Byrne Hessick, More on Corpus Linguistics and the Criminal Law, PRAWFSBLAWG (Sept. 11, 2017, 1:01 PM), https://prawfsblawg.blogs.com/prawfsblawg/2017/09/more-on-corpus-linguistics-andthe-criminal-law.html [https://perma.cc/T3NP-BHVP] ("To say that someone could conduct the same corpus search and obtain the same results is no different than saying someone could consult the same dictionary that I consult and find the same entries. But just as I might draw different conclusions from those dictionary entries, so too are people likely to draw different conclusions based on their corpus analyses.").

104. Judges most frequently use the Corpus of Historical American English (COHA), which covers American English from 1810 to 2009, and the Corpus of Contemporary American English, (COCA), which covers American English from 1990 to 2019. Each can be searched by genre and date, essentially resulting in multiple smaller corpora that judges can use, depending on what data that they think is relevant. In addition to the COHA and the COCA, judges and scholars have also explored the Corpus of Founding Era American English (COFEA), a database of words collected from texts written between 1760 and 1799. *See* Carpenter v. United States, 138 S. Ct. 2206, 2238 n.4 (2018) (Thomas, J., dissenting) (using the COHA and the COFEA); Stephanie H. Barclay, Brady Earley & Annika Boone, *Original Meaning and the Establishment Clause: A Corpus Linguistics Analysis*, 61 ARIZ. L. REV. 505, 531 (2019) (using the COFEA to analyze the Establishment Clause); Jeffrey Bellin, *Fourth Amendment* 

Like the lexicographer, the originalist, having found either too few or too many instances of a word in the corpus, will have to decide what constitutes original public meaning. And like the lexicographer, the originalist will have other choices to make about how narrowly or broadly, thinly or thickly, to construe a relevant word. These choices are not strictly linguistic. They depend upon the commitments of the corpus's user, and these commitments depend upon the user's stance with respect to the language being analyzed.

different genres, sources, geographic areas, and time periods, all of which can significantly affect the usage of words. The concerns about corpus selection are twofold.

First, corpus linguistics proponents have not developed a process for deciding what corpus—or which data within the corpus—should be used in any legal interpretation question. As a result, corpus selection involves the very intuition of the method that its proponents hope to displace. There is no readily apparent method for determining which corpus to use in any particular legal interpretation context.<sup>105</sup>

Second, besides the subjectivity of corpus selection, critics have also questioned whether general corpora—which collect language from several genres, sources, and time periods and are the most used corpora in legal interpretation—are ever appropriate for legal interpretation. That a word is included in a statute, these critics argue, may suggest that it must be interpreted in a way that is particular to its statutory usage rather than to its every-day usage.<sup>106</sup> And the magazine articles, television show transcripts, and scientific journal publications included in general corpora may be wholly inappropriate for determining the meaning of a statutory provision.<sup>107</sup> But selecting more specialized corpora may put too much stock in the corpus creator's editorial decisions of what to include in a corpus.<sup>108</sup>

# 3. Judicial Competence

A third concern raised by corpus linguistics skeptics is the judiciary's competence—or lack of competence—in employing corpus linguistics methodology. Anya Bernstein, for example, has suggested that judges are as ill-equipped to perform corpus linguistics analysis as they are to perform other kinds of empirical work.<sup>109</sup> Ultimately, she argues that judicial

106. See Zoldan, supra note 99, at 438–39.

108. See Ehrett, supra note 107, at 66-67.

*Textualism*, 118 MICH. L. REV. 233, 254 & n.142 (2019) (using the COFEA to evaluate the meaning of *search* under the Fourth Amendment); Lee & Phillips, *supra* note 2, at 296, 300–10 (using the COFEA to analyze the meaning of *commerce* under Article I of the Constitution). Scholars and commentators have also used ad hoc corpora made for a particular purpose. *See, e.g.*, Mascott, *supra* note 96, at 1576 (using the text of a Founding Era dictionary to create a searchable corpus for analyzing the term *officer*).

<sup>105.</sup> See Zoldan, supra note 99, at 420; see also Donald L. Drakeman, Is Corpus Linguistics Better than Flipping a Coin?, 109 GEO. L.J. ONLINE 81, 87 (2020) ("Beyond these concerns about whether COFEA includes a reasonable representation of ordinary language use, it is not even clear that the COFEA collection fully represents elite American speech patterns [from the Founding Era].").

<sup>107.</sup> Bernstein, *supra* note 12, at 455–56; *id.* at 458 ("[W]hat counts as the relevant kind of ordinary language for purposes of legal interpretation is not at all clear. It is not even clear whether 'ordinary' should always mean the same thing."); *see also* John S. Ehrett, *Against Corpus Linguistics*, 108 GEO. L. J. ONLINE 50, 61–64 (2019) (arguing that a corpus's "flattening" of language—giving equal weight to all instances of a word's usage—inappropriately bypasses the judicial task of prioritizing and giving credibility to certain sources).

<sup>109.</sup> See Bernstein, supra note 12, at 454 ("Does the empirical analysis of language use for statutory interpretation differ from other kinds of empirical inquiry? We would not expect, for instance, the Justices in *Chapman v. United States* to do their own chemical analysis of LSD, or the Justices deciding *Rapanos v. United States* to perform hydrological studies." (citations omitted)).

2021]

confidence in corpus linguistics stems from the same false sense of competence that has driven more traditional judicial exercises in legal interpretation:

The idea that judges should do their own empirical investigation of language use seems to rest on an assumption that language patterns are pretty easy to figure out and generally available to competent speakers—an assumption very similar to the criticism of judicial intuitions that has prompted the promotion of corpus linguistics to begin with.<sup>110</sup>

Though this concern could be mitigated by inviting parties to submit their own corpus linguistics analyses, perhaps through expert witnesses, critics have argued that such an approach significantly and prohibitively raises the costs of litigation.<sup>111</sup>

Others have suggested that it is inappropriate for judges to use corpus linguistics in legal analysis because it sacrifices notice and accountability,<sup>112</sup> and because average Americans do not have the access or expertise required to run a corpus linguistics analysis on statutes to understand them.<sup>113</sup> Some critics have pointed out that dictionaries are a huge part of "our culture's 'common linguistic experience'" and that average Americans turn to dictionaries to define words that they do not know.<sup>114</sup>

# III. MEASURING HIDDEN BIAS

Missing from those criticisms of legal corpus linguistics, however, is any discussion of the possibility of structural bias within the corpora themselves. The potential problem is not that a corpus, such as the COHA, is improperly assembled. Rather, the problem is that a corpus, properly constructed, reflects the underlying prejudices against certain groups in society and that such animus infects the texts in the corpus. The legacies of racism, sexism,

<sup>110.</sup> *Id.* This is a concern that other critics share. *See, e.g.*, Ehrett, *supra* note 107, at 66–67; *id.* at 69 ("[Corpus linguistics] is decidedly *not* a 'modest and simple' burden to place upon sitting judges, many of whom do not rely extensively on modern Internet-driven technologies.").

<sup>111.</sup> Hessick, *supra* note 12, at 1515–16; *id.* at 1515 ("[M]embers of the general public cannot be expected to perform their own corpus searches and analyses. The process described in the corpus linguistics literature appears quite involved, and it hardly seems accessible to the average American."); *see also* State v. Rasabout, 2015 UT 72, ¶ 18, 356 P.3d 1258, 1265 (suggesting that litigants would need to hire experts to perform corpus linguistics analyses); Ehrett, *supra* note 107, at 69–70 (describing the "catch-22" that the larger and more representative a corpus becomes, the more burdensome it is to use); Daniel C. Tankersley, Comment, *Beyond the Dictionary: Why* Sua Sponte *Judicial Use of Corpus Linguistics Is Not Appropriate for Statutory Interpretation*, 87 MISs. L.J. 641, 673 (2018) ("Justice Durham suggested that corpus data should be presented by *the parties* so that the court can have 'meaningful tools' at its disposal for interpretive tasks. . . . [T]he adversarial process is not compromised, because the judge is now performing one of her established tasks: assessing the reliability of conflicting proofs brought before the court." (citation omitted)).

<sup>112.</sup> Tankersley, *supra* note 111, at 669.

<sup>113.</sup> Hessick, *supra* note 12, at 1515–16.

<sup>114.</sup> Mascott, supra note 96, at 1570–71.

homophobia, and other prejudices in the United States make the likelihood of such biases almost certain. Yet, no attention has been paid to how we might identify the ways in which such biases are reflected in a corpus and what we should do about it.

