

ARTICLES

The Corpus and the Critics

Thomas R. Lee[†] & *Stephen C. Mouritsen*^{††}

Most any approach to interpretation of the language of law begins with a search for ordinary meaning. Increasingly, judges, scholars, and practitioners are highlighting shortcomings in our means for assessing such meaning. With this in mind, we have proposed the use of the tools of corpus linguistics to take up the task. Our proposals have gained traction but have also seen significant pushback.

The search for ordinary meaning poses a series of questions that are amenable to evaluation and analysis using evidence of language usage. And we have proposed to use the tools of corpus linguistics—tools for assessing patterns of language usage in large databases of naturally occurring language—to introduce transparent, falsifiable evidence on the questions at stake. Our critics raise a series of challenges, asserting that our methods test the wrong language community, pose notice problems, are inaccurate measures, and rest on certain fallacies.

We show that the criticisms are largely in error and ultimately highlight some of the main selling points of our proposed methods. We do so in reference to two canonical Supreme Court cases that have been discussed in the literature in this field (Muscarello v. United States and Taniguchi v. Kan Pacific Saipan, Ltd.) and also a more recent case (Mont v. United States). In analyzing these cases (particularly the most recent one), we also outline a framework for some proposed refinements to the methods we have advocated previously.

INTRODUCTION276
I. LAW AND CORPUS LINGUISTICS: RESTATEMENT AND REFINEMENT281

[†] Associate Chief Justice of the Utah Supreme Court and Distinguished Lecturer in Law at Brigham Young University J. Reuben Clark Law School, Lecturer on Law at Harvard Law School, and Lecturer in Law at The University of Chicago Law School.

^{††} Shareholder at Parr Brown Gee & Loveless and Adjunct Professor at Brigham Young University J. Reuben Clark Law School. The authors acknowledge the editorial input of James Heilpern and Benjamin Lee, who contributed to an early draft of this paper, and express thanks to those who commented on earlier drafts or offered insights in response to presentations in various conferences, symposia, and talks. Thanks to the Association of American Law Schools and Brigham Young University, who each sponsored conference sessions at which the ideas in this piece were initially vetted. Thanks also to Seth Cannon, Brian Casper, Dante Chambers, Spencer Crawford, Josh Jones, Zachary Lutz, Christopher Melling, Elizabeth Nielson, Monick Perone, Jackson Skinner, and Kyle Tanner for their able research assistance.

A.	The Why of Corpus Linguistic Analysis: Deficiencies in Current Legal Interpretation.....	281
1.	Judicial intuition.	285
2.	Dictionaries.	286
3.	Etymology.	288
4.	Canons of construction.	288
B.	The How of Corpus Linguistic Analysis: A Primer	291
C.	The Contributions of Corpus Linguistic Analysis	296
1.	Transparency.	297
2.	Refining legal theory.	298
3.	Testing the tools of statutory interpretation.	299
II.	CORPUS LINGUISTICS AND THE SPEECH COMMUNITY.....	300
A.	The Purpose of Legal Language.....	301
B.	The Audience of Legal Language	302
C.	The Usage and Syntax of Legal Language	303
III.	CORPUS LINGUISTICS AND NOTICE	304
IV.	SURVEY-BASED CRITIQUES OF THE ACCURACY OF CORPUS LINGUISTIC ANALYSIS	311
A.	Tobia’s Concept Condition Results Do Not Establish Baseline Truth.....	315
1.	Survey methods do not give privileged access to linguistic “competence.”	316
2.	Tobia’s data are unhelpful given his study design.....	323
3.	Tobia’s data fail to consider syntactic and pragmatic context....	326
B.	Tobia’s Survey Does Not Show That Corpus Tools Are Biased	330
1.	Tobia’s results come from his flawed corpus condition design. ...	332
2.	Tobia’s fallacies misstate our views.	334
3.	Tobia’s critiques highlight the need to refine our methodology. ...	337
V.	ORDINARY MEANING AND FREQUENCY OF OCCURRENCE	341
	CONCLUSION	350
	APPENDIX.....	360
A.I	Corpus Analysis for <i>Mont</i>	360
A.II	Double-Dissociation Analysis	361

INTRODUCTION

A decade ago we proposed the use of the tools of corpus linguistics in the interpretation of legal language. First in a student

law review note¹ and then in a concurring opinion,² we began to highlight a series of shortcomings in traditional tools for assessing the ordinary meaning of legal language (etymology, dictionaries, judicial intuition) and to propose that corpus linguistic analysis could fill those voids. We extended our ideas in further opinions,³ blog posts,⁴ workshops,⁵ and conferences.⁶ And we eventually presented a thorough defense of the use of corpus linguistics in statutory interpretation in *Judging Ordinary Meaning*,⁷ an extensive article followed by other pieces proposing extensions of the methodology to constitutional⁸ and contract interpretation.⁹

Corpus linguistics is the study of language through the analysis of large bodies of naturally occurring text. By sampling and analyzing words and phrases in their natural setting we can access evidence about meaning and usage that before we could only guess about, making our assessment of the meaning of the language of the law more transparent and independently verifiable.

¹ Stephen C. Mouritsen, Note, *The Dictionary Is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning*, 2010 BYU L. REV. 1915, 1919 [hereinafter Mouritsen, *Definitional Fallacies*].

² J.M.W. v. T.I.Z. (*In re* Adoption of Baby E.Z.), 266 P.3d 702, 723–29 (Utah 2011) (Lee, J., concurring in part and concurring in the judgment).

³ See, e.g., State v. Rasabout, 356 P.3d 1258, 1275–82 (Utah 2015) (Lee, A.C.J., concurring in part and concurring in the judgment) (using corpus linguistic analysis to ascertain the meaning of “discharge” in a statutory scheme and describing how corpus linguistics works to aid judicial decision-making when resort to dictionary definitions does not reveal ordinary meaning in light of multiple possible definitions).

⁴ See, e.g., Thomas Lee & Stephen Mouritsen, *Judging Ordinary Meaning with Corpus Linguistics*, WASH. POST (Aug. 8, 2017), <https://perma.cc/8NSU-GFXF> [hereinafter Lee & Mouritsen, *Corpus Linguistics*].

⁵ See, e.g., *19th Annual Faculty Conference: Corpus Linguistics and Legal Interpretation*, FEDERALIST SOC’Y (Jan. 5, 2017), <https://perma.cc/6MQL-ZFFQ>.

⁶ The Brigham Young University J. Reuben Clark Law School has held a Law & Corpus Linguistics Conference for each of the last five years. See *Call for Papers: 5th Annual Law & Corpus Linguistics Conference*, BYU L., <https://perma.cc/UQS8-7YJX>. The Section on Law and Interpretation of the Association of American Law Schools also convened a program on “Corpus Linguistics: The Search for Objective Interpretation” at the annual meeting in New Orleans in January 2019. See *Program: 2019 AALS Annual Meeting*, ASS’N OF AM. L. SCHS. (Jan. 2019), <https://perma.cc/3K22-ZJTB>. Many of the ideas in this Article were presented at the BYU conference in 2020 and the AALS meeting in 2019.

⁷ See generally Thomas R. Lee & Stephen C. Mouritsen, *Judging Ordinary Meaning*, 127 YALE L.J. 788 (2018) [hereinafter Lee & Mouritsen, *Judging Ordinary Meaning*].

⁸ See generally Thomas R. Lee & James C. Phillips, *Data-Driven Originalism*, 167 U. PA. L. REV. 261 (2019).

⁹ Stephen C. Mouritsen, *Contract Interpretation with Corpus Linguistics*, 94 WASH. L. REV. 1337 (2019) [hereinafter Mouritsen, *Contract Interpretation*].

Judges, scholars, and advocates have begun to take note. In recent opinions, judges on various state supreme courts¹⁰ and federal courts of appeals¹¹ have accepted the invitation to bring corpus linguistic analysis to bear in the interpretation of legal language. And in 2018, Justice Clarence Thomas employed the tools of corpus linguistics in two separate opinions on questions of constitutional law.¹² Judges are not alone in this regard; legal scholars¹³ and practicing lawyers¹⁴ are also increasingly advocating the use of these tools.

Not everyone is impressed. Our advocacy for the increased use of corpus linguistic tools in legal interpretation has drawn skepticism and criticism from various quarters. And the criticism has exploded of late—with a range of scholars seeking to pump the brakes on or outright repudiate the utility of corpus tools in

¹⁰ See, e.g., *Richards v. Cox*, 450 P.3d 1074, 1079–81 (Utah 2019); *State v. Lantis*, 447 P.3d 875, 880–81 (Idaho 2019); *People v. Harris*, 885 N.W.2d 832, 838–39 (Mich. 2016); *Rasabout*, 356 P.3d at 1275–82 (Lee, A.C.J., concurring in part and concurring in the judgment).

¹¹ See, e.g., *Caesars Ent. Corp. v. International Union of Operating Engineers Local 68 Pension Fund*, 932 F.3d 91, 95–96 (3d Cir. 2019) (employing corpus linguistics in conjunction with dictionary definitions to interpret the phrase “previously required” and concluding, in part because “previously” most commonly “co-occurred with . . . ‘had’ (35%) and ‘been’ (15%)—perfect tense verbs that connote completed action,” that “to say something is ‘previously required’ is to suggest it is no longer required”); *Wilson v. Safelite Grp., Inc.*, 930 F.3d 429, 439–45 (6th Cir. 2019) (Thapar, J., concurring in part and concurring in the judgment) (proposing the use of corpus linguistics as one “important tool” in the “judicial toolkit” and relying on searches in the Corpus of Historical American English to further support the majority’s statutory conclusion); see also, e.g., *Wright v. Spaulding*, 939 F.3d 695, 700 n.1 (6th Cir. 2019) (taking the unusual step of asking the parties to submit supplemental briefing on whether corpus linguistic analysis helped illuminate the original meaning of Article III’s case-or-controversy requirement but noting “that corpus linguistics turned out not to be the most helpful tool” in this particular case).

¹² See *Carpenter v. United States*, 138 S. Ct. 2206, 2238–39, 2238 n.4 (2018) (Thomas, J., dissenting) (using corpus linguistics to argue that the term “search” in the Fourth Amendment “did not mean a violation of someone’s reasonable expectation of privacy” “[a]t the founding”); *Lucia v. SEC*, 138 S. Ct. 2044, 2056–57 (2018) (Thomas, J., concurring) (quoting *NLRB v. SW Gen., Inc.*, 137 S. Ct. 929, 946 (2017) (Thomas, J., concurring)) (relying on a law review article, Jennifer L. Mascott, *Who Are “Officers of the United States”?*, 70 STAN. L. REV. 443, 564 (2018), that used corpus linguistic analysis and concluding, like that article, that “Officers of the United States” “encompassed all federal civil officials ‘with responsibility for an ongoing statutory duty’”).

¹³ See generally, e.g., Jennifer L. Mascott, *Who Are “Officers of the United States”?*, 70 STAN. L. REV. 443 (2018); Lawrence M. Solan, *Can Corpus Linguistics Help Make Originalism Scientific?*, 126 YALE L.J.F. 57 (2016); Lawrence M. Solan & Tammy Gales, *Corpus Linguistics as a Tool in Legal Interpretation*, 2017 BYU L. REV. 1311.

¹⁴ For example, the emoluments litigation after President Donald Trump’s election drew substantial attention from linguistics scholars. See generally, e.g., Brief of Amici Curiae Professor Clark D. Cunningham and Professor Jesse Egbert on Behalf of Neither Party, *In re Trump*, 928 F.3d 360 (4th Cir. 2019) (No. 18-2486).

law. Some of the criticism questions the viability of our proposed methodology by challenging the speech community that our methods have sought to test¹⁵ or suggesting that there is a public notice problem with a resort to tools that the public lacks the means or expertise to access.¹⁶ A more recent piece in the *Harvard Law Review* presents a more fundamental, empirical critique.¹⁷ In that article, Professor Kevin Tobia advances the results of an extensive series of surveys aimed at demonstrating that corpus tools are an inaccurate measure of ordinary meaning¹⁸ and are systematically biased in favor of the wrong sense of ordinary meaning (“prototypical” meaning).¹⁹ Tobia also presents his survey results in support of a further criticism (echoed by other scholars)—that corpus tools credit only the “most frequent” sense of a given legal term—the wrong measure of ordinary meaning.²⁰

We welcome the pushback. The method that we have proposed is novel and disruptive. And the opportunity to respond to our critics will help refine the methodology that we advocate, situate it more carefully within existing theories and practices of interpretation, and either refute or credit the major critiques that have been identified.

