

DUELING DICTIONARIES AND CLASHING CORPORA

KEVIN TOBIA[†]

INTRODUCTION

Textualism has broad support—at the Supreme Court,¹ within the lower federal courts’ new cohort of young “Trump judges,”² within many state courts,³ and even within the legal academy.⁴ Textualism comes in several variations,⁵ and a new “populist” version is taking hold.⁶ Modern textualists claim to interpret law from the perspective of *an ordinary person*,⁷ which includes giving terms in law their

Copyright © 2022 Kevin Tobia.

[†] Associate Professor of Law, Georgetown University Law Center. This essay originated from a conference on Corpus Linguistics and the Second Amendment, at the Duke Center for Firearms Law. Thanks to Joseph Blocher, Jacob Charles, and Darrell Miller for the invitation and to the co-panelists and participants for their comments, especially Dennis Baron, William Baude, Anya Bernstein, and Stephen Mouritsen. Great thanks to John Macy and the *Duke Law Journal* for outstanding editorial assistance.

1. See Victoria Nourse & William N. Eskridge, *Textual Gerrymandering: The Eclipse of Republican Government in an Era of Statutory Populism*, 96 N.Y.U. L. REV. 1718, 1722 (2021) (“Should interpreters focus on the readers and consumers of statutes (We the People) or the authors and producers of statutes (Congress)? . . . On its face, the now-dominant Supreme Court approach elevates the *consumer* perspective and belittles or ignores that of the *producers*. This is an alarming development.”); Kevin Tobia, Brian Slocum, & Victoria Nourse, *Statutory Interpretation from the Outside*, 122 COLUM. L. REV. 213, 216 (2022) (“[O]rdinary meaning is regularly deployed by all members of the current Supreme Court.”).

2. Jason Zengerle, *How the Trump Administration Is Remaking the Courts*, N.Y. TIMES MAG. (Aug. 22, 2018), <https://www.nytimes.com/2018/08/22/magazine/trump-remaking-courts-judiciary.html> [<https://perma.cc/UG99-J2QZ>] (President Trump was committed to “nominating and appointing judges that are committed originalists and textualists.”).

3. Abbe R. Gluck, *The States as Laboratories of Statutory Interpretation: Methodological Consensus and the New Modified Textualism*, 119 YALE L.J. 1750, 1758 (2010) (“[I]n the states studied, textualism is more than merely alive and well; it is the *controlling* interpretive approach—the consensus methodology chosen by the courts.”).

4. Eric Martínez & Kevin Tobia, *The Legal Academy and Theory Survey*, (unpublished manuscript) (on file with author).

5. See, e.g., Gluck, *supra* note 3; Tara Leigh Grove, *Which Textualism?*, 134 HARV. L. REV. 265, 265 (2020) (comparing “formalistic” and “flexible” forms of textualism).

6. See Nourse & Eskridge, *supra* note 1, at 1723; see also generally Anya Bernstein & Glen Staszewski, *Judicial Populism*, 106 MINN. L. REV. 283 (2021) (commenting on judicial populism).

7. Anya Bernstein, *Democratizing Interpretation*, 60 WM. & MARY L. REV. 435, 440 (2018) (“Textualism instructs judges to interpret a statute as its addressees would understand it.”); Amy

ordinary meanings.⁸ This commitment is taken to promote rule of law values (e.g. publicity), fair notice, and a “democratic” mode of interpretation.⁹ Even non-textualist Justices have begun to appeal to the “ordinary speaker.”¹⁰

How do today’s textualists go about finding ordinary meaning? They regularly appeal to sources including “dictionaries, corpus linguistics, and canons of construction.”¹¹ The flexibility of dictionaries and canons of construction is well-documented.¹² Judges can cherry-pick helpful dictionary definitions,¹³ and for many canons of

Coney Barrett, *Congressional Insiders and Outsiders*, 84 U. CHI. L. REV. 2193, 2195 (2017) (“[Textualists] view themselves as agents of the people rather than of Congress and as faithful to the law rather than to the lawgiver”).

8. E.g., Amy Coney Barrett, *Assorted Canards of Contemporary Legal Analysis: Redux*, 70 CASE W. RES. L. REV. 855, 856 (2020) (noting the significance of “ordinary meaning”).

9. See Bernstein, *supra* note 7, at 442; see also Kevin Tobia, Brian Slocum & Victoria Nourse, *Progressive Textualism*, 110 GEO. L.J. (forthcoming 2022) (documenting modern textualism’s motivations).

10. Consider Justice Roberts’s recent question in *Facebook v. Duguid*’s oral argument:

[O]ur objective is to settle upon the most natural meaning of the statutory language to an ordinary speaker of English, right? . . . So the most probably useful way of settling all these questions would be to take a poll of 100 ordinary – ordinary speakers of English and ask them what [the statute] means, right?