This Part of the Article takes a first step toward identifying hidden bias in a corpus. It introduces a quantitative method for measuring gender bias in text. It then uses that method to measure gender bias in the COHA, a popular corpus used in legal corpus linguistics. Unsurprisingly, the results of that study provide evidence of gender bias in the COHA.

The results also provide additional, more specific findings. The methods used here allow us to measure how gender bias changes in the COHA over time. Bias appears to wax and wane across the decades, and although there is modest evidence that the texts in the COHA have become less sexist over the years, there is still ample evidence of bias in twenty-first century texts. Furthermore, and of particular interest, our findings show that the use of discrete words changes over time. Our methods allow us to track the change of a word's usage with precision, and it is not uncommon for word usage to change materially over the years. Finally, we can also measure gender bias in texts written by male and female authors. We find evidence of bias regardless of author gender, although bias is less pronounced in texts written by female authors.

This Part of the Article proceeds as follows. First, we introduce the data that was used in the analysis and the hypotheses that we tested. Second, we provide an introduction to the machine learning methods that we used to measure gender bias in the COHA. We took pains to draft this introduction in plain English, appropriate for a nontechnical audience. Third, we present the results of the analysis using accessible, descriptive statistics and visualizations.

#### A. DATA

The study uses two sets of data to analyze gender bias in the COHA. The primary dataset is the COHA itself, a corpus of approximately 115,000 individual texts totaling over 385 million words; covering U.S. publications from 1810 to 2009; and balanced between fiction books, nonfiction books, magazines, and newspapers.<sup>115</sup> To test the hypothesis that author gender affects the extent of gender bias in the COHA, we used a second dataset that provides basic COHA information, including author names.<sup>116</sup>

<sup>115.</sup> See Introduction: Overview, FULL-TEXT CORPUS DATA, https://www.corpusdata.org [https:// perma.cc/XK32-XVCF] (last visited Jan. 10, 2021) (describing the composition of the COHA). The underlying text of the COHA is publicly available for a licensing fee. See Purchase, FULL-TEXT CORPUS DATA, https://www.corpusdata.org/purchase.asp [https://perma.cc/LG4S-MCP9] (last visited Nov. 6, 2020).

<sup>116.</sup> See Format/Samples: Overview, FULL-TEXT CORPUS DATA, https://www.corpusdata.org/ formats.asp [https://perma.cc/2A3P-ZXCS] (last visited Jan. 10, 2021). This second dataset is publicly available without a licensing fee. *Id*.

#### B. HYPOTHESES

This study's primary concern is the possibility that hidden bias in the texts of a corpus may infect judicial decisionmaking that relies upon corpus linguistics methods. There are, of course, many types of bias—gender, race, sexual orientation, age, among others—that could be hidden in a body of text. Here, we focus exclusively on gender bias for two reasons. First, gender bias is itself a policy concern of critical importance. Second, as explained in more detail below, analyzing gender bias in a corpus is more methodologically straightforward than other biases. Examining additional biases is an important priority for subsequent research, and we hope that the gender bias analysis introduced here can serve as a conceptual and methodological foundation for those later studies.

The study tests the following primary hypothesis: The texts comprising the COHA tend to associate certain ostensibly neutral words with either males or females in a way that exhibits gender bias. It also tests two subsidiary hypotheses. First, the extent of gender bias in the COHA decreases with time. Second, the extent of gender bias increases if the author of a given text is male rather than female. The first subsidiary hypothesis is tested to provide a sense of how gender bias evolves over time, an important issue in an historical interpretive approach like originalism. The second subsidiary hypothesis is tested to illuminate a possible origin of any observed gender bias in the text.

#### C. METHODS

To analyze gender stereotypes in the COHA, we use a machine learning method that examines "word embeddings" in natural language text.<sup>117</sup> This method has the same starting point as the corpus linguistics methods discussed in Part I. All of these methods assume that the meaning of a word is "usage based." In other words, how a word is used in context tells us something important about what it means.<sup>118</sup> For instance, analyzing usage in context provides a way to determine whether two words have similar or different meanings.<sup>119</sup> Linguistics scholars refer to this as the "distributional hypothesis"—the degree of semantic similarity between two words is a

<sup>117.</sup> See generally Tolga Bolukbasi, Kai-Wei Chang, James Zou, Venkatesh Saligrama & Adam Kalai, Man Is To Computer Programmer As Woman Is To Homemaker?: Debiasing Word Embeddings, 2016 ADVANCES NEURAL INFO. PROCESSING SYSS. 1; Aylin Caliskan, Joanna J. Bryson & Arvind Narayanan, Semantics Derived Automatically from Language Corpora Contain Human-Like Biases, 356 SCIENCE 183 (2017); Nikhil Garg, Londa Schiebinger, Dan Jurafsky & James Zou, Word Embeddings Quantify 100 Years of Gender and Ethnic Stereotypes, 115 PROC. NAT'L ACAD. SCI. U.S. E3635 (2018); Jieyu Zhao, Tianlu Wang, Mark Yatskar, Vicente Ordonez & Kai-Wei Chang, Men Also Like Shopping: Reducing Gender Bias Amplification Using Corpus-Level Constraints, PROC. 2017 CONF. ON EMPIRICAL METHODS NAT. LANGUAGE PROCESSING 2979 (2017).

<sup>118.</sup> Alessandro Lenci, *Distributional Semantics in Linguistic and Cognitive Research*, 20 RIVISTA LINGUISTICA 1, 1 (2008).

<sup>119.</sup> Id. at 2.

function of the similarity of the linguistic contexts in which those words appear.<sup>120</sup> Put more simply, we can know the meaning of "a word by the company it keeps."<sup>121</sup>

The analysis of word embeddings differs from the corpus linguistics methods introduced above in an important way, however. Analyzing word embeddings converts words into numerical *vectors* and, in turn, a corpus into a collection of those vectors called a *vector space*.<sup>122</sup> Representing a corpus in vector space allows a quantification of a word's context, which in turn, allows for statistical tools to be employed to analyze relationships among words across the text of a corpus. Let's unpack how this is done.

The following simple, lighthearted example provides a starting point for understanding what it means to quantify the relationships between words in this fashion. This example is derived from Allison Parrish's excellent overview for beginners of how word vectors are calculated and used for text analysis.<sup>123</sup> The example illustrates how we can think of a collection of related items as vectors and how depicting them as vectors allows us to study their relationships. Once that is demonstrated, we can then take the next step of understanding how words in a corpus can be depicted as vectors.

Imagine that we are interested in categorizing a variety of animals along two characteristics: their size and a subjective assessment of their "cuteness." We can then represent each animal as a vector in that two-dimensional space. Here, there are two numbers in the sequence that identify an animal vector: the animal's size and its cuteness. On that basis, some of the animals in our sample are small in size but quite cute, such as kittens. Others are large but not cute, such as crocodiles. Some are both small and not cute, such as tarantulas, or large and cute, such as panda bears. And so on. We can then plot the animals as vectors in a two-dimensional vector space, as visualized in Figure 1 below.

<sup>120.</sup> Id. at 2–3. For the original work on the distributional structure of language, see Zellig S. Harris, *Distributional Structure*, 10 WORD 146 (1954).

<sup>121.</sup> J. R. Firth, A Synopsis of Linguistic Theory, 1930–1955, in STUDIES IN LINGUISTIC ANALYSIS 1, 11 (1962).

<sup>122.</sup> A vector is a sequence of numbers identifying the location of a point in space.

<sup>123.</sup> See Allison Parrish, Understanding Word Vectors, GITHUB, https://github.com/aparrish/rwet/ blob/master/understanding-word-vectors.ipynb [https://perma.cc/LMF6-JX6C] (last visited Jan. 10, 2021) (including explanations and a sample Python code). Ratuerberg and Talley also provide an accessible introduction to machine learning methods, which they employ to analyze corporate opportunity waivers. See generally Gabriel Rauterberg & Eric Talley, A Machine Learning Classifier for Corporate Opportunity Waivers (Columbia Law & Econ., Working Paper No. 553, 2016), https://ssrn. com/abstract=2849491 [https://perma.cc/TJL5-9EKZ].