That said, some of the criticisms have posited a caricature of our proposed approach. With that in mind, we begin in Part I with clarification and refinement of our central idea—that the tools of corpus linguistics can address shortcomings of existing tools for assessing the ordinary meaning of legal language. The balance of the Article is a response to the main criticisms that have been raised to date. In Part II we address concerns about whether corpus tools are aimed at the right speech community. Here we clarify that the choice of speech community presents a problem for legal theory—an intentionalist who is interested in crediting the meaning attributed to a lawmaker would want to study specialized legislative speech, while a textualist who is concerned about public notice would be more interested in the everyday language of the

¹⁵ See Evan C. Zoldan, *Corpus Linguistics and the Dream of Objectivity*, 50 SETON HALL L. REV. 401, 430–35 (2019); Anya Bernstein, *Democratizing Interpretation*, 60 WM. & MARY L. REV. 435, 458–60 (2018).

¹⁶ See Carissa Byrne Hessick, *Corpus Linguistics and the Criminal Law*, 2017 BYU L. REV. 1503, 1514–16.

¹⁷ Kevin P. Tobia, *Testing Ordinary Meaning*, 134 HARV. L. REV. 727, 753–77 (2020).

¹⁸ See *id.* at 766.

¹⁹ See *id.* at 770.

²⁰ See *id.* at 759, 792; see also Hessick, *supra* note 16, at 1506.

public. We also acknowledge that there may be circumstances where language evidence from different regions, industries, social backgrounds, or racial, ethnic, or community identities may appropriately be brought to bear on an interpretive question. And we emphasize that this concern is not a weakness but a strength of corpus tools, which allow the interpreter to construct a corpus of any appropriate speech community and analyze the language of that community.

Part III speaks to questions about whether a lack of public access to corpus analysis raises notice concerns. We concede the need for fair notice and acknowledge that corpus tools will not be accessible to the general public. But we show that concerns about fair notice cut against maintaining the status quo and strongly in favor of the addition of corpus tools to the judge's existing toolbox.

In Part IV we turn to Tobia's "accuracy" criticism—the claim that corpus analysis is problematic because it fails to paint an accurate picture of ordinary meaning. Tobia's accuracy charge is based on the premise that the results of his "concept condition" surveys paint an accurate picture of ordinary meaning.²¹ Tobia also argues that corpus analysis is biased in a particular way—that it measures only language "prototype," and is premised on a fallacy of "nonappearance" (that the nonappearance in a corpus of any uses of a tested term means that such use does not count as "ordinary").²² We highlight a range of problems with these criticisms. We note that Tobia's starting premise (the accuracy of Tobia's "concept" surveys) is merely presumed with no open attempt to defend it and little engagement with relevant social science literature on survey methods. We engage with that literature and identify a broad range of reasons why this premise is wrong. And we show that Tobia's critiques of corpus methods are also flawed—that the purported fallacies are based on a misunderstanding of our methods and that the "prototype" results that Tobia complains about flow not from corpus analysis generally but from Tobia's survey design.

In Part V we consider what seems to be the most persistent claim of our critics (Tobia, and also Professor Carissa Byrne Hessick)—that our methods always credit only the most frequent sense of a tested term in a corpus. This misstates our position. We

²¹ Tobia, *supra* note 17, at 772–74.

²² *Id.* at 789–90, 795.

respond to it by highlighting the utility of corpus tools in assessing any of a range of senses of ordinary meaning.

We conclude with some observations about next steps. Though most of the criticisms either misstate our proposals or highlight concerns that backfire on our critics, some of them help underscore the need for us to enhance and refine the interpretive methodology that we envision. We close with some thoughts on specific steps that can be taken.

I. LAW AND CORPUS LINGUISTICS: RESTATEMENT AND REFINEMENT

The critics of corpus methods have raised important questions. But along the way they have also missed some fundamental tenets and important nuances of our proposed approach. With that in mind, we begin with a restatement and refinement of the potential contributions of corpus linguistic analysis set forth in our prior writings. To do so, (A) we highlight the deficiencies of the principal tools and techniques currently used by judges to discern the communicative content of legal texts; (B) we describe the corpus linguistic enterprise and explain how tools imported from this field can help address existing shortcomings; and (C) we synthesize and summarize the principal contributions of the use of the methodology of corpus linguistics in the law of interpretation.

A. The Why of Corpus Linguistic Analysis: Deficiencies in Current Legal Interpretation

The first step for any inquiry into the meaning of a legal text is the search for its “communicative content”—the intended and understood meaning of the words of the law.²³ Jurists have long described the communicative content of legal texts as the “ordinary” or “plain” meaning of the law.²⁴ The inquiry into communicative content is sometimes distinguished from an attempt to discern the law’s legal effect—its legal content.²⁵ Scholars sometimes

²³ Lawrence B. Solum, *Communicative Content and Legal Content*, 89 NOTRE DAME L. REV. 479, 480 (2013).

²⁴ See William Baude & Ryan D. Doerfler, *The (Not So) Plain Meaning Rule*, 84 U. CHI. L. REV. 539, 545 (2017).

²⁵ See Solum, *supra* note 23, at 480.

refer to the first of these tasks as “interpretation” and the latter as “construction.”²⁶ Our focus is on the former.

Ordinary meaning is crucial to the interpretative enterprise because, as Professor William Eskridge, Jr., has written, “[a] polity governed by the rule of law aspires to have legal directives that are known to the citizenry, that are predictable in their application, and that officials can neutrally and consistently apply based upon objective criteria.”²⁷ And while the search for “ordinary meaning does not always yield predictable answers,” it does “yield greater predictability than any other *single* methodology.”²⁸

Almost everyone agrees that a search for the communicative content of the law is the starting point when interpreting the language of law. And countless cases thus frame the interpretive question in terms of a search for the ordinary meaning. And where a court can discern the ordinary or plain meaning, that is the end of the interpretive analysis.²⁹

Yet the general consensus on the appropriate starting point masks latent conflicts—acknowledged and subconscious—about our search for ordinary meaning. One set of unresolved problems goes to our legal theory of ordinary meaning. Another is in operationalizing the theory. In highlighting these two sets of problems we will refer to two of the canonical Supreme Court cases, which we considered in our recent scholarship—*Muscarello v. United States*³⁰ and *Taniguchi v. Kan Pacific Saipan, Ltd.*³¹ And we show that the problems are ongoing by including a more recent case, *Mont v. United States*.³²

Each of these cases presents a question of lexical ambiguity (the core kind of problem for which we have proposed the use of corpus tools)—a choice between alternative senses of a given statutory term or phrase. In *Muscarello* the question was whether a person “carries a firearm,” as that term is used in the mandatory minimum sentencing provision in 18 U.S.C. § 924(c)(1)(A), when

²⁶ See *id.* at 483; Amy Barrett, *The Interpretation/Construction Distinction in Constitutional Law: Annual Meeting of the AALS Section on Constitutional Law*, 27 CONST. COMMENT. 1, 1 (2010).

²⁷ WILLIAM N. ESKRIDGE, JR., INTERPRETING LAW: A PRIMER ON HOW TO READ STATUTES AND THE CONSTITUTION 35 (2016).

²⁸ *Id.* at 36 (emphasis in original).

²⁹ Presumably this is what Justice Elena Kagan means when she says, “[W]e’re all textualists now.” *The 2015 Scalia Lecture: A Dialogue with Justice Elena Kagan on the Reading of Statutes*, HARV. L. TODAY at 8:29 (Nov. 17, 2015), <https://perma.cc/APM3-Y9EU>.

³⁰ 524 U.S. 125 (1998).

³¹ 566 U.S. 560 (2012).

³² 139 S. Ct. 1826 (2019).

he transports it in the glove box of a truck he drives to a drug deal.³³ The interpretive question has no plain answer—each side has a plausible definition of “carry,” and the *Muscarello* Court was deeply divided, producing a 5–4 vote.³⁴ *Taniguchi* presented a similar kind of problem. There the question was whether the ordinary meaning of “interpreter” under 28 U.S.C. § 1920 is limited to real-time translation of oral speech or also includes translation of written texts.³⁵ Again the Court was split (this time 6–3),³⁶ and again each side has a plausible interpretation of the statutory term. *Mont* was along the same lines. There the question was whether the ordinary meaning of “imprison[ment] in connection with a conviction” for a crime under 18 U.S.C. § 3624(e) is limited to terms served after the conviction is entered or encompasses terms of pretrial detention that are later credited in the ultimate sentence.³⁷ The *Mont* Court was also deeply split (5–4) and again the dispute was over two plausible senses of the statutory term.³⁸

The threshold problem in these kinds of cases is in our imprecision in the theory of what we are seeking to measure. Are we looking for the understanding of a legislator, or of a member of the public who is governed by the law? Should we interpret a legal instrument through the lens of modern, linguistic conventions, or conventions from the time that it was drafted and went into effect? What are we looking for when we search for the plain or ordinary meaning of the language of law? Are we asking about the most common sense of a given term or phrase, or just whether a term is often or permissibly used in a certain way? And how do we resolve ambiguity if the meaning of the text is not clear?

Judges speak confidently about the need to credit ordinary meaning. But we have no “ordinary meaning of ordinary meaning.”³⁹ When we speak of ordinary meaning, we sometimes seem to be seeking the understanding of a legislator and sometimes seem to be seeking public meaning. More fundamentally, we are

³³ *Muscarello*, 524 U.S. at 126–27.

³⁴ *See id.* at 126.

³⁵ *Taniguchi*, 566 U.S. at 562.

³⁶ *See id.* at 561.

³⁷ *Mont*, 136 S. Ct. at 1829 (alteration in original).

³⁸ *See id.*

³⁹ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 798 (quotation marks omitted).

inconsistent in the standard that we have in mind (permissible versus most frequent use, etc.).⁴⁰

These problems are on display in *Taniguchi*, *Muscarello*, and *Mont*. In *Taniguchi*, the majority says that only a real-time translator falls within the ordinary meaning of “interpreter.” In so concluding it asserts that “an interpreter is *normally understood* as one who translates orally from one language to another.”⁴¹ Although the majority concedes that a written translator could also be spoken of as an interpreter, that sense is rejected as “hardly a common or ordinary meaning.”⁴² The dissent seems to have a different sense of ordinary meaning in mind. It concludes that a written translator is an interpreter because we use the term that way “more than occasionally.”⁴³ The majority and dissenting opinions in *Muscarello* are perhaps even more problematic. The majority opinion is internally inconsistent—asserting that transporting a gun in a glove box is within the ordinary sense of “carry” both because it is a common sense of the term and insisting that this is the “primary” sense of the term.⁴⁴ And the dissent has parallel problems.⁴⁵

These imprecisions are also reflected in the more recent opinions in *Mont*. In concluding that pretrial detention falls within the meaning of a provision that tolls a term of supervised release, the majority said that the phrase “imprisoned in connection with a conviction” “may well include” and can “encompass” a term of “pretrial detention credited toward another sentence for a new conviction.”⁴⁶ The dissent had a different conception of ordinary meaning in mind. It sought for “normal usage” or “the colloquial sense” of the statutory phrase—the one “most naturally understood in context.”⁴⁷

⁴⁰ See *id.* at 800.

⁴¹ *Taniguchi*, 566 U.S. at 569 (emphasis added).

⁴² *Id.*

⁴³ *Id.* at 576 (Ginsburg, J., dissenting).

⁴⁴ *Muscarello*, 524 U.S. at 128–29 (asserting that “many”—“perhaps more than one-third”—of the uses of “carry” with respect to a “vehicle” and “weapon” in the *New York Times* and *U.S. News* databases reflect the “transport-in-a-car” sense of carry, while also stating that this sense is the “primary” sense that Congress had in mind).

⁴⁵ See *id.* at 143–44 (Ginsburg, J., dissenting) (asserting “that ‘carry’ is a word *commonly used* to convey various messages,” and that it “*could mean*” either personally bear or transport in a vehicle (emphasis added)); *id.* at 149 (stating that the personally bear sense of “carry” is “hardly implausible nor at odds with an *accepted meaning*” of the statutory terms (emphasis added)).

⁴⁶ *Mont*, 139 S. Ct. at 1832.

⁴⁷ *Id.* at 1837–38 (Sotomayor, J., dissenting).