Transcript of Oral Argument at 51–52, *Facebook, Inc. v. Duguid*, 141 S. Ct. 1163 (2021) (No. 19-511). Justice Alito, in his concurring opinion in *Duguid*, noted that

[t]he strength and validity of an interpretive canon is an empirical question, and perhaps someday it will be possible to evaluate these canons by conducting what is called a corpus linguistics analysis, that is, an analysis of how particular combinations of words are used in a vast database of English prose.

Facebook, Inc. v. Duguid, 141 S. Ct. at 1174 (Alito, J., concurring).

11. Nourse & Eskridge, *supra* note 1, at 1727.

12. On canons, see Karl Llewellyn, *Remarks on the Theory of Appellate Decision and the Rules or Canons about How Statutes Are to Be Construed*, 3 VAND. L. REV. 395, 401 (1950) (“[T]here are two opposing canons on almost every point.”); see also generally Anita Krishnakumar, *Dueling Canons*, 65 DUKE L.J. 909 (2016); Anita Krishnakumar & Victoria Nourse, *The Canon Wars*, 97 TEX. L. REV. 163 (2018); Ryan Doerfler, *Late-Stage Textualism*, 2022 SUP. CT. REV. (forthcoming 2022). On dictionaries, see generally Samuel A. Thumma & Jeffrey L. Kirschmeier, *The Lexicon Has Become a Fortress: The United States Supreme Court’s Use of Dictionaries*, 47 BUFF. L. REV. 227 (1999); Stephen C. Mouritsen, *The Dictionary is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning*, 2010 B.Y.U. L. REV. 1915 (2010); Ellen P. Aprill, *The Law of the Word: Dictionary Shopping in the Supreme Court*, 30 ARIZ. ST. L.J. 227 (1998); James J. Brudney & Lawrence Baum, *Oasis or Mirage: The Supreme Court’s Thirst for Dictionaries in the Rehnquist and Roberts Eras*, 55 WM. & MARY L. REV. 483 (2013).

13. Aprill, *supra* note 12, at 300 (“[O]pinions often cite or rely on only one definition in only one dictionary For the most part, opinions fail to explain or justify the basis for their choice.”); Brudney & Baum, *supra* note 12, at 491 (arguing that the Supreme Court has a “tendency to cherry-pick definitions that support results reached on other grounds”); Kevin Tobia, Brian Slocum & Victoria Nourse, *Ordinary Meaning and Ordinary People*, 171 U. PA. L.

interpretation, there is an opposing canon that could support the opposite result.

This essay explores textualism's newest tool: corpus linguistics. Over the past five years, the tool has been increasingly employed by U.S. courts.¹⁴ Legal corpus linguistics has also caught the attention of the U.S. Supreme Court. Justice Thomas mentioned corpus linguistics in his 2018 *Carpenter* dissent, and Justice Alito noted it again in his 2021 *Duguid* concurrence.¹⁵ Most recently, Justices Roberts and Barrett discussed corpus linguistics in a 2022 oral argument.¹⁶

Broadly speaking, legal corpus linguistics treats collections of texts (“corpora”) as data.¹⁷ To learn about the ordinary meaning of a statutory or constitutional term (e.g. “commerce”), a textualist would evaluate how that term is commonly used in different written sources (e.g. books) and what other words tend to appear near it in those sources. For example, the interpreter might consider different senses of a term (e.g. “commerce” in the narrow sense of “the trading . . . and selling of goods,” versus “commerce” in the broader sense of “all forms of social and economic intercourse”).¹⁸ Next, the interpreter could evaluate how often each of those senses appears in the corpus. Perhaps, for example, a scholar may find that the narrower trade sense appears more frequently than the broader sense. Some scholars suggest that these data evince the constitutional or statutory meaning of the term. For example, a recent article suggests that corpus linguistics data about “commerce” “at least arguably, tells us that the original [constitutional] meaning of commerce is the trade sense of the term.”¹⁹

The essay argues that corpus linguistics—an important and useful method in linguistics—is unlikely to achieve textualist’s theoretical aims. Many have criticized legal corpus linguistics’ prospects,

REV. (forthcoming 2023) (documenting the Supreme Court’s citation of dozens of ordinary and legal dictionaries).

14. Kevin Tobia, *The Corpus and the Courts*, U. CHI. L. REV. ONLINE (2021) (documenting judge’s appeals to corpus linguistics, rising sharply over the past five years).