Figure 1: Plotting a Sample of Animals Along Two Criteria: Size and "Cuteness"

The information presented in Figure 1 allows us to compare the animals in our imaginary sample. We can, for instance, assess the similarity among the sampled animals along the two dimensions of interest. Simply eyeballing the visualization of the data in Figure 1 reveals many relationships: puppies and kittens are similar in both size and cuteness; kittens and lobsters are similar in size but exceptionally different in terms of cuteness; dolphins and elephants are similar in cuteness but quite different in size; and so on. We can go a step further, however, and quantify the similarity. For example, we can calculate the Euclidean distance<sup>124</sup> between animals. Doing so reveals that the distance between capybara (30 size, 70 cuteness) and panda (40 size, 74 cuteness) is 10.77. That distance is shorter than, say, the distance between tarantula (3 size, 8 cuteness) and elephant (90 size, 65 cuteness), which is 104.01, reflecting the intuition that pandas are more similar to capybaras than tarantulas are to elephants.

So far, so good. Now we can take the next step of creating a similar model for words occurring in a corpus of text. In the previous model, we used two characteristics to represent a particular animal as a vector. Now, we are going to use the context in which words are employed in a corpus to represent each word as a vector. This is accomplished by tallying the adjacencies among the words in a corpus in a large—often extremely large—spreadsheet. That gives us a sense of "where"

<sup>124.</sup> Euclidean distance is the distance between two points in a two-dimensional vector space.

in a corpus words are found. Consider Table 1 below, which is a matrix that records the associations among words in the famous line opening Dickens's *A Tale of Two Cities*: "It was the best of times, it was the worst of times."<sup>125</sup> The first column in Table 1 includes all of the words that occur in the famous line. The first row contains the combinations with adjacent words that are possible for each of the words occurring in the line—that is, the rows contain the "context" for each word listed in the first column. The intersecting cells then tally how often the word in the first column occurs within that combination. So, for example, the word "was" occurs in the combination "it was the" two times.

	was	it the	was best	the of	best times
it	1	0	0	0	0
was	0	2	0	0	0
the	0	0	1	0	0
best	0	0	0	1	0
of	0	0	0	0	1
times	0	0	0	1	0
worst	0	0	0	1	0

 Table 1: Exemplary Word Adjacency Matrix

	of it	times was	was worst	worst times	of
it	0	1	0	0	0
was	0	0	0	0	0
the	0	0	1	0	0
best	0	0	0	0	0
of	0	0	0	1	0
times	0	0	0	0	1
worst	0	0	0	0	0

That matrix gives us a sense of the company that each word in the corpus keeps. And the sequence of numbers in the intersecting cells of that matrix can be used to represent each of the words in the first column as a vector. For instance, the vector for the word *it* is [1, 0, 0, 0, 0, 0, 1, 0, 0, 0]. That is a more complicated sequence of numbers than the two numbers that we used to identify our animals above, but the same principles apply. We can calculate the vectors for all of the words in our corpus using the matrix and then perform simple calculations to analyze their relationships, just like we did in our animals example. For instance, in the phrase, "It was the best of times, it was the worst of times," the words *best* and *worst* have exactly the same vector, [0, 0, 0, 1, 0, 0, 0, 0, 0], because they occur the same number of times in the same context—"the <u>best</u> of" and "the <u>worst</u> of." In other words, the Euclidean distance between the two words is zero, suggesting that they are similar. And, intriguingly, they in fact are: *best* and *worst* have related meanings because they are antonyms.

<sup>125.</sup> CHARLES DICKENS, A TALE OF TWO CITIES 1 (1859).

Word vectors provide a quantifiable way to measure associations among words. Measuring the distance between words in the vector space tells us whether they are closely or distantly associated with one another.<sup>126</sup> Words that are more closely associated with one another in the text will have vectors with a shorter Euclidean distance between them.<sup>127</sup> That distance provides the basis for measuring gender bias in a corpus of text.

By analyzing word vectors, we can see whether words that should be neutral, such as an occupation (for example, *doctor* or *nurse*) or an adjective (for example, attractive or brave), are more closely associated with male or female words (for example, he, she, man, or woman). Specifically, this Article follows a recent article by Professors Garg, Schiebinger, Jurafsky, and Zou that uses the distance between word vectors to measure gender bias in a number of corpora.<sup>128</sup> Their method produces a simple metric that compares the distance of female words and male words to a set of words that on their face should be gender neutral, such as occupations or adjectives.<sup>129</sup> For instance, the distance between the word man and the word *doctor* can be compared to the distance between the word *woman* and the word *doctor*. If *woman* is closer to *doctor* than *man* is, we can infer that there is a gender stereotype with respect to how the word *doctor* is used—that is, women are more closely associated with doctor than men are (and vice versa). The comparison between male and female words is undertaken by subtracting the average distance between a set of male words and a neutral word (such as an occupation or an adjective) from the average distance of a set of female words to that same neutral word.<sup>130</sup> If the value is negative, then the text more closely associates the occupational word or adjective with men than women; if the value is positive, then the text more closely associates the occupational word or adjective with women than men.<sup>131</sup>

We study the same set of occupational words and set of adjectives that Garg et al. use. The occupational word set includes seventy-six occupations, such as *judge*, *lawyer*, *teacher*, and *engineer*.<sup>132</sup> The adjective set includes 230 words,

130. Id. at E3636.

131. Id.

<sup>126.</sup> See Garg et al., supra note 117, at E3636.

<sup>127.</sup> See id.

<sup>128.</sup> Id. at E3635.

<sup>129.</sup> *Id.* at E3636. The approach of Garg et al., as well as our own, assumes that gender is binary. *See id.* at E3636 n.†. Our approach does not reflect a normative or positive position that gender is, in fact, binary; rather, it simply reflects the common assumption through much of U.S. social history of binary gender identification. This means that the shift in pronoun usage to reflect a nonbinary conception of gender, which is currently underway in parts of U.S. society, may make this approach less reliable when analyzing corpora of contemporary texts going forward.

<sup>132.</sup> See NIKHIL GARG, LONDA SCHIEBINGER, DAN JURAFSKY & JAMES ZOU, APPENDIX A: DATA TO WORD EMBEDDINGS QUANTIFY 100 YEARS OF GENDER AND ETHNIC STEREOTYPES 2 (2018). Software tools for applying the analysis of Garg et al. are available on a GitHub repository. EmbeddingDynamicStereotypes, GITHUB, https://github.com/nikhgarg/EmbeddingDynamicStereotypes [https://perma.cc/YG5Z-EV22] (last visited Jan. 11, 2021). The method of Garg et al. is one of the most recent techniques in the young but growing field of research that uses word embeddings to measure gender bias, ethnic bias, or both in text. See,

# such as resourceful, forgiving, thrifty, and resentful.<sup>133</sup>

Once the bias metric is calculated for each word in the occupational word set and adjective set, we follow Garg et al. by using simple descriptive statistics to assess the overall gender bias in the COHA.<sup>134</sup> Common measures of central tendency, such as mean and median figures, tell us whether the sets of occupational words and adjectives skew male or female. For instance, a mean or median of -0.20 for a set of neutral words would indicate that those words tend to be more associated with males. Measures of dispersion, such as the standard deviation, are also important, because they provide a sense of how extreme the gender bias is, toward either males or females, for the set of neutral words that are studied. For instance, a greater standard deviation would indicate that a number of words are particularly male associated and others particularly female associated.

#### D. RESULTS

Our analysis finds evidence of gender bias in the COHA. That bias appears in how both occupational words and adjectives are used; many occupations are associated with a particular gender, as are many adjectives. The dynamics of gender bias over time are particularly interesting. Overall, gender bias becomes less pronounced over time, although it still persists to the latest texts in the COHA. Importantly, the bias measures for any single word can change materially from decade to decade, and it is not unusual for a given word to experience significant bias "swings" over time. Finally, the trends with respect to authorship are also interesting. Texts authored by both males and females exhibit gender bias. However, gender bias is less extreme in texts that are authored by females.

# 1. Both Occupational Words and Adjectives in the COHA Reflect Gender Bias

We observer gender bias in both the set of occupational words and the set of adjectives. Figures 2a and 2b below indicate how the embedded bias measures for each word are distributed in the respective word sets. Evidence of gender bias would be weak if the distributions have means at or close to zero, the value at which a given word is equally associated with males and females, and the distributions are more concentrated around the mean.

*e.g.*, Bolukbasi et al., *supra* note 117; Caliskan et al., *supra* note 117; Elliott Ash, Daniel L. Chen & Arianna Ornaghi, *Stereotypes in High-Stakes Decisions: Evidence from U.S. Circuit Courts* (Working Paper, 2020), https://users.nber.org/~dlchen/papers/Stereotypes\_in\_High\_Stakes\_Decisions.pdf [https://perma.cc/XM9Z-MPMW].