Imprecisions in our theory of ordinary meaning are one problem. Yet additional problems are apparent when we examine the tools that jurists employ in operationalizing ordinary meaning. Judges typically measure ordinary meaning by resort to linguistic intuition and by reference to dictionaries, etymology, and canons of construction. Yet these tools are insufficient. Their weaknesses have been discussed at length elsewhere, both by us and others.⁴⁸ But we briefly sketch the major problems associated with each below.

1. Judicial intuition.

Judges often begin and end their assessment of a text's linguistic meaning with a gut-level intuitive determination. As a competent English speaker, a judge may present her view of the ordinary understanding of a phrase like "carry a firearm" simply based on her intuitive understanding of the language. This is an inevitable—and entirely appropriate—starting point for interpretation. But there are ample grounds for questioning the wisdom of relying exclusively on the intuition of an individual judge as the end point.

Absent corpus linguistic analysis, "a judge has no way of determining whether . . . her own interpretation is widely shared."⁴⁹ And because "humans tend to notice unusual occurrences [of words] more than typical occurrences," judges run the risk of overcrediting the frequency of obscure word senses when relying exclusively on intuition.⁵⁰ This risk is exacerbated when a word has many different senses.

The inquiry into ordinary meaning is typically framed as a search for how the ordinary, reasonable person would understand a given term or phrase.⁵¹ Yet the judge represents only a single data point—a data point that is unlikely representative of the

⁴⁸ See, e.g., Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 806–07; Mouritsen, *Definitional Fallacies*, *supra* note 1, at 1925–26, 1939–41; Solan & Gales, *supra* note 13, at 1331–36; James A. Heilpern, *Dialects of Art: A Corpus-Based Approach to Technical Term of Art Determinations in Statutes*, 58 *JURIMETRICS* 377, 380–81 (2018).

⁴⁹ Lawrence Solan, Terri Rosenblatt & Daniel Osherson, *False Consensus Bias in Contract Interpretation*, 108 *COLUM. L. REV.* 1268, 1273 (2008).

⁵⁰ DOUGLAS BIBER, SUSAN CONRAD & RANDI REPPEN, *CORPUS LINGUISTICS: INVESTIGATING LANGUAGE STRUCTURE AND USE* 3 (1998).

⁵¹ See, e.g., Frank H. Easterbrook, *The Role of Original Intent in Statutory Construction*, 11 *HARV. J.L. & PUB. POL'Y* 59, 65 (1988) ("We should look at the statutory structure and hear the words as they would sound in the mind of a skilled, objectively reasonable user of words. . . . The meaning of statutes is to be found . . . in the understanding of the objectively reasonable person.").

general population. Judges are generally wealthier and better educated than the average American and have often spent significant time studying at elite institutions. Such judges may not be ideally suited to make gut-level assessments of ordinary usage. If the question is whether “carry a firearm” is typically limited to personal bearing (rather than transporting in a vehicle), the judge may not be in a good position to answer the question.

The problem may be compounded by the introduction of a time-dimension problem. A modern originalist judge, for example, may posit the need to interpret historical documents in light of the semantic conventions of the time they were drafted and ratified. But he may not be in a good position to intuitively assess the meaning of language enacted in prior eras due to “linguistic drift—the notion that language usage and meaning shifts over time.”⁵² His twenty-first-century linguistic intuition may be out of sync with the original ordinary meaning of the text presented. Some changes in the language of law can be quite dramatic and occur for no apparent reason and can be nigh unto impossible for a judge to detect.⁵³

The judge may also bring latent biases or prejudices about preferred outcomes in cases that come before her. Opening the door to judicial resort to intuition risks motivated reasoning and confirmation bias. And it obscures the basis of judicial decisions, precluding the ability of parties, counsel, or commentators to assess and challenge the decision. An implicit or explicit “take my word for it” decision is impossible to question or falsify.

2. Dictionaries.

Perhaps because of these limitations on human linguistic intuition, judges frequently turn to dictionaries for help.⁵⁴ When a judge cites a dictionary, she is seeking to avoid the appearance of a subjective decision by a single individual and to signal objectivity and authoritativeness. But there is often more veneer than

⁵² Lee & Phillips, *supra* note 8, at 265.

⁵³ See John W. Welch & James A. Heilpern, *Recovering Our Forgotten Preamble*, 91 S. CAL. L. REV. 1021, 1065 (2018).

⁵⁴ See, e.g., *Taniguchi*, 566 U.S. at 566–69 (citing a wide range of dictionaries in support of the conclusion that the ordinary sense of “interpreter” is limited to real-time translation of oral speech); *Muscarello*, 524 U.S. at 128 (first citing 2 THE OXFORD ENGLISH DICTIONARY 919 (2d ed. 1989); then citing WEBSTER’S THIRD NEW INTERNATIONAL DICTIONARY 343 (1986); and then citing RANDOM HOUSE DICTIONARY OF THE ENGLISH LANGUAGE UNABRIDGED 319 (2d ed. 1987)) (supporting the conclusion that transporting something in a vehicle is within the primary sense of the verb “carry”).

substance to this impulse—not just because dictionaries do not emerge from some lexical Mount Sinai (a fallible “someone sat there and wrote the dictionary”)⁵⁵—but also because dictionaries typically do not answer the ordinary meaning questions presented to courts. In most cases, the dictionary will simply show that both sides have a plausible argument.

This is true in the canonical cases that we have studied. In *Muscarello* and *Taniguchi*, both the majority and dissenting opinions cited dictionaries in support of their views: The *Muscarello* majority cited a “convey . . . in any vehicle” definition of “carry”⁵⁶ and the dissent cited a “bear . . . upon the person” definition.⁵⁷ The *Taniguchi* majority cited “oral[]” translator definitions of “interpreter”⁵⁸ and the dissent cited a broader sense (that included the translation of written texts).⁵⁹

The competing opinions in *Mont* are a bit more nuanced. The majority cites dictionary definitions of “imprison” in support of its view that this term can “encompass pretrial detention.”⁶⁰ And the dissent focuses more on the need to assess phrasal meaning, while insisting that the “natural[]” or “normal” usage of “imprisoned in connection with a conviction” is “postconviction incarceration.”⁶¹ But again dictionaries are front and center.

A further problem is in the details of the courts’ reliance on dictionaries. Judges tend to cite dictionaries for propositions they do not establish. A principal problem is in what we have called the “sense-ranking fallacy”—the notion that a given sense of a statutory term is ordinary (in the sense of being more commonly used) because it is listed higher in the dictionary’s list of senses. This is on display in *Muscarello*, where the majority credits the “transport in a vehicle” sense of carry because it is listed ahead of

⁵⁵ Mouritsen, *Definitional Fallacies*, *supra* note 1, at 1916 (quoting Lawrence Solan, *When Judges Use the Dictionary*, 68 AM. SPEECH 50, 50 (1993)).

⁵⁶ *Muscarello*, 524 U.S. at 128 (quoting 2 THE OXFORD ENGLISH DICTIONARY 919 (2d ed. 1989)).

⁵⁷ *Id.* at 143 (Ginsburg, J., dissenting) (quoting *Carry*, BLACK’S LAW DICTIONARY (6th ed. 1990)).

⁵⁸ *Taniguchi*, 566 U.S. at 568–69, 568 n.2 (asserting that “only a handful” of dictionaries include the written translator sense of “interpreter” but “all” of them speak of the oral translator sense).

⁵⁹ *See id.* at 576 (Ginsburg, J., dissenting).

⁶⁰ *Mont*, 139 S. Ct. at 1832 (stating that “the term ‘imprison’ has meant ‘[t]o put in a prison,’ ‘to incarcerate,’ [and] ‘[t]o confine a person, or restrain his liberty, in any way’” (first quoting *Imprison*, BLACK’S LAW DICTIONARY (5th ed. 1979); and then quoting 5 THE OXFORD ENGLISH DICTIONARY 113 (1933) (citing *Imprison*, BLACK’S LAW DICTIONARY (10th ed. 2014))).

⁶¹ *Id.* at 1837–38 (Sotomayor, J., dissenting).

the “personally bear” sense in the *Oxford English Dictionary*’s list of senses.⁶² But the analysis is fallacious, as evidenced by the fact that the very dictionaries relied upon by courts explain that the order of definitions is either arbitrary or reflects historical usage (oldest first).⁶³ And this again underscores the need to look elsewhere to check judicial intuition, or to more reliably discern and determine the ordinary meaning of the language of the law.

3. Etymology.

Judges also sometimes look to a word’s etymology—its historical pedigree in other languages—to help decipher its linguistic meaning. The majority made this move in *Muscarello*. It noted that the verb “carry” traces from “the Latin ‘carum,’ which means ‘car’ or ‘cart,’” and “from Old French ‘carier’ and late Latin ‘carricare,’ which meant to ‘convey in a car.’”⁶⁴ And on that basis the Court concluded that the ordinary meaning of the phrase “carry a firearm” must include transporting it in a locked glove compartment.⁶⁵

This again is fallacious. The idea that the ordinary meaning of words in phrases in our language can be assessed by the way they were used in completely different languages in centuries past is incoherent. That’s not how language works. “[W]ord[s] needn’t mean forever what they meant in Greek or Latin.”⁶⁶ If they did, “*December* would mean the tenth month, and an *anthology* would mean a bouquet of flowers.”⁶⁷

4. Canons of construction.

Judges may also look to “linguistic” or “semantic” canons—often Latin-labeled principles that claim to identify shared

⁶² See *Muscarello*, 524 U.S. at 128 (citing 2 THE OXFORD ENGLISH DICTIONARY 919 (2d ed. 1989)) (crediting the former sense as “primary” and dismissing the latter as “special” and noting that the “first” definition in the *Oxford English Dictionary* is the transport sense).

⁶³ See 1 THE OXFORD ENGLISH DICTIONARY xxix (2d ed. 1989) (“[T]hat sense is placed first which was actually the earliest in the language: the others follow in the order in which they appear to have arisen.”).

⁶⁴ See *Muscarello*, 524 U.S. at 128 (first quoting THE BARNHART DICTIONARY OF ETYMOLOGY 146 (Robert K. Barnhart ed., 1988); and then quoting 2 THE OXFORD ENGLISH DICTIONARY 919 (2d ed. 1989)).

⁶⁵ See *id.*

⁶⁶ R.L.G., *The Etymological Fallacy*, THE ECONOMIST (Aug. 2, 2011), <https://perma.cc/R9MB-QCZB>.

⁶⁷ Lee & Mouritsen, *supra* note 7, at 809 (emphasis in original) (providing the etymology of “anthology,” THE BARNHART CONCISE DICTIONARY OF ETYMOLOGY 29 (Robert K. Barnhart ed., 1995), and “December,” *id.* at 188).

premises or conventions of language usage—to inform their search for the ordinary meaning of statutory language.⁶⁸ Examples include the notion that the meaning of a word may be informed by the meaning of surrounding terms and phrases (*eiusdem generis* and *noscitur a sociis*) and the presumption that each word or phrase in the law should be given independent meaning (the surplusage canon).⁶⁹

Canons have been around for centuries and purport to articulate shared premises of how language is ordinarily understood in context, potentially providing an air of objectivity and authoritativeness like dictionaries. But like dictionaries, canons also fall short for a number of reasons. First, there is internal tension in the way the canons are framed. At most they state rebuttable presumptions. And the law surrounding canons has never done a good job of identifying the circumstances in which the presumption should be rebutted. This highlights a related (and more central) problem. It is unclear whether they in fact capture ordinary usage. Thus, courts have cited these canons for centuries without questioning whether (or when) they are consistent with ordinary language usage, which means that the linguistic premises that undergird them are untested. That is problematic. This leads to a third, related problem. Many of the canons are subject to counter-canons,⁷⁰ which open the door to the very subjectivity and motivated reasoning that resort to ordinary meaning claims to avoid.

Canon-based reasoning is invoked in *Mont*. There the majority buttresses its interpretation of “imprisonment in connection with a conviction” by reference to the surplusage canon.⁷¹ Because the statute speaks not just of imprisonment but “imprisonment in connection with a conviction,” the majority says that imprisonment must be interpreted broadly.⁷² “[I]f imprisonment referred only to ‘confinement that is the result of a penalty or sentence,’” the majority says that “the phrase ‘in connection with a conviction’ [would] becom[e] entirely superfluous.”⁷³

⁶⁸ William Baude & Stephen E. Sachs, *The Law of Interpretation*, 130 HARV. L. REV. 1079, 1084, 1088 (2017).

⁶⁹ See *id.* at 1087, 1126.