15. *Carpenter v. United States*, 138 S. Ct. 2206, 2238 n.4 (2018) (Thomas, J., dissenting); *Facebook, Inc. v. Duguid*, 141 S. Ct. 1163, 1175 (2021) (Alito, J., concurring).

16. Transcript of Oral Argument at 9–11, *ZF Automotive U.S., Inc. v. Luxshare, Ltd.* (2022) (No. 21-401).

17. Thomas R. Lee & Stephen C. Mouritsen, *Judging Ordinary Meaning*, 127 YALE L.J. 788, 795 (2018) (“Corpus linguists study language through data derived from large bodies—*corpora*—of naturally occurring language.”).

18. Thomas R. Lee & James C. Phillips, *Data-Driven Originalism*, 167 U. PA. L. REV. 261, 300 (2019).

19. *Id.* at 323.

cautioning that judges do not have the training or expertise to employ the tool or pointing to fundamental flaws of the current method, as applied to legal debates.²⁰ This essay starts from a different perspective, noting that judges are *already* employing and referencing corpus linguistics in legal interpretation,²¹ as scholars advance provocative corpus linguistics arguments about statutory and constitutional language.²²

Corpus linguistics has been offered as a *preferred* interpretive tool, avoiding the pitfalls of dueling canons or cherry-picked dictionary definitions.²³ However, this essay proposes, there is one important commonality among textualists' current use of dictionaries, canons, and legal corpus linguistics: *Flexibility*. The essay articulates ten emerging “arguments” and “counterarguments” of legal corpus linguistics. Alongside the pitfalls of dueling canons and dueling dictionaries, legal interpreters should be aware of the similar possibility of “clashing corpora.” Corpus linguistics can greatly enrich our understanding of language and cognition, but—at least in the form employed by textualist judges and commentators—it does not provide inexorable determinations of how ordinary people understand legal language in contested cases of legal interpretation.

20. See, e.g., Bernstein, *supra* note 7; Anya Bernstein, *What Counts as Data?*, 86 BROOK. L. REV. 435 (2021); Anya Bernstein, *Legal Corpus Linguistics and the Half-Empirical Attitude*, 106 CORNELL L. REV. 1397 (2021); John S. Ehrett, *Against Corpus Linguistics*, 108 GEO. L.J. ONLINE 50 (2019); Ethan J. Herenstein, *The Faulty Frequency Hypothesis: Difficulties in Operationalizing Ordinary Meaning Through Corpus Linguistics*, 70 STAN. L. REV. ONLINE 112 (2017); Donald L. Drakeman, *Is Corpus Linguistics Better than Flipping a Coin?*, 109 GEO. L.J. ONLINE 81 (2020); Stanley Fish, *The Interpretive Poverty of Data*, BALKINIZATION (Mar. 2, 2018) <https://balkin.blogspot.com/2018/03/the-interpretive-poverty-of-data.html> [<https://perma.cc/4X4S-7QZ8>]; Carissa Byrne Hessick, *Corpus Linguistics and the Criminal Law*, 2017 B.Y.U. L. REV. 1503 (2018); Brian G. Slocum & Stefan Th. Gries, *Judging Corpus Linguistics*, 94 S. CAL. L. REV. POSTSCRIPT 13 (2020); Kevin Tobia, *Testing Ordinary Meaning*, 134 HARV. L. REV. 726 (2020); Evan C. Zoldan, *Corpus Linguistics and the Dream of Objectivity*, 50 SETON HALL L. REV. 401 (2019). But see Thomas R. Lee & Stephen C. Mouritsen, *The Corpus and the Critics*, 88 U. CHI. L. REV. 275 (2021) (defending legal corpus linguistics).

21. Tobia, *supra* note 14 (documenting judge's appeals to corpus linguistics, rising sharply over the past five years).

22. E.g., Lee & Phillips, *supra* note 18, at 300–11 (providing a corpus linguistic analysis of “commerce”).

23. E.g., Lee & Mouritsen, *supra* note 17, at 877 (suggesting that corpus linguistics offers better evidence of ordinary meaning than dictionaries).

I. POPULAR TEXTUALISM

Today's textualism is *popular* in two different senses. First, it has significant support from judges and scholars.²⁴ (That said, it is not *universally* approved; modern critics lambast textualist practice as flawed²⁵ and even “bogus.”²⁶). Today's textualism is also “popular” in a second sense: its interpretive inquiry is focused on the public.²⁷ As Justice Barrett puts it, modern textualists “view themselves as agents of the people rather than of Congress and as faithful to the law rather than to the lawgiver.”²⁸ Textualists thus “approach language from the perspective of an ordinary English speaker.”²⁹ Judges increasingly adopt this popular stance, committing to interpret statutory and constitutional language *empirically*, in line with its “ordinary” or “public” meaning.³⁰