<sup>133.</sup> GARG ET AL., *supra* note 132, at 2–3, 33 (citing, among others, JOHN E. WILLIAMS & DEBORAH L. BEST, MEASURING SEX STEREOTYPES: A MULTINATION STUDY (rev. ed. 1990) and John E. Williams & Deborah L. Best, *Sex Stereotypes and Trait Favorability on the Adjective Check List*, 37 EDUC. & PSYCHOL. MEASUREMENT 101 (1977) as sources of the words included in the adjective list).

<sup>134.</sup> Garg et al. find evidence of gender bias in a number of corpora, including the COHA, although they do not analyze nineteenth-century texts in the COHA. Garg et al., *supra* note 117, at E3637–38.



What the results show, however, are skewed distributions with nonzero median values. In the occupational word set, whose distribution is depicted in Figure 2a, the distribution is skewed to the right and the mean embedded bias measure is -0.0827502, indicating that most of the occupational words analyzed in the COHA are associated with men. Interestingly, the mean embedded bias measure in the adjectives word set is 0.0176891, which indicates that the adjectives analyzed in the COHA are more associated with women. Figure 2b below depicts the distribution for the adjective word set.

What are the most biased usages identified by this method? Figures 3 and 4 below unpack the tails of the gender bias distributions for the occupational word and adjective sets primarily to give the reader a better sense of whether the method is accurately measuring stereotypes. Figure 3 plots the ten most gender-biased occupations in the occupational word set. There are few surprises. The most female-biased words in the set are *nurse*, *house-keeper*, and *midwife*. The most male-biased are *postmaster*, *sheriff*, and *surveyor*.

Figure 4 plots the ten most gender-biased adjectives. Again, there are few surprises. The most female-biased adjectives analyzed are *feminine*, *charming*, and *gentle*. The most male-biased are *obnoxious*, *unscrupulous*, and *autocratic*.

In summary, the results of the general bias analysis are unsurprising. We see evidence of biased usage of both occupational words and adjectives in the COHA, suggesting that the authors of the text in the COHA were influenced by existing gender stereotypes. For anyone familiar with the arc of U.S. social history, this is what we would expect to find.

2. The Intensity of Gender Bias Changes over Time

This Section reports the results of tests of the subsidiary hypothesis that gender bias grows worse as one goes further back in time in the COHA. Texts in the COHA are analyzed by decade from 1810 to 2010. The results from both sets of

# Figure 3: Top Ten Most Gender-Biased Selected Occupational Words in the COHA



Figure 4: Top Ten Most Gender-Biased Selected Adjectives in the COHA



words show more pronounced bias in the nineteenth century, moderating but by no means fully disappearing, over the twentieth century and into the twenty-first. In the occupational word set, the intensity of gender bias waxes and wanes somewhat over the decades. In the adjective set, the pattern is particularly clear, with 2021]

both means approaching zero and standard deviations decreasing as time progresses.

We turn first to the results for the occupational word set. Here, key highlights are presented and discussed.

Gender bias in occupational words was most extreme in the nineteenth century. The 1890s have the most male-biased mean value (M = -0.1125098) followed by the 1820s (M = -0.1067916). Many of the other decades in the nineteenth century have mean bias values close to that -0.1 level. Interestingly, however, the decade with the greatest standard deviation is later: the 1940s (SD = 0.092157). Box plots of the gender bias distribution for each decade are presented in Figure 5 below, though a number of decades are not included to provide a clear presentation.

# Figure 5: Distribution of Bias Among Selected Occupational Words in the COHA by Decades



To test the hypothesis that the level of gender bias changes over time, we performed a one-way, between-groups analysis of variance (ANOVA) test.<sup>135</sup> As Table 2 reports below, that analysis yielded a statistically significant effect, and therefore the null hypothesis that there are no differences between the decade means is rejected. A statistical comparison of the means from each decade, which is not reported here, reveals that the differences between decades change incrementally. No differences between consecutive decades are statistically significant; it is only when decades are compared with decades thirty or more years later that differences between decade gender bias means are significant.

<sup>135.</sup> Prior to conducting the test, we evaluated ANOVA's assumption of normality, which we found to be satisfied: The decades' gender bias distributions are associated with skew and kurtosis less than |2.0| and |9.0| respectively. *See generally* Emanuel Schmider, Matthias Ziegler, Erik Danay, Luzi Beyer & Markus Bühner, *Is It Really Robust?: Reinvestigating the Robustness of ANOVA Against Violations of the Normal Distribution Assumption*, 6 METHODOLOGY: EUR. J. RES. METHODS FOR BEHAV. & SOC. SCI. 147 (2010). Additionally, we tested the assumption of homogenous variances with the Brown–Forsythe modified Levene F test, which we found to be satisfied: F(19, 1421) = 1.5128810, p = 0.07207286.

The analysis of the adjective set produced similar results. Gender bias in the use of adjectives is also most extreme in the nineteenth century. However, adjec-

Source	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	0.618346487	19	0.032544552	5.46	0.0000
Within Groups	8.46285479	1421	0.005955563		
Total	9.08120127	1440	0.00630639		

 Table 2: ANOVA Results, Gender Bias in Occupations by Decade

tive use is biased toward women, a different trend than occupational words, which are biased toward men. In the COHA, men are more closely associated with work, and women are more often the objects of description. The 1860s have the most female-biased mean value (M = 0.0404672) followed by the 1900s (M = 0.03224602). The decades with the greatest standard deviations are the 1820s (SD = 0.10117093), the 1810s (SD = 0.09768301), and the 1940s (SD = 0.0880743). Box plots of the gender bias distribution for each decade are presented in Figure 6 below, though again, numerous decades are omitted to achieve clear presentation.

# Figure 6: Distribution of Bias Among Selected Adjectives in the COHA by Decades



To test the hypothesis that the level of gender bias changes over time, we again performed a one-way, between-groups ANOVA test.<sup>136</sup> As Table 3 reports below, that analysis yielded a statistically significant effect, and therefore the null hypothesis that there are no differences between decades is rejected. A statistical

<sup>136.</sup> We found ANOVA's assumption of normality to be satisfied: The decades' gender bias distributions are associated with skew and kurtosis less than |2.0| and |9.0| respectively. We tested the assumption of homogenous variances with the Brown–Forsythe modified Levene F test, which we found not to be satisfied: F(19, 4394) = 13.002543, p = 0.0000. These results suggest caution when relying upon the results of this ANOVA test.

comparison of the means from each decade, which is not reported here, reveals that the differences between decades changed incrementally. As with the occupational word analysis, no differences between consecutive decades are statistically significant; gender bias differences are significant only when the means of nineteenth-century decades are compared with the means of twentieth-century decades.

Table 3: ANOVA Results, Gender Bias in Adjectives by Decade

Source	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	0.972236476	19	0.051170341	9.04	0
Within Groups	24.8728445	4393	0.005661927		
Total	25.845081	4412	0.005857906		

Changes over time are particularly interesting when the bias measure for single words are tracked over time. The bias associated with a word can change materially over decades. Figure 7 below plots the shifting bias measures for four occupational words—*dancer*, *librarian*, *nurse*, and *judge*—over time. *Nurse* and *judge* are persistently associated with females and males, respectively, over the years. *Librarian* and *dancer* shift from more to less biased associations over time.

Figure 7: Change in Bias of Four Occupational Words in the COHA by Decade



We observe similar dynamism in the four adjectives presented in Figure 8 below. Figure 8 plots the bias measures of *attractive*, *modest*, *organized*, and *tough* for each decade of the COHA. *Attractive* is persistently associated with females, and *organized* and *tough* are associated with males. *Modest* experiences

a dramatic shift in association over the years, going from a female-associated word to a neutral, or slightly male-associated, word by the end of the COHA.



Figure 8: Change in Bias of Four Adjectives in the COHA by Decade

Taken together, the results of the gender bias analysis across the decades of the COHA teach us many lessons. First, gender bias may grow less extreme over time, but it never disappears entirely. Gender bias is a twenty-first-century problem as much as a nineteenth-century one. Thus, as far as the COHA is concerned, there is no purely unbiased part of the corpus that could be used for legal interpretation. Second, the bias associated with any given word can be highly dynamic over time. A corpus linguistics analysis of a word may incorporate exceptionally different levels of bias depending on which era of the COHA is being studied. This variation means that any attempts to "control" for bias associated with a certain word (for example, a judge adjusting an interpretation of the word *modest* by assuming that the word is always associated with women) may lead to inaccurate results.