⁷⁰ See Karl N. Llewellyn, *Remarks on the Theory of Appellate Decision and the Rules or Canons of About How Statutes Are to Be Construed*, 3 VAND. L. REV. 395, 401–06 (1950).

⁷¹ *Mont*, 139 S. Ct. at 1832 (alteration omitted).

⁷² *Id.* (alteration omitted).

⁷³ *Id.* (second and third alterations in original) (quoting *United States v. Goins*, 516 F.3d 416, 421 (6th Cir. 2008)).

As with dictionaries or the use of judicial intuition, the point is not that linguistic canons are never helpful. It's just that they may not always be independently up to the task of gauging ordinary meaning. These canons should "stand or fall" by whether they reflect ordinary use in a given linguistic setting.⁷⁴ And they will not always do that.

The *Mont* analysis is illustrative. The surplusage canon has great force in some linguistic settings. But no one thinks that every term in every legal phrase always has independent meaning. We also have a countercanon that speaks to the idea of surplus language included in an abundance of caution or by way of emphasis. The Supreme Court invoked the countercanon in the same term in which it decided *Mont*—in *Rimini Street, Inc. v. Oracle USA, Inc.*⁷⁵ There the Court acknowledged prior cases in which it had said that "some redundancy is hardly unusual" in some statutory settings.⁷⁶ Thus, the Court embraced a general rule favoring an "interpretation of a statute" that "would avoid redundancy" over an alternative interpretation that "would cause some redundancy." But it emphasized that this is only a presumption—noting that "[s]ometimes the better overall reading of the statute contains some redundancy."⁷⁷

This is the difficulty with the surplusage canon, and many others. A canon isn't a hard-and-fast rule. They are general presumptions subject to rebuttal. And unless and until we can speak clearly about the proper basis for the presumptions and the grounds for rebuttal, we will not have a set of tools that can properly cabin our judges' intuition in legal interpretation.

* * *

If judicial intuition, dictionaries, etymology, and linguistic canons are insufficient tools for discovering a statute's ordinary meaning, what is to be done? We see three possible answers: First, we can abandon the "the view that we can explain our legal norms by pointing to the ordinary communicative content of our

⁷⁴ Baude & Sachs, *supra* note 68, at 1084.

⁷⁵ 139 S. Ct. 873 (2019).

⁷⁶ *Id.* at 881 (quotation marks omitted) (quoting *Marx v. Gen. Revenue Corp.*, 568 U.S. 371, 385 (2013)).

⁷⁷ *Id.*

legal texts” (the “Standard Picture”)⁷⁸ completely and revive the premises of “strong purposivism”—freeing judges to make decisions based on some sort of sense of the purpose or intent of the legislature, with the text as a secondary consideration.⁷⁹ Second, we can bury our heads in the sand and continue to pretend that there are no problems with our faulty tools in the name of convention or tradition because we’ve always done it that way. Or third, we can open our eyes to the problems with our existing methods and search for new tools and solutions that will do a better job of helping judges and practitioners evaluate claims about the meaning of legal instruments. We can evaluate whether and to what extent existing tools successfully provide evidence of meaning, and whether such tools (like the canons of interpretation) can be reformulated to better represent prevailing linguistic conventions. We see the third answer as the only viable one, and we see the tools of corpus linguistics as an important step forward in the law’s attempt to better capture the “standard picture.”

B. The How of Corpus Linguistic Analysis: A Primer

Corpus linguistics is a subfield of linguistics that investigates real language use and function by analyzing electronic collections of naturally occurring texts. These collections are called “corpora,” and are digitally searchable compilations of real-world sources—books, newspapers, speeches, scholarly articles, television transcripts, etc.—drawn from a particular speech community and annotated with linguistic metadata. They are designed to be representative samples of all of the language produced by a community, and therefore reflect (both in diversity and relative frequency) the language patterns within that community.⁸⁰

Linguists have designed a number of different tools for analyzing the evidence drawn from these databases, many of which we have barely begun to think about in terms of how they might assist the judge in her judicial task. But here we will focus on just two—what linguists refer to as “collocation” analysis and “concordance line” analysis.

⁷⁸ Baude & Sachs, *supra* note 68, at 1086 (citing Mark Greenberg, *The Standard Picture and Its Discontents*, in 1 OXFORD STUDIES IN PHILOSOPHY OF LAW 39, 48 (Leslie Green & Brian Leiter eds., 2011)).

⁷⁹ See Jonathan T. Molot, *The Rise and Fall of Textualism*, 106 COLUM. L. REV. 1, 23 (2006).

⁸⁰ See Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 828–33.

Collocation is simply “the tendency of words to be biased in the way they co-occur.”⁸¹ By examining collocation in a corpus, we can identify words that tend to occur in close proximity with terms we are assessing—a collocate is just a word neighbor. This can give us a starting point for an attempt to discern the range of senses of a given word that we may wish to test further.⁸²

Corpora also permit the analysis of concordance lines. A concordance line is a sample sentence from real-world language that shows how a queried word or phrase has been used in the past.⁸³ For a judge or lawyer, this can make the corpus “like Lexis on steroids,”⁸⁴ enabling her—through the use of computer-aided searches—to “assemble a greater number of examples” of a particular word or phrase used in a particular grammatical and topical context than she could “summon by memory.”⁸⁵ By doing so, the judge can check her intuition to ensure that it conforms with the way language is actually used, and not the result of her own idiosyncratic linguistic tendencies.

Justice Stephen Breyer took a step in this direction in his opinion in the *Muscarello* case when he cited some sample sentences from the *New York Times* database in LexisNexis to support his view that we often speak of “carry” in terms of transporting something in a vehicle.⁸⁶ Corpus linguistics takes this impulse a step further by allowing us to use more reliable tools employed by linguists to gauge the most common usage of terms like “carry” a firearm and “harbor” an alien through searches of large databases producing replicable, falsifiable datasets of actual language usage. This is the premise—and promise—of corpus linguistic analysis.

Some have characterized the corpus linguistics project as introducing a new “theory of interpretation.”⁸⁷ Others dismiss it as useful only for textualists or originalists, or as presupposing the

⁸¹ SUSAN HUNSTON, CORPORA IN APPLIED LINGUISTICS 68 (2002); see also John R. Firth, *A Synopsis of Linguistic Theory, 1930–1955*, in STUDIES IN LINGUISTIC ANALYSIS 1, 14 (1957) (“Collocations are actual words in habitual company.”).

⁸² See HUNSTON, *supra* note 81, at 76.

⁸³ See Mouritsen, *Definitional Fallacies*, *supra* note 1, at 1958.

⁸⁴ Brief for the Project on Government Oversight, the Brechner Center for Freedom of Information, and Tax Analysts as Amici Curiae in Support of Petitioners at 14, *FCC v. AT&T, Inc.*, 562 U.S. 397 (2011) (No. 09-1279).

⁸⁵ *State v. Rasabout*, 356 P.3d 1258, 1271 (Utah 2015) (Lee, A.C.J., concurring in part and concurring in the judgment).

⁸⁶ *Muscarello*, 524 U.S. at 129–30.

⁸⁷ Hessick, *supra* note 16, at 1526 n.96.

primacy of these theories.⁸⁸ But these are mistaken premises. Corpus linguistic analysis is not a theory of interpretation; it is a tool for assessing ordinary usage of language in context. Such a tool takes no position on the judge's preferred theory of interpretation. Indeed, the tool may be useful to question some of the premises of textualism and originalism. As noted above, the insights from the corpus linguistics project can offer ammunition to those who seek to question the objectivity or determinacy of the search for ordinary (or original) meaning. To that extent this tool may be quite valuable to the antitextualist, or living constitutionalist—who may use corpus linguistic methods to highlight the indeterminacy of language, and open the door to more pragmatic, purposivist inquiries. Such theories, moreover, typically at least start with a nod to ordinary meaning.⁸⁹ So corpus linguistics is not a theory, and it's hardly just for textualists.

How does corpus linguistic analysis work? The starting point is to answer a series of threshold questions:

(1) *What is the relevant speech community I want to investigate?* Should the judge assume that the legal text she is interpreting is written in ordinary American English? Or is it written in a legal “dialect”?⁹⁰ The answer to this question may depend on whether the judge agrees with Justice Oliver Wendell Holmes, Jr., who argued that “it is reasonable that a fair warning should be given to the world *in language that the common world will understand*, of what the law intends to do if a certain line is passed,”⁹¹ or Justice Stanley Reed who felt that “[i]n the interpretation of statutes, the function of the courts is . . . to construe the language so as to give effect to the intent of Congress”⁹²—a Congress that is

⁸⁸ See, e.g., John S. Ehrett, *Against Corpus Linguistics*, 108 GEO. L.J. ONLINE 50, 72 (2019) (“Judges committed to textualism and originalism may see in this new methodology the potential to resolve longstanding debates over linguistic subjectivity and finally triumph over the demon of arbitrariness. But corpus-based research is a temptation the judiciary should resist.”).

⁸⁹ See, e.g., John F. Manning, *What Divides Textualists from Purposivists?*, 106 COLUM. L. REV. 70, 87 (2006) (“While purposivism is characterized by the conviction that judges should interpret a statute in a way that carries out its reasonably apparent purpose and fulfills its background justification, purposivists start—and most of the time end—their inquiry with the semantic meaning of the text.”).

⁹⁰ See Heilpern, *supra* note 48, at 394.

⁹¹ *McBoyle v. United States*, 283 U.S. 25, 27 (1931) (emphasis added).

⁹² *United States v. Am. Trucking Ass'ns, Inc.*, 310 U.S. 534, 542 (1940).

demographically wealthier, better educated, and contains a greater percentage of lawyers than the nation as a whole.

(2) *What is the relevant time period I want to investigate?* Should the Constitution or statute be interpreted according to the linguistic conventions at the time it was enacted? Or reflect the understanding of people reading it today?

The answers to these questions will determine which corpus a judge should use in a particular case. An original public meaning originalist interpreting the Commerce Clause might want a corpus composed of documents written or read by the general public during the Founding Era. An original methods originalist might want to limit her search to legal documents from that time period. And a living constitutionalist who believes that the Constitution should be interpreted in light of society's evolving standards would be more interested in examples drawn from a modern corpus. The same is true with statutes. Judges who prioritize fair notice in statutory interpretation may want to consult a modern, general corpus—especially for criminal statutes—while those who view themselves as “faithful agents of the legislature”⁹³ may wish for a corpus of congressional speech from the time of enactment, such as all of the legislative history created by a specific Congress. The evidence drawn from different corpora may point to different answers to the same question, but it is the judge's legal theory that makes the difference. Corpus linguistics simply helps judges operationalize their normative values—whatever they may be—in a systematic and scientific manner.

After answering these threshold questions, a judge could assemble concordance line evidence to inform the inquiry into the ordinary meaning of the language of law according to her interpretive theory. We have done so in an attempt to offer some evidence-based answers to the questions that the courts were opining on in *Muscarello*, *Taniguchi*, and *Mont*. The tools that the judges used in those cases weren't up to the task of discerning ordinary meaning. But corpus linguistic analysis can at least begin to fill in some of the gaps.

Our corpus analysis has uncovered linguistic evidence that can make the inquiries in these cases more transparent. For each of the above cases, we were able to assemble objective evidence of

⁹³ Cass R. Sunstein, *Interpreting Statutes in the Regulatory State*, 103 HARV. L. REV. 405, 435 (1989).

ordinary meaning that cannot be derived from the use of dictionaries, etymology, or linguistic canons.

In assessing the problem in *Muscarello*, we determined that the verb “carry” co-occurs with the word “firearm” (or its synonyms) in 109 codable concordance lines in the News on the Web (NOW) Corpus.⁹⁴ Of those lines, 104 refer to personally transport a gun, only 5 to transporting a gun in a vehicle.⁹⁵

As to *Taniguchi*, we found that the noun “interpreter” appears in 188 codable concordance lines in the NOW Corpus. Of those lines, we found none that refer to the text-to-text written translator idea.⁹⁶

We presented our corpus analyses of *Muscarello* and *Taniguchi* in our prior scholarship.⁹⁷ We refer the reader to that work for the details of our analyses. We have not previously considered the *Mont* case, however. And the corpus tools introduced above can also be used to analyze the ordinary meaning question presented there, in a manner highlighting some of the important contributions of these tools.