How does one find “ordinary public meaning”?³¹ Textualists appeal to linguistic evidence, like dictionary definitions.³² Commentators question that approach. Differing definitions allow judges to go “dictionary-shopping,”³³ and empirical studies suggest that judges’ dictionary use is often “ad hoc and subjective.”³⁴

“Picking and choosing” is a broader issue for textualists. Traditionally, textualism has aimed to constrain legal interpretation

24. See *supra* notes 1–4 and accompanying text.

25. E.g., Victoria Nourse, *Textualism 3.0: Statutory Interpretation After Justice Scalia*, 70 ALA. L. REV. 667 (2019).

26. Mitchell N. Berman & Guha Krishnamurthi, *Bostock was Bogus: Textualism, Pluralism, and Title VII*, 97 NOTRE DAME L. REV. 67 (2021).

27. Bernstein & Staszewski, *supra* note 6, at 287 (“[T]he brand of populism we address here . . . makes claims justifying action in the name of ‘the people.’”).

28. Barrett, *supra* note 7, at 2195.

29. *Id.* at 2194.

30. E.g., Lawrence B. Solum, *Triangulating Public Meaning: Corpus Linguistics, Immersion, and the Constitutional Record*, 2017 B.Y.U. L. REV. 1621 (2018).

31. This phrase, appearing in *Bostock v. Clayton Cnty.*, 140 S. Ct. 1731, 1738 (2020), reflects the synthesis of textualist’s ordinary meaning and originalist’s public meaning. As Victoria Nourse documents, “new” textualists are statutory originalists. Nourse, *supra* note 25, at 669.

32. Thumma & Kirschmeier, *supra* note 12, at 260–62 (documenting dictionary usage by Justices of the U.S. Supreme Court); Mouritsen, *supra* note 12, at 1918 (noting the “overarching trend to rely upon dictionaries to resolve lexical ambiguity”).

33. Aprill, *supra* note 12, at 318 (arguing that Justice Scalia sometimes treats dictionary definitions as authoritative, but other times rejects dictionary definitions).

34. Brudney & Baum, *supra* note 12, at 483 (“[T]he Court’s patterns of dictionary usage reflect a casual form of opportunistic conduct.”).

and limit judicial discretion.³⁵ For example, textualists offer the limitation of judicial discretion as one reason for courts to avoid evaluating legislative intent.³⁶ With respect to legislative history, many textualists seek to avoid judge's freedom to pick and choose their friends among the crowd. But recent critics note that textualist themselves can pick and choose what text to analyze,³⁷ pick and choose dictionaries,³⁸ and pick and choose definitions.³⁹ Textualists also choose among canons, "hundreds of interpretive presumptions that have no hierarchy among them."⁴⁰ The dueling canons and "dueling dictionaries" show that a mere commitment to "text" does not guarantee limited judicial discretion.

II. LEGAL CORPUS LINGUISTICS

Enter "legal corpus linguistics," an exciting new tool for textualists and other theorists committed to ordinary meaning.⁴¹ Corpora (the plural of "corpus") are samples of language-usage.⁴² To learn about the ordinary or public meaning of a term like "commerce" or "bear arms," interpreters might look beyond just a few dictionaries, to hundreds of uses of the phrase in a corpus. For example, a search of the Corpus of Founding Era American English revealed 281 instances of the phrase "bear arms." "[O]nly a handful don't refer to war, soldiering or organized armed action," suggesting to some that "the natural meaning of 'bear arms' in the framers' day was military."⁴³ With an appeal to

35. John F. Manning, *What Divides Textualists from Purposivists?*, 106 COLUM. L. REV. 70, 74–75 (2006).

36. See John F. Manning, *Second-Generation Textualism*, 98 CAL. L. REV. 1287, 1289 (2010).

37. Victoria Nourse, *Picking and Choosing Text: Lessons for Statutory Interpretation from the Philosophy of Language*, 69 FLA. L. REV. 1409, 1423–29 (2017); Nourse & Eskridge, *supra* note 1, at 1747–51.

38. Brudney & Baum, *supra* note 12, at 529–31.

39. Brudney & Baum, *supra* note 12, at 529–31.

40. Abbe R. Gluck, *Imperfect Statutes, Imperfect Courts: Understanding Congress's Plan in the Era of Unorthodox Lawmaking*, 129 HARV. L. REV. 62, 62 (2015); see also Llewellyn, *supra* note 12, at 401 ("[T]here are two opposing canons on almost every point.").