3. Both Male and Female Authors Write Biased Text, but Female Authors Less So

This Section reports the results of tests of the subsidiary hypothesis that gender bias differs according to the gender of the text's author. To identify author gender, we analyzed the first given names of the authors of the COHA.<sup>137</sup> The analysis used the *gender* package in the data analysis software R to estimate the

<sup>137.</sup> See Format/Samples: Overview, supra note 116. This second dataset is publicly available without a licensing fee.

genders of authors based on their first names.<sup>138</sup>

Identifying author gender in this way revealed, unsurprisingly, that the vast majority of identifiable authors in the COHA are men. In every decade of the COHA, male authors outnumber females, usually by several orders of magnitude. Only in the 1990s did female authors account for fifty percent or more of all authors in the corpus. The lack of identifiable female authors in the COHA means that we should interpret the results here with particular caution—some of the patterns that we observe may be affected simply by the low number of female authors in the COHA underscores the severity of the gender bias problem. If females lack visibility, it is because female authors are so underrepresented.

The results of the analysis are mixed. The results from the occupational word set, shown in Figure 9a, reveal no statistically significant difference in the way that female and male authors associate occupations with genders. The mean bias for texts with identifiable female authors is -0.0540046 (SD = 0.005002), which means that female authors tend to associate the occupational words in the sample with males. The mean bias value for texts with identifiable male authors is -0.0667847 (SD = 0.0088223), which also reflects a tendency to associate occupations with males. The difference between the texts authored by females and by males is small and not statistically significant.<sup>139</sup>

The results for the adjective set, however, do provide evidence of a significant difference in how female and male authors associate adjectives with genders. As Figure 9b shows, the mean bias value for texts with identifiable female authors is 0.0183497 (SD = 0.0602622), and the mean bias value for texts with identifiable male authors is 0.0489919 (SD = 0.69457). The difference between the texts authored by female and male authors is fairly small but nevertheless statistically significant.<sup>140</sup>

Taken together, these results are somewhat surprising. First, the difference between female and male authors is not as dramatic as one might expect. Second, although adjective usage suggests that male authors are more biased in the way that they write, both male and female authors nevertheless write biased text.<sup>141</sup>

<sup>138.</sup> We used the *gender* package to estimate author gender because the COHA user interface cannot filter between male and female authors. The *gender* package, which we understand to be frequently used by researchers, relies on the Integrated Public Use Microdata Series to estimate gender based on a given name for the years prior to 1930 and Social Security Administration data to estimate gender for the years 1930 to 2009. LINCOLN MULLEN, CAMERON BLEVINS & BEN SCHMIDT, PACKAGE 'GENDER' 3–4 (2020), https://cran.r-project.org/web/packages/gender/gender.pdf [https://perma.cc/PK7G-VRUQ].

<sup>139.</sup> Comparison of means using a simple t-test did not find a statistically significant difference between the samples of male- and female-authored texts: t(150) = 1.0432, p = 0.8507.

<sup>140.</sup> Comparison of means using a simple t-test found a statistically significant difference between the samples of male- and female-authored texts: t(458) = -5.0537, p = 0.000.

<sup>141.</sup> Again, care should be taken when interpreting these results given the limited incidence of female authors in the COHA.

#### E. SUMMARY

If we are not careful, using corpus linguistics to inform legal interpretation could result in courts importing the biases latent in the texts comprising a corpus into their decisionmaking. The empirical results presented in this Part provide evidence of gender bias in the COHA. Analysis of both occupational words and adjectives in the texts comprising the COHA provides evidence that the usage of those words reflects gender stereotypes. For example, nurses and housekeepers tend to be women, and women are more often described as charming and gentle. Sheriffs and carpenters tend to be men, and men are more often described as vindictive and autocratic.

The analysis also identifies notable patterns in the evolution of gender bias over time. Of course, one headline finding in the results is that gender bias persists to the present. Bias is not solely a matter of history. However, particularly with respect to adjective usage, gender bias is stronger in the nineteenth century. The analysis of occupational words highlights the dynamism of usage over time; the magnitude of gender bias waxes and wanes over the decades, suggesting that usage is not fixed. Analysis of individual words underscores that point: the usage of some words is dynamic—for example, some occupations, such as librarian, that were associated more with men in the nineteenth century became more associated with women in the twentieth.

Taken together, these results suggest that originalist approaches to legal interpretation should proceed with caution. On one hand, some of the evidence suggests that the further back we go, the worse bias grows, which has obvious implications for originalist approaches to the interpretation of the U.S. Constitution and longstanding statutes. The results also suggest, however, that interpretive questions involving more recent constitutional or statutory text are also vulnerable to the incorporation of bias if the COHA is referenced to determine public meaning. Finally, some evidence suggests that the gender of the author influences the magnitude of bias in a text, raising the possibility that historically uneven representation of authors in the COHA might skew interpretation.

## IV. NEXT STEPS FOR EMPIRICAL APPROACHES TO LEGAL INTERPRETATION

One way to understand the results in Part III is in terms of the algorithmic bias problem that has gained widespread notoriety in recent years. Algorithmic bias occurs when, for example, the machine learning algorithms that power the artificially intelligent "chatbots" that simulate human conversations produce output that reflects the prejudices of the underlying data on which the algorithms rely.<sup>142</sup> Microsoft's Tay, a chatbot, is a leading example of the problem. Tay's protocol fed into its algorithm language from the interactions the chatbot was having with humans on Twitter, which meant that biased comments in Tay's Twitter feed would be incorporated into the underlying data from which Tay constructed its

<sup>142.</sup> See, e.g., Ayanna Howard & Jason Borenstein, The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity, 24 SCI. & ENGINEERING ETHICS 1521, 1522 (2018).

conversation snippets.<sup>143</sup> In less than twenty-four hours, as people fed discriminatory material into Tay through their tweets, Tay "turned into a brazen anti-Semite and was taken offline."<sup>144</sup>

The results in Part III imply that a similar problem can happen in legal corpus linguistics. Our results describe one kind of hidden bias—gender bias—that might infect legal interpretation. Of course, the gender bias studied in this Article is unlikely to be relevant to all interpretive questions. Gender bias may not be directly relevant to interpreting Section 7 of the Clayton Act, which prohibits anticompetitive mergers and acquisitions, for instance.<sup>145</sup> At the same time, however, biases infect many legal issues, from family law and civil rights to sexual harassment and corporate governance.<sup>146</sup> Drawing definitive boundaries between legal issues that are or are not affected by gender bias is difficult.

How then should we address the problem of bias in a corpus? The growing literature on algorithmic bias provides a guide for how we might think of the problem. Prescriptions for addressing algorithmic bias tend to fall in two categories. The first is to make bias in the underlying data transparent.<sup>147</sup> The second goes a step further and attempts to "debias" the results of the algorithm or, in other words, to engage in "fairness-aware" data mining.<sup>148</sup>

This Part argues that the first approach—making bias transparent—is a clear priority for practitioners of legal corpus linguistics and for any other approach that leverages data for empirical insight on interpretive questions. It then turns to possibilities for the second approach—debiasing corpora. This discussion proceeds with caution because we do not yet have the transparency needed to fully

147. Sara Hajian, Francesco Bonchi & Carlos Castillo, *Algorithmic Bias: From Discrimination Discovery to Fairness-Aware Data Mining*, 2016 KDD 2125, 2126, https://perma.cc/9DNW-2NA7 (referring to this approach as "discrimination discovery").

148. Id.

<sup>143.</sup> Asha Barbaschow, *Microsoft and the Learnings from Its Failed Tay Artificial Intelligence Bot*, ZDNET (July 24, 2019, 3:46 AM), https://www.zdnet.com/article/microsoft-and-the-learnings-from-its-failed-tay-artificial-intelligence-bot [https://perma.cc/D59S-FGFR]; James Vincent, *Twitter Taught Microsoft's AI Chatbot to Be a Racist Asshole in Less than a Day*, VERGE (Mar. 24, 2016, 6:43 AM), https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist [https://perma.cc/3PG9-BHUL].

<sup>144.</sup> Barbaschow, supra note 143.

<sup>145.</sup> Clayton Act, Pub. L. No. 63-212, ch. 323, § 7, 38 Stat. 730, 731–32 (1914) (codified as amended at 15 U.S.C. § 18 (2018)).