A threshold question, as noted, concerns the relevant language community. This is a point on which the majority and dissent in *Mont* disagreed—albeit without really acknowledging that this was a focus of their disagreement. The majority cited both general-usage and legal dictionaries in support of its view of “imprison,”⁹⁸ while the dissent looked to usage of this term in the U.S. Code and the Sentencing Reform Act.⁹⁹ So the majority is imagining (at least in part) a general-usage language community, while the dissent has in mind a more specialized community of statutory language.

Corpus tools can provide evidence of the use of “imprison” in either of these communities. We can assess the former by looking to the use of this term in Corpus of Contemporary American English (COCA). A search for the verb “imprison” in COCA generates two hundred codable concordance lines where the object of the verb is a person. Of those, thirty make clear reference to a person being detained after conviction, thirty-three could be

⁹⁴ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 847.

⁹⁵ *Id.*

⁹⁶ *Id.* at 849–50.

⁹⁷ See Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 845–48 (presenting corpus analysis of the *Muscarello* case); *id.* at 848–50 (presenting corpus analysis of *Taniguchi*).

⁹⁸ *Mont*, 139 S. Ct. at 1832.

⁹⁹ See *id.* at 1838 (Sotomayor, J., dissenting).

referring to either pre- or post-conviction, and fifty-five refer to detainment without conviction. Another twenty-five lines were unclear or made reference to some other meaning (like a metaphorical use).

We can also assess use of “imprison” in a statutory setting. We did so by searching use of the term in a corpus of the U.S. Code prepared by linguist Jesse Egbert. Here we found two hundred codable concordance lines that use the verb “imprison” in connection with a person as a direct object. One hundred ninety-five of those lines referred to post-conviction detainment, one could be viewed to refer to either pre- or post-conviction detainment, and none refer to detainment pre-conviction.

This evidence can serve as a check on a single judge’s linguistic intuition. It can also help refine and focus our legal theory on the inquiry into ordinary meaning. Does the corpus evidence also paint a conclusive view of the “standard picture”—the ordinary meaning of the statutory text?

The answer to this crucial question is that it depends—on some refinements that need to be made to our (legal) interpretive theory of the nature of the “standard picture” (or in other words our view of what we are looking for in our search for ordinary meaning). Those refinements will require us to decide whether ordinary meaning encompasses only the most frequent sense of a term in a given context, to all permissible meanings of a term, or perhaps to the term’s prototypical meaning. We must also explore how the law’s search for ordinary meaning fits in the broader structure of the law of interpretation—whether and when we end the interpretive enterprise with the “standard picture,” and how the ordinary meaning inquiry interacts with other premises of the law of interpretation (such as substantive canons of construction).

We offer some answers to these questions in the course of our responses to critics in Parts II–V below. For now, we are focused on the ordinary meaning inquiry. And we are confident in concluding that corpus linguistic analysis can fill in some significant holes in our existing methodologies, even if it can’t resolve all of the problems it highlights with our existing approach to the “standard picture.”

C. The Contributions of Corpus Linguistic Analysis

Corpus linguistics cannot solve all of the problems described above. But its potential contributions are significant.

1. Transparency.

First, corpus linguistic analysis can promote greater transparency in the assessment of ordinary meaning. Without this tool, judicial analysis in this field resides within the black box of judicial intuition, or, worse, under tools that provide a false sense of objectivity and authoritativeness (as under the sense-ranking fallacy as applied to dictionaries, the fallacy of etymology, or the decision to credit an untested linguistic canon). The tools of corpus linguistics can begin to provide checks on these forms of judicial analysis.

Textualists have accused purposivists of abusing legislative history in a manner that amounts to “looking over a crowd and picking out your friends.”¹⁰⁰ Meanwhile, purposivists accuse textualists of doing the same thing with semantic tools and canons in an attempt to justify a preferred outcome.¹⁰¹ The problem of motivated reasoning is a real one. Some empirical work in the *Chevron* domain is a stark illustration. Scholars have shown that a judge’s partisan alignment is among the greatest predictors of whether the judge will defer to an agency under *Chevron*: “Where the [judges’] partisanship aligned with the direction of the agency outcome, those judges were more likely to [find the statute ambiguous and] affirm the agency’s interpretation.”¹⁰² The attempt to judge ambiguity by intuition overlaps substantially with the attempt to assess ordinary meaning. There will always be a risk of motivated reasoning or confirmation bias. The introduction of an evidence-based tool like corpus linguistic analysis can help reduce this risk.

The evidence-based methodology of corpus linguistic analysis can help force a judge to show her work. To begin with, a judge

¹⁰⁰ Patricia M. Wald, *Some Observations on the Use of Legislative History in the 1981 Supreme Court Term*, 68 IOWA L. REV. 195, 214 (1983) (“It sometimes seems that citing legislative history is still, as my late colleague [Judge] Harold Leventhal once observed, akin to ‘looking over a crowd and picking out your friends.’”).

¹⁰¹ See, e.g., Richard J. Pierce, Jr., *The Supreme Court’s New Hypertextualism: An Invitation to Cacophony and Incoherence in the Administrative State*, 95 COLUM. L. REV. 749, 752 (1995); Thomas W. Merrill, *Textualism and the Future of the Chevron Doctrine*, 72 WASH. U. L.Q. 351, 372 (1994).

¹⁰² Kent Barnett, Christina L. Boyd & Christopher J. Walker, *The Politics of Selecting Chevron Deference*, 15 J. EMPIRICAL L. STUD. 597, 601 (2018); see *id.* at 608–09, 614 (finding that this only held true for liberal panels, not conservative ones who applied *Chevron* consistently regardless); Thomas J. Miles & Cass R. Sunstein, *Do Judges Make Regulatory Policy? An Empirical Investigation of Chevron*, 73 U. CHI. L. REV. 823, 851 (2006); Frank B. Cross & Emerson H. Tiller, *Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals*, 107 YALE L.J. 2155, 2175 (1998).

will need to articulate her normative assumptions and values in order to justify her use of a particular corpus and to describe her search parameters. Ideally, a judge would then provide the public with a copy of her coded concordance lines (as exemplified in two amicus briefs submitted to the Supreme Court, which included links to spreadsheets available online¹⁰³) and detailed instructions about how to replicate her search (as one of us has done in a judicial opinion¹⁰⁴). Fellow judges on the panel, reviewing judges on appeal, and legal scholars could check her work to ensure that she has not manipulated the evidence or engaged in motivated reasoning. One of the great virtues of corpus linguistics is that it is falsifiable. Critics can replicate the judge's search to test her conclusions. And if the judge has engaged in cherry-picking or motivated reasoning, a replicated search will reveal this and expose the judge to public criticism.

2. Refining legal theory.

Second, corpus linguistics can help promote some needed and overdue refinements to our legal theory of interpretation. One of the more difficult problems highlighted by the corpus linguistics project is the law's failure to speak precisely about what we are looking for in inquiring into ordinary meaning. Are we looking for the most common sense of a given term or phrase in a given linguistic setting? Or is it enough that a given sense is attested and somewhat frequent? The debate in the *Muscarello* case highlights the law's imprecisions in this regard. In that case the majority opinion says it is seeking the ordinary meaning of "carry" a firearm, but in the Court's analysis it seems to be alternatively either looking for the most common sense of this term or for simply a permissible one.¹⁰⁵

It is tempting to say that this highlights a fatal problem with the attempt to assemble corpus evidence of ordinary meaning. But the problems highlighted above do not go away if we abandon the idea of evaluating claims about meaning with usage evidence. We could keep guessing about ordinary meaning using our

¹⁰³ See Brief of *Amici Curiae* Scholars of Corpus Linguistics at 18 n.11, *Lucia v. SEC*, 138 S. Ct. 2044 (2018) (No. 17-130); Brief of *Amici Curiae* Scholars of Corpus Linguistics Supporting Petitioners at 18 n.21, *Rimini Street, Inc. v. Oracle USA Inc.*, 139 S. Ct. 873 (2019) (No. 17-1625).

¹⁰⁴ See *J.M.W. v. T.I.Z. (In re Adoption of Baby E.Z.)*, 266 P.3d 702, 725 n.26 (Utah 2011) (Lee, J., concurring in part and concurring in the judgment).

¹⁰⁵ *Muscarello*, 524 U.S. at 128.

intuition, or pretend to find objective support for our sense of ordinary meaning using dictionaries or etymology or linguistic canons. Without the evidence, we could mask the fact that we have long been imprecise about what we mean by ordinary meaning. But the question would still remain unanswered.

The question, moreover, is one for legal theory—not linguistics. If we struggle to find an answer to the question whether the law is interested in the most common sense of a given term, simply a common sense, or even a linguistic prototype, then it is because the law hasn't decided what question it is asking when it seeks after ordinary meaning. Linguistic tools can help us measure whatever it is we decide we want to measure. We lawyers just need to decide what we're looking for. And interest in the use of corpus linguistics in legal interpretation has already helped us see the need for this refinement.

3. Testing the tools of statutory interpretation.

Corpus linguistics also provides the means for testing the validity of many of the traditional tools of statutory interpretation. One of the main contributions of corpus linguistics so far has been in highlighting the deficiencies of our current methods of interpretation, such as those discussed in Part I.A. Scholars have been able to show that dictionaries and etymology are poor tools for assessing ordinary meaning. But there is still so much work left to be done. Corpus linguistics provides a means of testing the so-called linguistic canons. Do they actually reflect linguistic usage? And if so, from when?

Further analysis of our linguistic canons is in order. And we can utilize corpus linguistic analysis to determine the extent to which the canons accurately reflect linguistic practices by beginning to identify the circumstances in which the presumptions set forth in the canons should hold and when those presumptions may be rebutted. Corpus analysis of extant canons may also help the legal profession identify new presumptions not yet reflected in our canons—a new set of canons, which are supported by linguistic evidence but not yet reflected in our law.

* * *

The law and corpus linguistics project is still in its infancy. But it has much to offer. Already it has helped highlight the deficiencies of the methods currently used by judges to discover the communicative content of legal text. If judges are serious about

our search for ordinary meaning, they should heartily embrace this new methodology, regardless of their jurisprudential leanings or views of the nature of legal interpretation. Doing so will improve transparency in judicial analysis and promote refinements in both our legal theory and canons of interpretation.

II. CORPUS LINGUISTICS AND THE SPEECH COMMUNITY

One critique of the use of corpus tools concerns the selection of the relevant language community. Professor Evan Zoldan, for example, has asserted that the language of statutes differs in potentially important ways from the language of general corpora like the COCA, the NOW Corpus, and the Corpus of Historical American English (COHA).¹⁰⁶ Zoldan suggests that “statutory language and the language of texts found in a general corpus have different purposes, audiences, and other linguistic characteristics.”¹⁰⁷ And from that premise he concludes that “[i]t is never appropriate to search for statutory meaning in a general corpus.”¹⁰⁸

Zoldan’s points are echoed to some degree by Professor Anya Bernstein, who suggests the possible need to identify specialized speech communities to study—such as the community “directly affected or addressed by [a] statute.”¹⁰⁹ Bernstein posits that the relevant language community for the cost-shifting statute at issue in *Taniguchi* may not be professional translators and interpreters but the judges who are charged with shifting costs.¹¹⁰ And she asserts that African Americans, who are “dramatically overrepresented in drug arrests, convictions, and incarceration,” may be the relevant speech community for “drug-related criminal and sentencing statutes.”¹¹¹

We agree with elements of the above premises. But we think the criticism is substantially overstated. And we think that the arguments about language community ultimately reinforce rather than undermine the need for and utility of corpus linguistic analysis.

¹⁰⁶ See Zoldan, *supra* note 15, at 423–41.

¹⁰⁷ *Id.* at 426.

¹⁰⁸ *Id.* at 438.

¹⁰⁹ See Bernstein, *supra* note 15, at 459.

¹¹⁰ *Id.* (citing *Taniguchi*, 566 U.S. at 576–77 (Ginsburg, J., dissenting)).

¹¹¹ *Id.*

A. The Purpose of Legal Language

We agree that “statutory language serves different purposes than the language found in the texts of a general corpus.”¹¹² The fact that the former is prescriptive—in “prescrib[ing] modes of conduct, grant[ing] rights, [and] impos[ing] obligations”—could highlight a significant point of distinction.¹¹³ It may be that the prescriptive nature of legal language could sometimes call into question the probity of evidence from a general corpus. Perhaps we use the terminology of “carry” a firearm differently when we are prohibiting it in a criminal law establishing a mandatory minimum sentence than we do when we are merely speaking descriptively. And if so, that could call into question the utility of a general corpus for assessing the meaning of the language of this statute.