41. Lee & Mouritsen, *supra* note 17, at 795.

42. Jesse Egbert, *The Corpus—A Sample By Another Name*, LINGUISTICS WITH A CORPUS (May 27, 2021), <https://linguisticswithacorpora.wordpress.com/2021/05/27/the-corpus-a-sample-by-another-name/> [<https://perma.cc/W3ND-YKQK>].

43. Dennis Baron, *Opinion: Antonin Scalia Was Wrong About the Meaning of "Bear Arms."* WASH. POST. (May 21, 2018), https://www.washingtonpost.com/opinions/antonin-scalia-was-wrong-about-the-meaning-of-bear-arms/2018/05/21/9243ac66-5d11-11e8-b2b8-08a538d9dbd6_story.html [<https://perma.cc/E2FA-QMFW>]; see also Dennis Baron, *Corpus Evidence Illuminates the Meaning of Bear Arms*, 46 HASTINGS CONST. L.Q. 509, 510 (2019);

“big data,” legal corpus linguistics offers a new, and perhaps more objective and empirically robust, basis to seek “ordinary meaning.”

Despite its promise, legal corpus linguistics approach has faced criticism.⁴⁴ Yet, over the past decade, scholars and judges have adopted corpus linguistic tools to address questions about public meaning. This trend is sharp in the past five years, with citation to corpus linguistics from several state and federal courts.⁴⁵ Scholars have advanced new corpus linguistics arguments about constitutional language including “commerce,”⁴⁶ and in 2018, Justice Thomas cited corpus linguistics evidence about the meaning of “search” at the U.S. Supreme Court.⁴⁷

As legal corpus linguistics gains prominence, one of its primary concerns should be its political (non-)neutrality and potential for abuse. Some critics accuse textualist theory to be more motivated by conservatism than fidelity to democracy or separation of powers,⁴⁸ suggesting textualism is a mere “smokescreen by conservative judges to reach ideologically acceptable outcomes.”⁴⁹ A similar concern applies to textualist tools, including legal corpus linguistics. Thus far, legal corpus linguistics has been discussed more frequently (and favorably) by Republican-appointed judges than by Democratic-appointed ones.⁵⁰

But some legal corpus linguistics research appears to have the opposite valence. For example, some of the most robust legal corpus

Alison L. LaCroix, *Historical Semantics and the Meaning of the Second Amendment*, PANORAMA (Aug. 3, 2018), <http://thepanorama.shear.org/2018/08/03/historical-semantics-and-the-meaning-of-the-second-amendment/> [https://perma.cc/5WKC-S4AY]; Josh Jones, Note, *The “Weaponization” of Corpus Linguistics: Testing Heller’s Linguistic Claims*, 34 B.Y.U. J. PUB. L. 135, 135 (2020). But see James C. Phillips & Josh Blackman, *Corpus Linguistics and Heller*, 56 WAKE FOREST L. REV. 609 (2021).

44. See *supra* note 20.

45. Tobia, *supra* note 14.

46. Lee & Phillips, *supra* note 18, at 300–11.

47. *Carpenter v. United States*, 138 S. Ct. 2206, 2238 n.4 (2018) (Thomas, J., dissenting).

48. See, e.g., Andrei Marmor, *The Immorality of Textualism*, 38 LOY. L.A. L. REV. 2063, 2065 (2005) (“I believe that the underlying motivation of textualism derives from a neoconservative conception of the regulatory state, much more so, anyway, than from a concern with principles of democracy and separation of powers.”).

49. Grove, *supra* note 5, at 266 (Grove does not endorse this idea, but cites others who do, including Neil H. Buchanan & Michael C. Dorf, *A Tale of Two Formalisms: How Law and Economics Mirrors Originalism and Textualism*, 106 CORNELL L. REV. 591, 640 (2020) (suggesting that textualism is “a rhetorical smokescreen for extremely conservative results”)); William N. Eskridge, Jr. & Philip P. Frickey, *The Supreme Court, 1993 Term — Foreword: Law as Equilibrium*, 108 HARV. L. REV. 26, 77 (1994); Margaret H. Lemos, *The Politics of Statutory Interpretation*, 89 NOTRE DAME L. REV. 849, 851 (2013)).

50. Tobia, *supra* note 14.

linguistics research, from a number of scholars, questions the conclusions of *Heller*, finding that, “the Supreme Court’s reasoning may be flawed.”⁵¹ The Second Amendment is an alluring test case: Can legal corpus linguistics attain textualism’s promise of objectivity, and will commentators and judges persuaded by corpus linguistics evidence concerning “commerce” and “search” be similarly moved by evidence about “bear arms”?