<sup>146.</sup> See, e.g., Cynthia A. McNeely, Lagging Behind the Times: Parenthood, Custody, and Gender Bias in the Family Court, 25 FLA. ST. U. L. REV. 891, 895 (1998); Richard A. Warshak, Gender Bias in Child Custody Decisions, 34 FAM. & CONCILIATION CTS. REV. 396, 396 (1996); Camille E. LeGrand, Note, Rape and Rape Laws: Sexism in Society and Law, 61 CALIF. L. REV. 919, 919 (1973) (arguing that, due to implicit biases "largely based on traditional attitudes about social roles and sexual mores," "rape laws are not designed, nor do they function, to protect a woman's interest in physical integrity"). Resnik's reflections on gender bias in legal academia and the court system more broadly is also indicative of the scale of the issue. See Judith Resnik, Gender Bias: From Classes to Courts, 45 STAN. L. REV. 2195 (1993). Gender bias is also an important issue in corporate governance, where issues such as the lack of representation of women on corporate boards is a concern. See SHEARMAN & STERLING LLP, CORPORATE GOVERNANCE & EXECUTIVE COMPENSATION SURVEY 2019: 17TH ANNUAL SURVEY OF THE 100 LARGEST U.S. PUBLIC COMPANIES 72 (2019), http://digital.shearman.com/i/1162884-2019-corporate-governance-executive-compensation-survey/0? \_ga=2.20327160.1104264344.1582561301-1742671351.1582561301 [https://perma.cc/NV5R-URDC].

diagnose the bias problem, as discussed above. Finally, this Part explores additional issues in law, beyond the use of corpus linguistics to interpret constitutions and statutes, where methods for measuring bias in text can be useful. We see a wide range of potential applications across many areas of legal practice and scholarship. Using word embeddings to identify hidden bias can operate as a general-purpose methodology.

#### A. MAKING HIDDEN BIAS TRANSPARENT

This Article demonstrates how one machine learning method can shed light on the existence of gender bias in a corpus. Gender bias is a serious problem in its own right, and in that respect, providing a tool for measuring gender bias in text is a contribution in itself. However, there are many more types of hidden bias to uncover. Developing tools for making other forms of bias transparent is an important priority for future research.

Tools for illuminating other forms of hidden bias already exist. For instance, the Garg et al. article, which forms the basis for the methods used in Part III, also demonstrates how we might use word embeddings to measure bias related to ethnicity.<sup>149</sup> In addition to gender bias, that article also studies the relationship between certain ethnicities in the United States—Caucasians, Asian-Americans, and Latinxs-and references to occupations and adjectives. The article identifies ethnicities using surnames and analyzes whether certain surnames are associated more closely with the occupational words and adjectives studied above in Part III.<sup>150</sup> The results provide evidence of ethnic stereotypes in the corpora that the authors study. For instance, the top five occupations most closely associated with Caucasians are smith, blacksmith, surveyor, sheriff, and weaver; for Asian-Americans, they are professor, official, secretary, conductor, and physicist; and for Latinxs, they are housekeeper, mason, artist, janitor, and dancer.<sup>151</sup> Recent research also analyzes word embeddings to identify patterns in the way that different socioeconomic classes are discussed in text.<sup>152</sup> This research finds that certain words are associated with different classes-such as bowling's association with the working class and golf's association with the rich-and thus provides a basis for measuring classism systematically in a corpus.<sup>153</sup>

In many respects, however, the tools we have available today are quite simple. The results presented here, for instance, are illuminating, but more importantly serve as a call for further academic investment in developing better methods for empirically identifying bias in text. Surely there are aspects of gender bias that the method applied here does not capture.<sup>154</sup>

<sup>149.</sup> See Garg et al., supra note 117, at E3639-41.

<sup>150.</sup> Id. at E3640.

<sup>151.</sup> Id. at E3638 tbl. 1.

<sup>152.</sup> Austin C. Kozlowski, Matt Taddy & James A. Evans, *The Geometry of Culture: Analyzing the Meanings of Class Through Word Embeddings*, 84 AM. SOC. REV. 905, 911–14 (2019).

<sup>153.</sup> Id. at 913 fig. 2.

<sup>154.</sup> For instance, gender bias is studied here in only two types of context—the use of occupational words and of adjectives—although bias may be manifest in other contexts. Methods might be developed

The process for developing those improved methods will require the interaction of three bodies of literature. The first is the multidisciplinary literature that articulates theories of bias, which provides us with a framework for understanding how discrimination acts in social life.<sup>155</sup> The second is the technical literature, such as the articles discussed immediately above, that constructs new techniques for quantitatively measuring bias.<sup>156</sup> The third is the work on bias specifically in the legal system, which provides the institutional detail needed to accurately apply social theory and empirical methods in the context of law.<sup>157</sup> We hope that this Article helps spark the interdisciplinary interest needed to further that dialogue.

#### B. DEBIASING LEGAL CORPUS LINGUISTICS

Revealing the latent bias in a corpus necessarily raises the question of what, if anything, judges and scholars can do about that bias. Though our ability to recommend potential ways to account for and mitigate bias is necessarily limited by the need for further exploration of the bias that exists in corpora used for legal interpretation, here we identify some avenues for additional exploration. First, scholars of legal interpretation and corpus linguistics must identify the kinds of interpretation questions that may be impermissibly affected by a hidden bias in the corpus. Second, scholars must consider potential measures to correct a corpus itself or modify the methodology used to perform a corpus linguistics analysis to neutralize bias that may infect decisionmaking. Third, scholars must consider whether an individual decisionmaker in isolation can effectively identify and mitigate bias embedded in corpus linguistics when bias cannot be corrected for with modified corpus construction or adjusted methodology. Below, we offer some observations on each of these remaining inquiries.

to study whether male and female subjects are depicted more or less as self-actuating agents, an analysis that may measure their relationship with certain verbs. Furthermore, the single-word analyses undertaken here could be expended to study relationships between males and females and larger phrases, adding greater nuance to the bias studied. We hope that future work will explore opportunities such as this, and we note that it may be potentially fruitful to expand the dialogue between critical theory and the research developing these empirical methods.

<sup>155.</sup> See, e.g., GORDON W. ALLPORT, THE NATURE OF PREJUDICE (1954); Peter Glick & Susan T. Fiske, *The Ambivalent Sexism Inventory: Differentiating Hostile and Benevolent Sexism*, 70 J. PERSONALITY & SOC. PSYCHOL. 491 (1996).

<sup>156.</sup> See, e.g., Bolukbasi et al., supra note 117; Bruno Lepri, Nuria Oliver, Emmanuel Letouzé, Alex Pentland & Patrick Vinck, Fair, Transparent, and Accountable Algorithmic Decision-Making Processes, 31 PHIL. & TECH. 611 (2018); Jieyu Zhao, Tianlu Wang, Mark Yatskar, Ryan Cotterell, Vicente Ordonez & Kai-Wei Chang, Gender Bias in Contextualized Word Embeddings, N. AM. CHAPTER ASS'N COMPUTATIONAL LINGUISTICS (2019); Jieyu Zhao, Tianlu Wang, Mark Yatskar, Vicente Ordonez, Kai-Wei Chang, Gender Bias in Coreference Resolution: Evaluation and Debiasing Methods, 2018 N. AM. CHAPTER ASS'N FOR COMPUTATIONAL LINGUISTICS 15; Zhao et al., supra note 117.

<sup>157.</sup> See, e.g., AT THE BOUNDARIES OF LAW: FEMINISM AND LEGAL THEORY (Martha Albertson Fineman & Nancy Sweet Thomadsen eds., 1991); CATHARINE A. MACKINNON, TOWARD A FEMINIST THEORY OF THE STATE (1989); Katherine M. Franke, *Theorizing Yes: An Essay on Feminism, Law, and Desire*, 101 COLUM. L. REV. 181 (2001); Leti Volpp, *Feminism Versus Multiculturalism*, 101 COLUM. L. REV. 1181 (2001).

# 1. Identifying Potential Instances of Bias in Corpus-Based Legal Interpretation

To better understand whether bias in a corpus can be sufficiently mitigated, scholars will first need to identify classes of legal interpretation questions that are affected by corpus bias in a way that delegitimizes the results of a corpus linguistics analysis. Not every legal interpretation question will fit into this class, and determining whether an interpretive issue raises a bias problem will be a difficult task. The effect of hidden bias will likely be indirect, and determining whether the effect of bias renders the results suspect will require additional work.