Yet that premise does not render corpus linguistic analysis untenable. Zoldan and Bernstein offer no alternative mechanism for assessing the communicative content of the language of law. This project, as noted, is as much about underscoring inadequacies in our existing theories and tools as it is in proposing a range of new tools. Critics questioning the utility of corpus linguistic analysis should choose a place on the spectrum of key questions we raised in Part I—whether they are proposing to abandon the search for the “standard picture” altogether, to stick with now-debunked tools for painting that picture (sense-ranking fallacy, fallacy of etymology, etc.), or to help in the quest for more reliable tools.¹¹⁴ Zoldan and Bernstein offer no alternatives. They only suggest that a general corpus may not be perfect. That may be true, but it is not enough to justify either the continued use of more problematic tools or the abandonment of the corpus linguistics project altogether.

That leads to a second response to Zoldan’s point about the purpose of statutory language: If and when his premise is correct, it may be tested using the very methods that Zoldan criticizes. There is nothing about corpus linguistic analysis that limits its use to a general corpus. BYU Law is in the initial stages of developing a corpus of legislative history. Such a corpus might allow us to test the premise of Zoldan’s argument. We could look to see how “carry” a firearm is used in legal language. And we could

¹¹² Zoldan, *supra* note 15, at 426.

¹¹³ *Id.* at 426 & n.158 (third alteration in original) (quoting Andrei Marmor, *The Pragmatics of Legal Language*, 21 *RATIO JURIS* 423, 425 (2008)).

¹¹⁴ *See supra* Part I.A.

assemble evidence of whether the use of this term in this context differs in any meaningful way from its use in a general corpus. Such analysis, moreover, would give us evidence that could never be derived from current methods.

B. The Audience of Legal Language

Zoldan also suggests that legal language is aimed at a different audience. He says that “[t]he audience of statutes always includes experts who interpret statutory language in their official capacity.”¹¹⁵ And he notes that general corpora like COCA are measuring informal language aimed at a different audience, in that they include transcripts of unscripted conversation from TV and radio programs, short stories and plays, etc.¹¹⁶ This premise is not wrong. But it is overstated. And, again, it fails to undermine the corpus linguistics project because Zoldan offers no better way to assess the communicative content of the language of law.

A key problem with this critique is that it sweeps too broadly. It is surely true that the audience of a statute “always *includes* public officials, subject-matter experts, lawyers, and judges” who will have specialized training about the language of law.¹¹⁷ But that doesn’t tell us that this is the *only* audience at which our laws are aimed. Much of the criminal and regulatory law of state and local government is aimed at individuals. This is Bernstein’s point. In the criminal law we assume that people read the law and are entitled to regulate their behavior in accordance with its terms.¹¹⁸ This may often be a fiction. But it advances a core function of law—a notice function, which is aimed at avoiding unfair surprise, or upsetting settled reliance interests.¹¹⁹

This helps to highlight a central point we have been making all along—that many of the problems we encounter in the law and corpus linguistics movement are rooted in imprecisions in legal theory, not in problems with corpus linguistics. What Zoldan has done is to take a position on an unsettled question of legal theory—inadvertently emphasizing the need to be clear about the nature of meaning that the law is trying to measure. There may be

¹¹⁵ Zoldan, *supra* note 15, at 430.

¹¹⁶ *See id.*

¹¹⁷ *Id.* (emphasis added).

¹¹⁸ *See* *McBoyle v. United States*, 283 U.S. 25, 27 (1931) (Holmes, J.) (“[I]t is reasonable that a fair warning should be given to the world in language that the common world will understand, of what the law intends to do if a certain line is passed.”).

¹¹⁹ *See* Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 818 & n.131.

reasons (rooted in the goals of advancing the coherence of the broader fabric of the law) for crediting specialized legal meaning in the words of our law. But there are also strong arguments in favor of a search for ordinary meaning. At least sometimes, moreover, the concepts will overlap. Zoldan has failed to identify any examples of cases in which he thinks the audience problem would distort the utility of corpus linguistic analysis. We see no reason to doubt the utility of our corpus analysis of “carry” a firearm, or “harbor” an alien, especially since these cases involve criminal laws where concerns for fair notice call for application of the rule of lenity. And Zoldan has offered no better way to assess the meaning of language within the specialized community that he imagines.

Again, this highlights a feature—not a bug—of corpus linguistic analysis. To the extent the language of law is specialized, we can build a specialized corpus to assess its meaning. If “carry” a firearm has a distinct meaning in the dialect of law, we can build a specialized legal corpus to measure it. So the critique ultimately comes back as a ground for endorsing (and of course refining) the corpus linguistic project, not abandoning it.

C. The Usage and Syntax of Legal Language

Zoldan’s last point is to assert that legal language may differ from ordinary language in terms of word usage and syntax. He says that “differences in syntax and drafting conventions [] ‘render[] legislative texts incomprehensible to all except the specialist reader’”¹²⁰ and that “[legal texts] contain ‘word usages that have no parallel in ordinary conversation.’”¹²¹

Zoldan may be right that a general corpus search could be unhelpful as to these sorts of phrases and constructions. If a word or phrase is uncommon or unattested in a general corpus, then of course a corpus search will not be helpful. But once again, Zoldan has cited no examples of this in the case studies we have presented. And it doesn’t undermine the general utility of corpus analysis to suggest that it will not always be useful. Our principal responses to this critique are made above, but we remind the reader that our existing tools for assessing ordinary meaning are

¹²⁰ Zoldan, *supra* note 15, at 436 (quoting Yon Maley, *The Language of Law*, in LANGUAGE AND THE LAW 11, 25 (John Gibbons ed., 1994)).

¹²¹ *Id.* at 435 (quoting Paul E. McGreal, *Slighting Context: On the Illogic of Ordinary Speech in Statutory Interpretation*, 52 U. KAN. L. REV. 325, 326 (2004)) (giving as examples “interplead” and “demurrer”).

even more problematic and that we can construct corpora to assess the specialized language communities that Zoldan is interested in.

Corpus searches—especially in large corpora—allow the user to find many examples of comparatively infrequent linguistic phenomena, like “carry” as a verb within so many words of some type of firearm or gun. Questions like the one raised in *Muscarello* are not concerned with the general syntax of the law, but with the particular syntax of the relevant statute. And in this case, the NOW Corpus provided many examples of language containing the relevant information, including the relevant syntax, of the statute. Because these corpora range from hundreds of millions of words to several billion words, they often provide more than enough evidence, even if the syntax of a particular clause differs from how most people speak and write.

III. CORPUS LINGUISTICS AND NOTICE

Another objection to corpus linguistic analysis of the language of the law concerns whether the tool gives fair notice to litigants and legislatures who don’t have access to or knowledge of corpora and wouldn’t know how to use them if they did. The critics seem to have two forms of notice in mind: notice of what the law requires and notice of the methods used by judges to decide their cases.

Professor Hessick observes that our criminal law in particular “must give us fair notice of what conduct is permitted and what conduct is prohibited.”¹²² And because the average person “cannot be expected to perform their own corpus searches and analyses,” Hessick asserts that the public will not have fair notice.¹²³ With that concern in mind, she warns that a person subject to criminal liability based on corpus linguistic analysis of a criminal statute “may accidentally engage in illegal conduct,” or “may choose to avoid large swaths of legal conduct because she does not know for certain whether such conduct is permitted.”¹²⁴

Hessick starts from valid premises—about the importance of fair notice, and the inability of defendants to perform corpus linguistic analysis on their own. But those premises are hardly

¹²² Hessick, *supra* note 16, at 1514.

¹²³ *Id.* at 1515.

¹²⁴ *Id.* at 1514–15 (first citing *Grayned v. City of Rockford*, 408 U.S. 104, 108 (1972); and then citing *United States v. Nat’l Dairy Prods. Corp.*, 372 U.S. 29, 36 (1963)).

enough to undermine the use of this set of tools. Indeed, concerns about fair notice cut against maintaining the status quo and strongly in favor of the addition of corpus tools to the judge's existing toolbox.

The concern about fair notice underpins our attempt to discern the meaning of the language of the law.¹²⁵ In fact, this concern drives most any interpretive inquiry—whether textualist, purposivist, pragmatic, or otherwise—to begin with an attempt to discern the ordinary or plain meaning of legal language.¹²⁶ We start (and often end) with that inquiry because we think that people are entitled to fair notice of the legal consequences of their behavior—and because we believe that our shared conventions about language usage lend more determinacy and predictability to the inquiry into ordinary meaning than we will find in other interpretive inquiries.¹²⁷

Hessick is also right that most people don't know how to do corpus linguistic analysis. But it doesn't follow that the tools of this field cannot be used to assess the ordinary meaning of the language of law. The viability of these tools must turn on whether they improve the accuracy of the judge's assessment of that meaning. Our courts use a broad range of tools (canons, legislative history, etc.) that the general public lacks the expertise necessary to employ on its own. And the viability of those tools has always turned on whether they accurately assess the public behavior they purport to measure—not on whether the public has the expertise necessary to employ such tools on their own.

To flesh out this point it may be helpful to distinguish two categories of factual inquiry performed in our courts. One such inquiry concerns adjudicative facts. These are facts about the real-world dispute between the parties to a given case—facts about what/where/when/how/why that require courts to decide which side's view is deemed correct.¹²⁸ We would not put up with a criminal justice system that subjects people to criminal (or even civil) liability based on factual inquiries that turn on standards or methods that are opaque, hidden, or inaccessible to the parties. Transparency and fair notice are crucial. Yet the transparency

¹²⁵ See Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 818.

¹²⁶ See ESKRIDGE, *supra* note 27, at 34–35.

¹²⁷ See Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 792–95.

¹²⁸ As to these facts, judges concededly are beholden to the evidence presented by the parties; sua sponte judicial inquiries into adjudicative facts are prohibited. See *State v. Rasabout*, 356 P.3d 1258, 1284 (Utah 2015) (Lee, A.C.J., concurring in part and concurring in the judgment); Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 868–71.

and notice that we require does not turn on a layperson's ability to reproduce the means of proving a particular fact; it just turns on whether those means are shown to be sufficiently reliable in proving the underlying fact.¹²⁹ Where means of proof are inaccessible to parties directly, we expect them to turn to and rely on experts who can inform them and represent their interests in court—their counsel and the expert witnesses that their counsel may employ.

Consider a criminal price-fixing case under antitrust law,¹³⁰ or a criminal counterfeiting case under trademark law.¹³¹ A defendant in these cases is entitled to fair notice of the terms and conditions for proving the facts necessary to establish the elements of the charged crimes. And the law of evidence will dictate whether a given means of proving those elements is sufficiently reliable to be admitted into evidence. But we do not require the defendant himself to be capable of understanding or replicating the methodology. Instead, we expect him to rely on his own experts. In the criminal antitrust case example, the prosecution may put on expert testimony from an economist, presenting a multiple regression analysis supporting an allegation that the defendant exercised market power in setting prices. The defendant may not be in a position to replicate or even understand the economist's regression analysis. Yet that would not foreclose the admissibility of this testimony. That would rise or fall on the competency and reliability of the expert analysis;¹³² and the defendant could retain his own expert to challenge it.

The same goes for the trademark case. Here we could expect the prosecution to call a marketing or consumer behavior expert to present a consumer survey supporting the allegation that the defendant's trademark creates a likelihood of confusion with a senior trademark. Again, the defendant would likely have no capacity to replicate this sort of evidence. And again, the

¹²⁹ See *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 593–94 (1993) (establishing admissibility factors for scientific testimony); *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 147 (1999) (applying *Daubert* to all expert testimony under Federal Rule of Evidence 702).

¹³⁰ See 15 U.S.C. § 1 (making illegal “[e]very contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States, or with foreign nations”).

¹³¹ See 18 U.S.C. § 2320(a) (defining federal criminal trademark counterfeiting offenses).

¹³² See, e.g., *In re Processed Egg Prods. Antitrust Litig.*, 81 F. Supp. 3d 412, 434 (E.D. Pa. 2015) (“The Court finds Dr. Rausser’s regression sufficiently reliable for purposes of *Daubert*. . . . Plaintiffs persuasively note that Dr. Myslinski’s methods for testing Dr. Rausser’s regression model hardly demonstrate that the model is, as a whole, unreliable for demonstrating an expected overcharge.”).

admissibility of the evidence would turn not on that concern but on whether the expert is competent and the proposed survey is reliable.¹³³ The defendant would of course also have access to his own expert if he wishes to challenge the prosecution's survey or present his own countersurvey.