As an example, consider a recent case. In *Jones v. Becerra* (concerning a Second Amendment challenge to California’s ban on firearm purchases by those between age 18 and 21), the Ninth Circuit ordered supplemental briefing concerning legal corpus linguistics and the Second Amendment. Specifically, the parties were instructed to address the “original public meaning” of the phrases: “A well regulated Militia”; “the right of the people”; and “shall not be infringed”—and to address “[h]ow does the tool of corpus linguistics help inform the determination of the original public meaning of those Second Amendment phrases?”⁵²

The parties’ responses were striking. Both the plaintiff-appellants and defendant-appellees criticized the method of corpus linguistics.⁵³ At the same time, both the plaintiff-appellants and defendant-appellees conducted corpus linguistic analyses and managed to find corpus linguistics data to support opposing conclusions about the original public meaning of the Second Amendment.⁵⁴

51. Baron, *supra* note 43; *see also* Baron, *supra* note 43; LaCroix, *supra* note 43; Jones, *supra* note 43; Neal Goldfarb, *A (Mostly Corpus-Based) Linguistic Reexamination of D.C. v. Heller and the Second Amendment* (unpublished manuscript), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3481474 [<https://perma.cc/8E4E-3SE3>]. *But see* Phillips & Blackman, *supra* note 43.

52. *Jones v. Becerra*, Order, Case 20-56174, at 1 (9th Cir. Mar. 26, 2021).

53. Supplemental Brief for Appellees at 2, *Jones v. Bonta*, 2022 WL 1485187 (9th Cir. 2022) (No. 20-56174), 2021 WL 1727661, at *2 (“[I]nitial results suggest that a corpus linguistics analysis would likely be of limited utility in answering [the] question.”); Supplemental Brief for Appellants at 2–3, *Jones v. Bonta*, 2022 WL 1485187 (9th Cir. 2022) (No. 20-56174), 2021 WL 1727665, at *2–*3 (“Because of the weaknesses inherent in the methodology of corpus linguistics, however, it ultimately sheds little light on the matter—and it certainly can do nothing to upset the interpretation of the Second Amendment adopted by binding Supreme Court precedent.”).

54. Supplemental Brief for Appellees, *Jones v. Bonta*, 2022 WL 1485187 (9th Cir. 2022) (No. 20-56174), 2021 WL 1727661, at *25–*26. In their supplemental brief, the Appellees noted that preliminary searches in COHA and COFEA for the phrase ‘right of the people’ return a relatively manageable number of hits: approximately 200 in each database. They do not appear to provide clear evidence that this phrase, as used in the Second Amendment, was originally understood to protect an individual right for persons under 21 to keep or bear arms (much less to purchase or receive them from a commercial dealer), however.

This result is not surprising, but *Jones v. Becerra* portends a new time of “clashing corpora.” Like judicial use of dictionaries, judicial use of corpus linguistics admits of interpretive choice and flexibility. Judges and advocates have flexibility in terms of which selection from the legal text to analyze, which corpus or corpora to search, which search(es) to conduct, and what conclusions to draw from the results returned from the corpus.

The phenomenon of “dueling dictionaries” is well-known. But this essay concludes by sketching some of the emerging “moves” of legal corpus linguistic argumentation (inspired by the style of Llewellyn’s dueling canons).⁵⁵

Argument	But Counterargument
1. The corpus data supports that the term ordinarily reflects <i>this</i> meaning; so <i>this</i> is its public meaning. ⁵⁶	1. The term is a legal term of art and should be given its legal meaning, <i>that</i> meaning. ⁵⁷
2. The corpus reveals that the term is always used in <i>this</i> sense; <i>this</i> is its public meaning.	2. A corpus is not exhaustive of ordinary understanding; the meaning might not be <i>this</i> sense.

Id.; Supplemental Brief for Appellants, *Jones v. Bonta*, 2022 WL 1485187 (9th Cir. 2022) (No. 20-56174), 2021 WL 1727665, at *2 (“We have conducted a corpus-linguistics analysis of the three phrases identified by the Court, and we set forth the results below—results that are fully consistent with the conventional evidence of the original public meaning of those phrases (and with the determinations in *Heller*).”).

55. Llewellyn, *supra* note 12, at 401–06. Here, thrust and parry 1 is nearly identical to Llewellyn’s pair concerning ordinary versus legal meaning. Here, the thrust and parry arguments imply conflicting, although not necessarily opposite, conclusions.

56. See *Vermont v. Misch*, 256 A.3d 519, 530 (Vt. 2021) (“Analyzing these databases . . . several studies have reviewed hundreds of instances of ‘bear arms’ and found that the phrase was overwhelmingly used in a collective or military sense.”).