For some questions, bias in a corpus may have no meaningful effect on the corpus results. Whether *vehicle*, for example, refers to automobiles or bicycles—or both—may not be affected at all by the gender bias described in Part III above. For some questions, bias may be implicated in the results but not in a way that compromises the results. A decisionmaker might determine that, even if *bicycle* is more often associated with females and *car* is more often associated with males, or vice versa, *vehicle* in a statute can be properly understood to refer to cars and other motor vehicles and not to bicycles because *vehicle* most often refers to cars and other motor vehicles. Likewise, a regulation detailing the licensing requirements for nurses may not be inappropriately affected by the association between *nurse* and females.

But in other situations, bias inherent in a corpus may have a material effect on a question of legal interpretation. Could gender bias, for example, skew corpus linguistics results for the terms *professor* and *researcher*, two occupations with favored preference status in the Immigration and Nationality Act?<sup>158</sup> Bias may be especially prevalent and problematic when using historical corpora for an originalist interpretation of a statute or constitutional provision that remains in force today. For instance, how might a corpus linguistics analysis view the word *spouse* in the federal tax code provisions or state family laws enacted decades ago?<sup>159</sup> In that context, there is a high risk of corpus bias that is at odds with contemporary understandings of justice and equality but that escapes detection through available corpus linguistics tools.

# 2. Corpus Construction and Methodology

Having identified the ways in which hidden bias in a corpus may infect and delegitimize legal interpretation, scholars will be able to consider corrective measures. Ideally, the bias would be corrected ex ante, either through corpus

<sup>158.</sup> See Immigration and Nationality Act § 203, 8 U.S.C. § 1153(b)(1)(B) (2018) (providing visa allocation to "[o]utstanding professors and researchers" without requiring an employment offer or labor certification from the Department of Labor).

<sup>159.</sup> See, e.g., Internal Revenue Code of 1954, pt. II, § 6013, 68A Stat. 1, 733–36 (codified as amended at 26 U.S.C. § 6013 (2018)); An Act to Revise the Marriage and Dissolution of Marriage Laws of This State by Adopting the Provisions of the Uniform Marriage and Divorce Act as Recommended by the National Conference of Commissioners on Uniform State Law, ch. 536, § 12, 1975 Mont. Laws 1514, 1519 (codified as amended at MONT. CODE ANN. § 40-1-404 (2019)) (legislating a definition and rights of a "[p]utative spouse").

construction or by the use of additional, more sophisticated tools of corpus analysis. However, either of these approaches would introduce problems of its own.

It is unclear how one would construct a corpus free from bias, and any deliberate effort to select certain kinds of text over others would make the corpus unrepresentative in different ways. In addition, any attempt to construct a corpus primarily for the purpose of legal interpretation magnifies concerns that some corpus linguistics critics have already raised. As Part II discusses, some critics worry that reliance on corpora in legal interpretation places undue weight on the choices of a corpus creator rather than those of an elected or appointed decisionmaker.<sup>160</sup> This concern may become even more problematic where the corpus creator has gathered text knowing that the corpus would be used for legal interpretation.

#### 3. Judging and Mitigating Bias

In the absence of a suitable correction to corpora themselves or to corpus linguistics methodology, scholars must consider whether judges can correct for biased results themselves. Can a judge, having explored the frequency, context in usage, and collocation for a term identify bias in the results and account for it? Without exploring this possibility in detail, we note that this approach to bias correction potentially raises additional concerns of indeterminacy and lack of transparency by introducing another juncture at which a judge must make a decision. In the absence of a robust set of accepted norms by which to make that decision, a judge's effort to scrub bias from the results relies on the very intuition that corpus linguistics was meant to reduce. To the extent that a judge uses intuition to identify and eliminate bias, the judge risks either compounding the hidden biases that infect the results or introducing new biases. Corpus linguistics proponents and critics alike will agree that increasing the role of intuition in this way significantly reduces or negates any benefit that corpus linguistics offers.

An alternative to a judge's intuitive adjustment of the results of a corpus linguistics analysis to account for bias is the machine learning techniques used in this Article. As mentioned above, there is an active community of researchers exploring ways to debias the results of these techniques.<sup>161</sup> If effective, debiasing would allow a corpus to remain as is, warts and all, while problematic biases are accounted for through the algorithm itself.<sup>162</sup> Shifting our attention from corpus linguistics to machine learning tools could help a user both identify and control for particular biases while also measuring the relationships between words in a more sophisticated way than collocation and frequency can.

There are trade-offs, however, that arise with increasing the quantitative sophistication of empirical legal interpretation methods. The attraction (and limitation) of corpus linguistics in legal interpretation is that it can at least be

<sup>160.</sup> See supra Section II.B.2.

<sup>161.</sup> See supra note 156.

<sup>162.</sup> Of course, this proposition presumes that the research efforts to develop methods for debiasing algorithms are successful, and we do not suggest that this success may be easily assumed.

understood and used by generalist judges. Even so, we see examples of inexpert applications of the technique in opinions and scholarship.<sup>163</sup> That risk of misapplication may only increase with the introduction of machine learning techniques, which are often so complex as to render their operation opaque to laypeople and, at times, even to experts.<sup>164</sup> Or, much more likely, it will take the empirical analysis of public meaning entirely out of the hands of judges and give it to outside experts. Relatedly, allowing or requiring human judges to relinquish more of their decisionmaking to technology may prove problematic where more reasoning or a welfare criterion is necessary to supply legal content.<sup>165</sup>

The machine learning methods used here require some fluency in statistics, linguistics, and coding, and it is rather unrealistic for generalist judges to achieve such a skill set in addition to their substantive mastery of the law. Because judicial resources are scarce, the use of quantitatively sophisticated interpretive methods will most likely occur through expert advice. Outsourcing quantitative analysis to experts in this way has obvious advantages—the court can leverage the expertise of others. For that reason, the use of experts in litigation has a long history among certain dispute types, from personal injury and securities fraud litigation to shareholder appraisal actions.

However, access to legal services is catastrophically uneven in the United States, and the ability to retain expert quantitative linguists to opine on matters of legal interpretation will likely be available only to well-heeled litigants outside of the increasingly limited range of actions where claim aggregation is possible. Thus, applying more and more sophisticated corpus linguistics methods may improve judicial accuracy but, in the end, exacerbate distributional concerns. Legal interpretation may soon have its own version of the digital divide.<sup>166</sup>

#### C. NEW QUESTIONS THAT WE SHOULD BE ANSWERING

As important as it is to address the issue of hidden bias in empirical approaches to legal interpretation, it is also crucial to note how much value the methods for identifying bias introduced here can add to many other areas of the law. Legal corpus linguistics is just one area where quantitatively measuring gender bias matters. Future scholarship should explore these new avenues as a pressing matter of public policy.

<sup>163.</sup> Stefan Th. Gries, In Defense of Corpus-Linguistic Approaches to Ordinary Meaning: What Critics Say and a Cheatsheet for How to Preempt Them or Respond 1 (Feb. 7, 2020) (unpublished manuscript) (on file with the authors).

<sup>164.</sup> See, e.g., ELLIOTT ASH, SOC. MKT. FOUND., JUDGE, JURY, AND EXECUTE FILE: THE BRAVE NEW WORLD OF LEGAL AUTOMATION 5–6 (2018), http://www.smf.co.uk/wp-content/uploads/2018/06/CAGE-FINAL-VF.pdf [https://perma.cc/8ZZP-GF4Y].

<sup>165.</sup> Id. at 7. The authors thank the editors for raising this point.

<sup>166.</sup> See generally Donna L. Hoffman & Thomas P. Novak, The Growing Digital Divide: Implications for an Open Research Agenda (July 8, 1999) (unpublished manuscript) (available at https://www.markle.org/sites/default/files/digitaldivide\_openresearch.pdf [https://perma.cc/ML6K-R3AJ]) (describing the origins and meaning of the term *digital divide*).

For instance, the methods introduced here can be used to shed light on hidden bias in U.S. case law, which can be considered a corpus in its own right. Are there patterns of gender or other types of bias in certain types of cases? In certain time periods? In certain jurisdictions? In the decisions of particular judges? When particular types of litigants are involved?

These methods can also identify hidden bias in other bodies of text that are important to legal issues. For example, they could measure bias in public company disclosures at a time when laws for increasing the role of women on corporate boards are under debate.<sup>167</sup> These methods could illuminate hidden bias in federal or state regulations, particularly as they apply to workplaces. They could be used to identify bias in documentary evidence in sexual harassment cases. And so on.