Corpus linguistic analysis goes to a different sort of factual inquiry—to the relevant legislative facts of the case. These are facts that inform the court's understanding of the law that applies to the resolution of the case.¹³⁴ In a common law case, the legislative facts may concern psychological, economic, or sociological materials that inform the court's decision as to the optimal rule to adopt in a particular field.¹³⁵ In a constitutional case, the legislative facts may encompass a historical inquiry into the original meaning of a term or clause of the Constitution.¹³⁶ On these sorts of facts we allow and even encourage our judges to perform their own inquiries. Judges are not limited to the legislative factual material submitted by the parties or their counsel.¹³⁷ They may perform their own inquiries into relevant social science literature to determine the best common law rule, just as they may do their own historical research on the original meaning of the

¹³³ See 6 J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 32:182 (5th ed. 2020).

¹³⁴ See *Rasabout*, 356 P.3d at 1284 & n.36 (Lee, A.C.J., concurring in part and concurring in the judgment); see also *Bulova Watch Co. v. K. Hattori & Co.*, 508 F. Supp. 1322, 1328 (E.D.N.Y. 1981) (explaining that the “court’s power to resort to less well known and accepted sources of data to fill in the gaps of its knowledge for legislative and general evidential hypothesis purposes must be accepted because it is essential to the judicial process”); Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 868–71 (discussing how corpus linguistic tools are used to assess and determine legislative, rather than adjudicative facts); ROBERT E. KEETON, *JUDGING* 38–39 (1990) (clarifying the distinction between legislative and adjudicative facts).

¹³⁵ See, e.g., Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 871 n.316 (listing examples); *Rasabout*, 356 P.3d at 1276–77, 1277 nn.14–16 (Lee, A.C.J., concurring in part and concurring in the judgment) (discussing examples).

¹³⁶ See *Lucia v. SEC*, 138 S. Ct. 2044, 2056–57 (2018) (Thomas, J., concurring) (interpreting “Officers of the United States”); see also Lee & Phillips, *supra* note 8, at 264–71 (describing several examples).

¹³⁷ See *Rasabout*, 356 P.3d at 1284 (Lee, A.C.J., concurring in part and concurring in the judgment); see also Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 869–70; Kenneth Culp Davis, *An Approach to Problems of Evidence in the Administrative Process*, 55 HARV. L. REV. 364, 402–03 (1942):

When an agency wrestles with a question of law or policy, it is acting legislatively, just as judges have created the common law through judicial legislation, and the facts which inform its legislative judgment may conveniently be denominated legislative facts. The distinction is important; the traditional rules of evidence are designed for adjudicative facts, and unnecessary confusion results from attempting to apply the traditional rules to legislative facts.

Constitution. The same goes for our judges' inquiries into the ordinary meaning of the language of the law. They are not stuck with the dictionaries or other materials cited by the parties; we expect them to use their best lights and the most reliable tools to assess the meaning of legal language.¹³⁸

This is not to say that "anything goes" when it comes to the inquiry into legislative facts. We expect both parties and judges alike to utilize competent methods and tools and to present reliable, probative evidence in support of their positions on these facts. And although a judge is entitled to perform his own inquiry and present his own evidence, most judges will prefer to have adversary briefing before relying on a new line of inquiry (such as corpus linguistic analysis).¹³⁹ Ultimately, however, we judge the competence of the methods and tools and the reliability of the evidence on the basis of their ability to accurately inform the court's determination of the legislative fact in question—and not on whether a party to the case could himself replicate the process leading to the court's determination. As with adjudicative facts, we understand that a lay party may lack the knowledge or expertise to perform a given factual inquiry; but we fall back on the assurance that the inquiry itself is reliable and that the defendant can retain his own expert to perform his own analysis of the issue.

Average criminal defendants may lack knowledge of the social science methods or literature of relevance to the court's determination of the appropriate common law standard or defense to apply in a given case, just as they may be out of their depth on the tools or methods for assessing the original meaning of the Double Jeopardy Clause or the Ex Post Facto Clause in a case where those provisions may provide a defense. But we do not foreclose a court from employing these tools for determining the content of the law on this basis. We allow defendants to acquire the expertise they may need through counsel and expert witnesses.

For all these reasons, the viability of corpus linguistic analysis of the language of the law must turn on the reliability of these tools for assessing ordinary meaning. And a core premise of the corpus linguistics movement in law is the notion that our

¹³⁸ See *Rasabout*, 356 P.3d at 1276–77 (Lee, A.C.J., concurring in part and concurring in the judgment); see also Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 869–71.

¹³⁹ See *Rasabout*, 356 P.3d at 1283 (Lee, A.C.J., concurring in part and concurring in the judgment) ("Our opinions are better when adversary briefing is complete and in-depth.").

traditional tools for assessing ordinary meaning—judicial intuition, dictionaries, etymology, and canons—are demonstrably inadequate. The first two of these listed tools are usually just smoke screens. Judges cite them as dispositive, but they usually don't give a reliable indication of which of two proffered senses of a statutory term is more ordinary.¹⁴⁰ Hessick and other critics do not address with this problem. They certainly don't defend the reliability of these tools. And without an answer to this problem, they are effectively arguing in favor of judicial reliance on tools that are clearly incapable of providing a reliable answer to the ordinary meaning inquiry. If notice is the concern, this prospect ought to be much worse. One of the motivating concerns behind the corpus linguistics project is the need for greater transparency, and the avoidance of tools that can mask and facilitate confirmation bias and motivated reasoning.¹⁴¹ When judges are allowed to justify their answers to questions about legislative facts with malleable tools that don't really provide an answer, they are certain to undermine the goal of fair notice. Corpus linguistic tools are aimed at minimizing this risk. To the extent they accurately inform the inquiry into the ordinary meaning of the language of the law, they enhance rather than undermine the goal of promoting fair notice to the public. And for reasons explored above and in our prior writings we are confident that these tools do in fact improve the accuracy of this judicial inquiry.

That leaves the use of judicial intuition. We are quite in favor of the use of this tool. We just think that intuition about linguistic facts, unchecked by evidence, runs the risks (if not the guarantee) of confirmation bias and motivated reasoning. And again, these

¹⁴⁰ See Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 808–09 (arguing that relying on the ordering of definitions in a dictionary to determine the sense of the word used in a given context is often problematic because “[t]he dictionaries typically cited by our courts . . . make no claims about the relative frequency of the listed senses of a given word”); *id.* at 809–10 (describing the “etymological fallacy” as the assumption that a word’s etymology reveals the word’s true meaning in a given context (quoting HENRY SWEET, *THE PRACTICAL STUDY OF LANGUAGES: A GUIDE FOR TEACHERS AND LEARNERS* 88 (1900))); Lee & Phillips, *supra* note 8, at 283 (noting that “because the human brain understands words not in isolation but in their broader semantic (and pragmatic) context, we may often miss the import of a given constitutional term if we just separately look up its component words in the dictionary”).

¹⁴¹ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 877:

At a minimum, the data that can be compiled through corpus linguistic analysis will allow lawyers and judges to have a transparent debate informed by real data instead of inferences from sources (like dictionaries or etymology or intuition) that are both opaque and ill-suited to the task to which they are applied.

are the very risks that a proponent of fair notice ought to be interested in avoiding.

On this point (as with every other argument addressed herein) our critics fall short in their failure to back up their criticism with a defense of the status quo. A central part of the law and corpus linguistic thesis is its challenge to the viability of traditional tools for assessing ordinary meaning. And, unless our critics can show that the status quo is better at protecting their concerns about fair notice, they have failed to counter a central element of the project. Indeed, if fair notice about judicial tools for assessing ordinary meaning is the concern, then the status quo is the approach that ought to cause the most heartburn. A party to a judicial proceeding may not be able to anticipate the linguistic intuitions of the individual judge or panel of judges that will decide their case. So if the inaccessibility of a decisive interpretive methodology is a sticking point on fair notice grounds, then the bare use of judicial intuition (unchecked by any evidence) is the worst approach we could possibly imagine.

Hessick misstates the premises of the corpus linguistic inquiry. She accuses us of “[s]eeking to prevent judges from relying on their own judgment in statutory interpretation,” in a manner that ultimately “reject[s] [] the judicial power to interpret the laws.”¹⁴² Professor Bernstein advances a similar charge. She says that corpus linguistic analysis improperly “outsourc[es] interpretation,” giving the “impression of certainty” but “actually undermin[ing] . . . the discretion we thrust on [judges].”¹⁴³ Both of these criticisms are mistaken. As discussed in more detail below, the corpus linguistics project has never been about eliminating judicial discretion. The only discretion we wish to limit is the discretion to make claims about language usage based on the single data point of a judge’s “take my word for it” assurance—an assurance unsupported by (and at times directly contradicted by) evidence from a tool specifically designed to measure language usage.

We do not disagree with Bernstein’s observations about the discretion that remains in a judge’s use of corpus linguistic evidence. To use such evidence reliably and appropriately, a judge must be informed of the premises of this corpus linguistic inquiry and exercise the judgment necessary to separate reliable corpus linguistic analysis from junk science. We concede that more work

¹⁴² Hessick, *supra* note 16, at 1512.

¹⁴³ Bernstein, *supra* note 15, at 439.

needs to be done to refine the proper terms and conditions for corpus analysis of the language of the law. But that concession does not undermine the basis for the project.

IV. SURVEY-BASED CRITIQUES OF THE ACCURACY OF CORPUS LINGUISTIC ANALYSIS

Recent criticisms of corpus linguistic analysis also question whether corpus tools yield an accurate measure of the ordinary understanding of statutory language. A prominent example of this work is Professor Tobia's recent article in the *Harvard Law Review*.¹⁴⁴ Tobia tests the performance of survey respondents on a series of interpretive tasks. Some of the respondents were asked to respond based on their own intuitive concept of a given question. Others were either given dictionary definitions or corpus evidence.¹⁴⁵ Because respondents who made concept-based judgments about the scope of tested terms responded differently than those who were given dictionary definitions and corpus evidence, Tobia concludes that the corpus-based (and dictionary-based) analysis must be in "error"¹⁴⁶ and "inconsistent with ordinary meaning."¹⁴⁷

Tobia states that "dictionary use and legal corpus linguistics carry serious risks of diverging from ordinary understanding."¹⁴⁸ He also concludes that the results of his experiments "shift the argumentative burden to theorists and practitioners who rely on these tools to determine legal outcomes: in light of the data, these views must articulate and demonstrate a nonarbitrary and reliable method of interpretation."¹⁴⁹

Tobia's thesis is premised on a series of surveys administered to judges, law students, and, primarily, respondents recruited from Amazon's Mechanical Turk.¹⁵⁰ The surveys addressed the meaning of several statutory terms. But here we will focus on Tobia's surveys of the scope of the term "vehicle."

Tobia's experiment proceeded in four parts. In the first part, he conducted surveys assessing the meaning of the term "vehicle" under three separate conditions—a concept condition, a

¹⁴⁴ See generally Tobia, *supra* note 17.

¹⁴⁵ *Id.* at 754–56.

¹⁴⁶ *Id.* at 772–74.

¹⁴⁷ *Id.* at 749 tbl.1, 772.

¹⁴⁸ *Id.* at 805.

¹⁴⁹ Tobia, *supra* note 17, at 805.

¹⁵⁰ *Id.* at 754, 763, 765.

dictionary condition, and a corpus condition.¹⁵¹ In the concept condition, Tobia's respondents were asked to respond to a series of questions about whether a given noun (car, bus, truck, bicycle, airplane, toy car, etc.) is properly classified as a "vehicle."¹⁵²

Tobia's dictionary and corpus condition surveys took a somewhat different approach. Instead of asking survey respondents to say whether a given noun is properly classified as a "vehicle," Tobia introduced a dummy replacement word ("ailac"), asserting that the use of such a placeholder would prevent respondents from having any associations with the term "vehicle" that would interfere with their use of the dictionary and ensure that the responses were attributable to the dictionary prompts, rather than their "conceptual competence concerning vehicles."¹⁵³

In the dictionary condition, participants were given a set of dictionary definitions defining "ailac" as "1) a means of carrying or transporting something" and "2) an agent of transmission: carrier."¹⁵⁴ Respondents were then asked to characterize ten items (vehicle, automobile, car, bus, truck, bicycle, airplane, ambulance, golf cart, and toy car), responding to question prompts like: "Is a car an ailac?" [Yes / No].¹⁵⁵

In the corpus condition, Tobia attempted to model the use of linguistic corpora in legal interpretation. Participants in the corpus condition were prompted to "[c]onsider the noun, 'ailac'" and given "some information" about the term to help them "understand" it.¹⁵⁶ First, he gave respondents "the top common words used in connection with 'ailac'"—words that "might appear before

¹⁵¹ *Id.* at 754–56.