57. This counterargument could be offered on the basis of precedent or common law, but could also be supported with corpus linguistics evidence. For a compelling example, see Lawrence Solan & Tammy Gales, *Revisiting a Classic Problem in Statutory Interpretation: Is a Minister a Laborer?*, 36 GA. ST. L. REV. 491, 505–513 (2020) (stating that “[t]he term ‘labor or service’ may not be a matter of ordinary meaning at all but may rather be a legal term of art” and examining a corpus of statutory language).

3. The corpus reveals that the term was never used in *that* sense; *that* cannot possibly be its meaning.⁵⁸

4. The corpus reveals that, generally, the term is (most) frequently used in *this* sense; *this* is its meaning.⁵⁹

5. The corpus reveals that, in the relevant context, the term is (most) frequently used in *this* sense; *this* is its meaning.⁶¹

6. The corpus shows that *this* is at least a possible sense of the term, a candidate for its ordinary meaning.⁶³

3. See Counterargument 2. Absent evidence is not evidence of absence.

4. Given the full context of the legal text, the term takes *that* sense.⁶⁰

5. The “context” shared by the examples of language-use in the corpus is not adequately similar to that of the statutory context.⁶²

6. Some language-use is figurative, metaphorical, sarcastic, or otherwise inapt as evidence of public meaning; *this* is not be a possible meaning in the legal text.⁶⁴

58. E.g., *Carpenter*, 138 S. Ct. at 2238 (Thomas, J., dissenting) (“At the founding, ‘search’ did not mean a violative of someone’s reasonable expectation of privacy The phrase ‘expectation(s) of privacy’ does not appear in . . . collections of early American English texts.”).

59. E.g., Lee & Phillips, *supra* note 18, at 300–11 (illustrating the concept using “commerce”).

60. A classic example is Justice Scalia’s opinion in *Smith v. United States*, 508 U.S. 223, 241–46 (1993) (arguing that offering a firearm in exchange for cocaine does not fit within the statutory language of “using” a firearm, since the broader context of “using a firearm” expresses “using a firearm as a weapon,” not any possible “use,” broadly construed).

61. E.g., *United States v. Costello*, 666 F.3d 1040, 1044 (7th Cir. 2012) (using Google News to assess how “harbor” is used with a human object, concluding that it most often implies hiding the human).

62. E.g., Phillips & Blackman, *supra* note 43, at 672 (acknowledging that one possible response to their corpus analyses is that the relevant phrase might have a different meaning in different contexts); see also *id.* at 680 (calling for analysis of words and phrases in only the “appropriate context”).

63. E.g., Lee & Mouritsen, *supra* note 17, at 828–29.

64. See generally Raymond Gibbs & Herbert Colston, *Figurative Language*, in HANDBOOK OF PSYCHOLINGUISTICS 835 (Matthew Traxler & Morton Gernsbacher ed., 2006).

7. The corpus shows that a term often appears with “*this*” and rarely with “*that*”; thus, *this* is more informative than *that* of the term’s public meaning.⁶⁵

8. The corpus provides evidence about the meaning of multi-word expressions by providing evidence about the meaning of each individual word.

9. The corpus provides evidence about the meaning of sentences by providing evidence about the meaning of each word and expression in that sentence.

10. Corpus evidence about “*this*” is not evidence of *public* meaning, where the corpus over-represents elite writers, and thus elite meaning.⁶⁹

7. Co-location frequency of “*this*” over “*that*” does not always imply that *this* is more central to the term’s meaning; in fact, it could imply the opposite.⁶⁶

8. Meanings of expressions are not always the simple sum of their parts.⁶⁷

9. Meanings of sentences are not always the simple sum of their parts.⁶⁸

10. Without good reason to think elite writings diverge relevantly from non-elite ones with respect to “*this*,” corpus evidence from the former provides evidence

65. E.g., Lee & Mouritsen, *supra* note 17, at 839 (describing common collocates as informative of a term’s ordinary meaning).

66. *Language Bias and Black Sheep*, NAT. LANGUAGE PROCESSING BLOG (June 24, 2016), <https://nlpers.blogspot.com/2016/06/language-bias-and-black-sheep.html> [<https://perma.cc/T7C2-TFMB>] (noting that, often in writing, “black” appears more frequently than “white” before “sheep”).

67. Solan & Gales, *supra* note 57, at 505–13 (considering the meaning of “labor or service”); *Smith*, 508 U.S. at 241–46 (considering the meaning of “uses a firearm”).

68. See generally Peter Hagoort & Jos van Berkum, *Beyond the Sentence Given*, 362 PHIL. TRANSACTIONS ROYAL SOC. B 801 (2007) (presenting evidence against a simple two-step compositional model of sentence representation); see also generally Nourse & Eskridge, *supra* note 1 (arguing that textualists inappropriately strip statutory language out of its statutory context and define individual terms (in a different context)).