Importantly, the quantitative methods for measuring bias can be an important tool in moving beyond purely descriptive studies, like this Article, to making causal inferences. Using word embeddings to measure bias gives us a finely grained tool for assessing the quantity and quality of bias in text. That level of detail can be useful as we move from identifying the problem to understanding its underlying causes.

#### CONCLUSION

The idea that words may be neutral on their face but gendered in their usage is unsurprising. Words reflect larger cultural phenomena, including pervasive gender stereotypes and biases that scholars have analyzed and deconstructed for years.<sup>168</sup> The machine learning methods that we use here allow us to measure and quantify gender bias in the Corpus of Historical American English (COHA), which collects text from two hundred years of American English. Our results are consistent with our expectations: the COHA, like the linguistic community it is designed to represent, is sexist. Ostensibly neutral occupational terms like *judge* and *housekeeper* hide gendered identities. Likewise, adjectives like *gentle* and *autocratic* lean female or male. Those biases change over time, from the start of

<sup>167.</sup> See, e.g., CAL. CORP. CODE § 301.3 (West, Westlaw through 2020 Reg. Sess.); Moez Bennouri, Tawhid Chtioui, Haithem Nagati & Mehdi Nekhili, Female Board Directorship and Firm Performance: What Really Matters?, 88 J. BANKING & FIN. 267 (2018); Johanne Grosvold, Where Are All the Women? Institutional Context and the Prevalence of Women on the Corporate Board of Directors, 50 BUS. & Soc'Y 531 (2011); Paul B. McGuinness, João Paulo Vieito & Mingzhu Wang, The Role of Board Gender and Foreign Ownership in the CSR Performance of Chinese Listed Firms, 42 J. CORP. FIN. 75 (2017); Vathunyoo Sila, Angelica Gonzalez & Jens Hagendorff, Women on Board: Does Boardroom Gender Diversity Affect Firm Risk?, 36 J. CORP. FIN. 26 (2016).

<sup>168.</sup> See generally Michela Carlana, Implicit Stereotypes: Evidence from Teachers' Gender Bias, 134 Q.J. ECON. 1163 (2019); Amanda B. Diekman, Alice H. Eagly & Patrick Kulesa, Accuracy and Bias in Stereotypes About the Social and Political Attitudes of Women and Men, 38 J. EXPERIMENTAL SOC. PSYCHOL. 268 (2002); Jack B. Harrison, "Because of Sex," 51 LOY. L.A. L. REV. 91 (2018); Madeline E. Heilman, Gender Stereotypes and Workplace Bias, 32 RES. ORGANIZATIONAL BEHAV. 113 (2012); Elise Meyer, Designing Women: The Definition of "Woman" in the Convention on the Elimination of All Forms of Discrimination Against Women, 16 CHI. J. INT'L L. 553 (2016); Alexander I. Platt, Debiasing Statutory Interpretation, 39 OHIO N.U. L. REV. 275 (2012); Michela Menegatti & Monica Rubini, Gender Bias and Sexism in Language, OXFORD RES. ENCYCLOPEDIAS: COMM. (Sept. 26, 2017), https://oxfordre.com/communication/view/10.1093/acrefore/9780190228613.001.0001/acrefore-9780190228613-e-470 [https://perma.cc/8Q84-LPDZ].

the COHA in 1810 to its end in 2009, diminishing from nineteenth-century highs but nevertheless persisting into modern times. Finally, gender bias in the COHA differs depending upon whether text is written by a male or female author.

The implications of structural gender bias—and, likely, other kinds of bias—in the COHA are significant. To the extent that judges rely on the COHA for legal interpretation, they risk importing and embedding gender bias in their decisions. In this Article, we explore some of the next steps for better understanding that risk and potentially mitigating it. We identify some of the issues that might arise in attempts to mitigate the effects of bias, including concerns about judicial competency and transparency. Our discussion is merely the beginning of what we hope will become a serious inquiry into the effects of corpus bias on corpus-assisted legal interpretation.

A lingering question about legal interpretation as a whole, however, lurks under the surface of our discussion: Is the concept of ordinary meaning fatally flawed? The rise of corpus linguistics in legal interpretation has lived up to some of its proponents' claims. The task of finding ordinary meaning is more transparent in that it more clearly reveals and illustrates the premise that words in a statute should be interpreted in light of the way that those words are or were used in the relevant community.<sup>169</sup> To understand the ordinary meaning of "bear [a]rms"<sup>170</sup> or

<sup>169.</sup> See Lee & Mouritsen, supra note 2, at 792-93 ("We speak of a search for meaning 'not in the subjective, multiple mind of Congress but in the understanding of the objectively reasonable person.' And we generally conclude that the search for such meaning . . . assures notice to the public, protects reliance interests, assures consistency of application, and respects the will of the legislative body." (quoting Frank H. Easterbrook, The Role of Original Intent in Statutory Construction, 11 HARV. J.L. & PUB. POL'Y 59, 65 (1988))); id. at 795 ("When we speak of ordinary meaning, we are asking an empirical question-about the sense of a word or phrase that is most likely implicated in a given linguistic context."); id. at 827 ("By limiting a search for ordinary meaning to the relevant speech community and register in question, we can have greater confidence that information about the frequency of use of a given word is telling us something useful about ordinary meaning."); see also Baude & Doerfler, supra note 27 ("Courts and scholars sometimes use the phrase 'plain meaning' to denote something like ordinary meaning-that is, the normal meaning, or the meaning one would normally attribute to those words given little information about their context."); Slocum, supra note 5, at 17 ("The very premise of the ordinary meaning doctrine (i.e. presumed legislative adherence to normal principles of language usage) is that the test for meaning is an objective one that is external to the legislature's actual intentions or the concerns of the court. While courts assume that the ordinary meaning of a statute's language represents the legislature's intent, the intent being referenced is generalized in the sense that it is not connected to any particular Congress, subject matter, or statute."); Thomas Lee & Stephen Mouritsen, Opinion, Judging Ordinary Meaning with Corpus Linguistics, WASH. POST (Aug. 8, 2017, 8:59 AM), https://www.washingtonpost.com/news/volokhconspiracy/wp/2017/08/08/judging-ordinary-meaning-with-corpus-linguistics ("We argue that a complete theory of ordinary meaning requires us to take into account not only the comparative frequency of different senses, but also the context of an utterance, its historical usage and the speech community in which it was uttered. Context necessarily includes the formal aspects of an utterance, its syntactic structure and semantic features, as well as the pragmatic aspects of the utterance, including the physical, spatial and social environment in which it occurs. Ordinary meaning should also take into account historical usage, acknowledging the simple fact that language is in a constant state of change (but does not change at a predictable rate). Ordinary meaning should also take into account variations in meaning in the speech or writing of different speech communities and different linguistic registers.").

<sup>170.</sup> U.S. CONST. amend. II. See generally Dennis Baron, Corpus Evidence Illuminates the Meaning of Bear Arms, 46 HASTINGS CONST. L.Q. 509 (2019); Neal Goldfarb, A (Mostly Corpus-Based) Linguistic Reexamination of D.C. v. Heller and the Second Amendment (Nov. 5, 2019) (unpublished

"establishment"<sup>171</sup> or "[e]moluments,"<sup>172</sup> one might look at the usage of those terms in early American historical communities. But what if that community was sexist? Or racist? Or biased against particular religions? Does ordinary meaning, or at least an originalist understanding of the term, conserve biases that our society—and the law—now repudiate? What is left of ordinary meaning and originalism after bias is somehow scrubbed from inquiry? What other kinds of bias, we wonder, are petrified in language and embedded into the law by judges' search for ordinary meaning?

manuscript) (available at https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3481474 [https://perma. cc/7XMP-JPTV]).

<sup>171.</sup> U.S. CONST. amend. I. See generally Barclay et al., supra note 104.

<sup>172.</sup> U.S. Const. art. I, § 6, cl. 2; see also id. art. I, § 9, cl. 8; id. art. II, § 1, cl. 7. See generally Clark D. Cunningham & Jesse Egbert, Using Empirical Data to Investigate the Original Meaning of "Emolument" in the Constitution, 36 GA. ST. U. L. REV. 465 (2020); Phillips & White, supra note 67; Clark D. Cunningham & Jesse Egbert, Scientific Methods for Analyzing Original Meaning: Corpus Linguistics and the Emoluments Clauses (Ga. State Univ. Coll. of Law Legal Studies, Working Paper No. 2019-02, 2019), https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3321438 [https://perma.cc/XA24-TS93].