69. Anya Bernstein, *More Than Words*, DUKE CTR. FOR FIREARMS L. BLOG (July 7, 2021), <https://firearmslaw.duke.edu/2021/07/more-than-words/> [<https://perma.cc/9JHU-FEXE>].

about public meaning of
“this.”⁷⁰

To enumerate these clashing arguments is not to endorse or discredit any particular one. It is simply to question the legal corpus linguistics (and textualist) claim that the introduction of these new empirical methods will straightforwardly constrain legal interpreters or provide uncontroversial answers to hard interpretive questions.

To be sure, where textualists aim to uncover how ordinary people understand language, legal corpus linguistics seems no less promising than dictionaries, canons, intuition, and other textualist tools. But where textualists can freely leverage any of these arguments and counterarguments (and pick and choose which text to analyze and searches to run⁷¹), legal corpus linguistics is unlikely to provide much more predictability or constraint.

CONCLUSION

Legal corpus linguistics is not yet “popular” in the sense of receiving widespread approval.⁷² And it is not necessarily “popular” in the sense of relating to the ordinary public; the historical corpora that legal scholars have most often relied upon tend to overrepresent elites’ language.⁷³ But legal corpus linguistics is popular in the sense of being increasingly *encountered*—cited by commentators and judges as

70. See, e.g., Dennis Baron, *Corpus Linguistics, Public Meaning, and the Second Amendment*, DUKE CTR. FOR FIREARMS L. BLOG (July 12, 2021), <https://firearmslaw.duke.edu/2021/07/corpus-linguistics-public-meaning-and-the-second-amendment/> [https://perma.cc/7CB9-EUZY]. Baron suggested this line of argument with respect to the meaning of “bear arms”:

It’s true that ordinary people didn’t write as much as the framers. But there’s no proof that ordinary people in the federal period said they were bearing arms when they hunted deer, elk, buffaloes, or rabbits. Nor is there any evidence that elite writers like Madison and the members of Congress who carefully edited and revised the Second Amendment baked a non-elite, non-military sense of *bear arms* into the amendment as a concession to an unattested ‘ordinary’ usage.

Id.

71. See, e.g., Nourse & Eskridge, *supra* note 1, at 1721 (“As this Article suggests, in any difficult case, the textualist judge starts with two potentially outcome-determinative decisions: a *choice of text*—the scope of text the judge decides to focus on when interpreting a statute—and a *choice of context* surrounding this text.”).

72. See *supra* note 20 and accompanying text.

73. See Bernstein, *supra* note 69 (noting that a popular corpus of founding era language represents a “tiny minority” of the founding era population, consisting of the language of “political superstars, lawmakers and government agents, [and] a few legal scholars.”).

relevant to the public meaning of “commerce,” “search,” and maybe even “bear arms.”⁷⁴

This essay has argued that legal corpus linguistics is unlikely to provide easy answers in hard cases of interpretation. Will the fact that legal corpus linguistics often admits of “clashing” arguments sap its popularity? It is unlikely. For one, that an interpretive tool could support putatively clashing arguments does not imply that the tool is ultimately flawed—perhaps the clashes can be resolved. Some argue that legal corpus linguistics’ current problems are largely the result of proponents’ presentation of a “highly impoverished version of [corpus linguistics]” and caution that legal scholars and judges should avoid “reduc[ing it] to the point of caricature.”⁷⁵ As legal corpus linguistics develops, perhaps the judges and scholars who rely upon these tools will clarify the appropriate methodological moves.

But even if the clashing is more fundamental, I would bet that legal corpus linguistics is here to stay. Despite its critics, textualism is increasingly influential at the Supreme Court and lower courts.⁷⁶ And textualist argument requires apparent textualist evidence.

Perhaps “clashing corpora” will share the fate of the “dueling canons” and “dueling dictionaries.” It’s been seventy years since Llewellyn noted the “dueling canons” and at least twenty years since the observation of “dueling dictionaries.” Today’s Supreme Court regularly relies on both tools.⁷⁷

74. See also Tobia, *supra* note 14 (documenting judge’s appeals to corpus linguistics, rising sharply over the past five years).

75. Stefan Th. Gries, *Corpus Linguistics and the Law: Extending the Field from a Statistical Perspective*, 86 BROOK. L. REV. 321, 324 (2021).

76. See *supra* notes 1–10 and accompanying text.

77. Anita S. Krishnakumar, *Cracking the Whole Code Rule*, 96 N.Y.U. L. REV. 76, 97 (2021) (reporting that the Roberts court relies on language and grammar canons in 8.7% of statutory meaning cases, substantive canons in 14.9% of such cases, and dictionaries in 21.6% of such cases).