¹⁵² *Id.* at 754–55, 757 fig.1. Tobia speaks of a range of examples for this term in the language of law: an insurance policy covering all vehicles owned by an insured, a state statute requiring registration of all vehicles, and a local ordinance prohibiting vehicles in public parks. *See id.* at 739–40. But despite acknowledging the different context that each of these applications would introduce, Tobia does not seek to consider the context differences in any of his surveys. We take this point up further below.

¹⁵³ *Id.* at 755–56.

¹⁵⁴ Tobia, *supra* note 17, at 755. Notably, Tobia gave his survey respondents broad, extensive definitions of "vehicle"—the notions of "a means of carrying or transporting something" and of "an agent of transmission: carrier." He also claims that "this dictionary definition" is the one we "suggested" in our article. *Id.* at 755 n.134. This is incorrect. We identified a much wider range of definitions, including a more limited one referring to an automobile. *See* Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 800–02. And Tobia's choice of definition skewed his results in a predictable way that renders his survey uninteresting and his results unhelpful. We discuss this problem further below.

¹⁵⁵ Tobia, *supra* note 17, at 756 (brackets in original).

¹⁵⁶ *Id.* at 755.

or after ailac, or sometimes close to ailac.”¹⁵⁷ He then showed respondents a collection of approximately fifty collocates or “[t]op common words” purportedly associated with “ailac,” though they were in fact collocates of “vehicle.”¹⁵⁸

Next, Tobia showed respondents a collection of nine concordance lines featuring the word “ailac.”¹⁵⁹ Tobia characterizes these survey prompts as “*precisely* what recent advocates of legal corpus linguistics recommend.”¹⁶⁰ As with the dictionary condition, respondents were asked to characterize ten items responding to question prompts like: “Is a car an ailac?” [Yes / No].”¹⁶¹

In a second experiment, Tobia administered a different survey to respondents recruited from Amazon’s Mechanical Turk. In this survey, Tobia explained to respondents “the difference between prototypical and nonprototypical category members,” and then asked them to answer the following prompts for the same ten items as in the first experiment:

An airplane is a prototypical vehicle. 1 (strongly disagree) to 7 (strongly agree)

An airplane is technically a vehicle. 1 (strongly disagree) to 7 (strongly agree).¹⁶²

In light of the results of his first and second experiments, Tobia concludes, among other things, that “corpus linguistics elicits more prototypical uses of a term[,] but dictionaries elicit more extensive uses.”¹⁶³

In a third experiment, Tobia repeated the first experiment with minor changes,¹⁶⁴ only this time his respondents were a collection of ninety-six state and federal judges who agreed to

¹⁵⁷ *Id.*

¹⁵⁸ *Id.* These were: “electric, motor, plug-in, unmanned, armored, connected, cars, aerial, charging, pure, launch, owners, hybrid, traffic, fuel, driving, gas, autonomous, struck, operating, road, safety, accidents, battery, ownership, emergency, batteries, emissions, seat, advanced, driver, primary, demand, commandeered, fuel-efficient, automakers, demonstrators, excluding, lunar, passenger, fleet, gasoline, luxury, drove, parking, retirement, [and] infrastructure.” *Id.* at 755–56 (emphasis omitted).

¹⁵⁹ *Id.* at 756 & n.137.

¹⁶⁰ Tobia, *supra* note 17, at 756 (emphasis in original) (citing Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 840–42).

¹⁶¹ *Id.* (brackets in original).

¹⁶² *See id.* at 759.

¹⁶³ *Id.* at 761.

¹⁶⁴ Here, Tobia asked about twenty-five entities instead of ten entities, including some that he predicted would not be viewed as “vehicles.” *See id.* at 762.

respond to a survey.¹⁶⁵ He finds that the responses of state and federal judges are “strikingly similar” to those of “nonexpert[]” respondents, and also that the “judges’ use of legal corpus linguistics and dictionary methods did not consistently reflect their ordinary judgments about category membership.”¹⁶⁶

Lastly, Tobia conducted a fourth experiment in which he administered surveys similar to those in the first experiment, but this time assessing “vehicle” and nine additional terms.¹⁶⁷ The additional tested terms were: carry, interpreter, labor, tangible object, weapon, animal, clothing, food, and furniture.¹⁶⁸

Tobia says that the data from these experiments show “significant differences among Dictionary, Legal Corpus Linguistics, and Concept conditions.”¹⁶⁹ He questions the use of both dictionaries and corpus tools for assessing ordinary meaning, asserting that both approaches yield responses that are inconsistent with the results of his concept condition surveys.¹⁷⁰ And he thus asserts that both are problematic tools for accurately assessing ordinary meaning.¹⁷¹

Tobia is particularly critical of corpus tools. He says that “[b]roadly speaking, dictionary use was fairly consistent with people’s ordinary judgments,” in that “cars, buses, and trucks are vehicles, but a toy car definitely is not.”¹⁷² Yet he concludes that “corpus linguistics did not perform nearly as well,” asserting that “[a] bus is seemingly within our modern conception of a vehicle, but *only half* of the users of legal corpus linguistics made that categorization.”¹⁷³ Later, however, Tobia also raises concerns

¹⁶⁵ See Tobia, *supra* note 17, at 763. Tobia also administered the first survey to law students at Harvard, Columbia, and Yale. See *id.* at 762. The results are presented in Appendix C to Tobia’s article and included in some charts, but not discussed in detail in the main body. See *id.*; see, e.g., *id.* at 766 fig.5. This Article will primarily address the discussion in the main body of Tobia’s piece.

¹⁶⁶ *Id.* at 764.

¹⁶⁷ *Id.*

¹⁶⁸ See *id.* As in the third experiment, Tobia asked survey respondents about twenty-five separate entities for each term. In addition, rather than look for definitions in the dictionaries cited in Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, as in the first experiment, Tobia relied on the first full definition of the relevant term from *Merriam-Webster.com*. See *id.* at 764–65. It is not clear why Tobia elected to use only the first definition. Here again, Tobia finds a large degree of divergence in responses to his Concept, Dictionary, and Corpus conditions. *Id.* at 765.

¹⁶⁹ Tobia, *supra* note 17, at 756; see *id.* at 764–65.

¹⁷⁰ See *id.* at 772–73.

¹⁷¹ See *id.*

¹⁷² *Id.* at 757.

¹⁷³ *Id.* (emphasis in original).

about dictionary usage. He asserts that “in some cases the dictionary use indicated that clear nonvehicles were, in fact, vehicles”—noting that “dictionary-using judges overrated roller skates and baby-shoulder carriers as vehicles, compared to judges’ ordinary evaluation of those entities.”¹⁷⁴

Tobia makes some provocative points. We are open to the possibility that survey tools may hold some promise in introducing an additional set of empirical data for assessing the ordinary meaning of statutory text. And we welcome Tobia’s contribution to the dialogue about how best to assemble empirical data in this budding field. That said, we see flaws in the methods and premises of Tobia’s analysis. So, although we think he raises some crucial questions that merit further exploration, we don’t think he delivers defensible answers to his own questions. And we don’t think his surveys establish the ground truth by which the “accuracy” of corpus tools for discerning the ordinary meaning of legal language can be measured.

Our response proceeds in two main parts. In Part IV.A, we address the threshold premise of Tobia’s accuracy claim—his assertion that the baseline truth of ordinary meaning is established by the results of his concept condition survey. We identify a range of reasons to question this premise, highlighting shortcomings in Tobia’s survey design and in survey methods generally. In Part IV.B, we respond to the specific charge that corpus measures are biased in a particular way. Again, we highlight problems in Tobia’s measures. But we also show how the careful corpus methods that we have advocated can account for the concerns that Tobia has raised.

A. Tobia’s Concept Condition Results Do Not Establish Baseline Truth

Tobia assumes, without establishing, that his concept condition survey results are an accurate depiction of human perception of ordinary meaning. But his assumption falters on each of three grounds: (1) Tobia adopts an implicit preference for linguistic competence over linguistic performance, even though his choice introduces a range of concerns for which he has not accounted. (2) He ignores the nuance and complexity embedded in the legal notion of ordinary meaning, which renders his data unhelpful. (3) Human perception of language is informed by both syntactic

¹⁷⁴ Tobia, *supra* note 17, at 773.

and pragmatic context but is excluded from Tobia's oversimplified concept condition experiment (while such context may be incorporated into corpus linguistic analysis).

1. Survey methods do not give privileged access to linguistic "competence."

A threshold problem with Tobia's survey analysis is definitional. Tobia repeatedly uses the phrases "ordinary understanding" and "ordinary judgments."¹⁷⁵ He never defines these terms.¹⁷⁶ But he seems to equate them with the results of his concept condition surveys.¹⁷⁷ So when Tobia says that "both legal corpus linguistics and dictionary use diverged from [the] ordinary meaning" of the tested term or phrase,¹⁷⁸ all he is saying is that the results of his dictionary- and corpus-based surveys diverged from the results of his concept-based survey.¹⁷⁹ And when he states that the "judges' use of legal corpus linguistics and dictionary methods did not consistently reflect their ordinary judgments about category membership,"¹⁸⁰ all he is saying is that the results of dictionary- and corpus-based surveys diverged from the results of his concept-based survey. The divergence in the survey results across the three conditions is thus the basis for Tobia's conclusion that corpus linguistic tools cannot accurately assess "ordinary understanding" or "ordinary meaning."¹⁸¹

While Tobia doesn't define what he means by ordinary understanding, he does quote Justice Holmes for the proposition that, in construing a legal text, "we ask, not what this man meant, but what those words would mean in the mouth of a normal speaker of English," and that words are used "in the sense in which they would be used by the normal speaker of English."¹⁸² Justice Holmes also said that "[w]hen a rule of conduct is laid down in words that evoke [a picture] in the common mind," the proper

¹⁷⁵ See, e.g., *id.* at 739, 753.

¹⁷⁶ Tobia does appear to provide a definition of the related term "ordinary meaning," stating that "ordinary meaning is generally informed by considerations of *how readers of the text would actually understand it.*" *Id.* at 739 (emphasis in original).

¹⁷⁷ See *id.* at 754 ("[C]oncept participants received no information so that they would rely on their ordinary understanding.").

¹⁷⁸ *Id.* at 735.

¹⁷⁹ See Tobia, *supra* note 17, at 772–73.

¹⁸⁰ *Id.* at 764.

¹⁸¹ *Id.* at 805.

¹⁸² *Id.* at 736–37 (quoting Oliver Wendell Holmes, *The Theory of Legal Interpretation*, 12 HARV. L. REV. 417, 417, 419–20 (1899)).

interpretation of the legal text should be consistent with that picture.¹⁸³ Tobia seems to suggest that the survey methods he employs give us privileged access to the picture that is elicited in the mind by the words of a statute—access that cannot be obtained through linguistic corpora.

Though Tobia doesn't define what he means by "ordinary understanding," a possible framing might turn on the distinction between linguistic "competence" and linguistic "performance." In his influential work *Aspects of the Theory of Syntax*, Professor Noam Chomsky distinguishes between linguistic competence, which is "the speaker-hearer's knowledge of his language," and performance, which is "the actual use of language in concrete situations."¹⁸⁴ Competence is thus the unconscious, internal knowledge of language, while performance reflects the actual, real-world use of language by speakers and writers.

It is possible to view corpus analysis as aimed at uncovering evidence of performance, while surveys could be viewed as producing evidence of competence. But this view is problematic for reasons explained below, and it is one that Chomsky would flatly reject. Chomsky made clear that he viewed the speaker-hearer's competence as "neither presented for direct observation nor extractable from data by inductive procedures of any known sort."¹⁸⁵

Survey responses do not present a direct window into the linguistic perceptions of survey respondents. And they do not give us direct evidence of Chomskyan competence. Like corpora, they provide *indirect* evidence of linguistic perception, use, and knowledge. In Chomskyan terms, survey responses, like corpus evidence, are just another type of performance.

Tobia does not reference Chomsky's *Aspects of the Theory of Syntax*, nor does he make explicit claims about *performance* or *competence*. But Tobia does argue that ordinary meaning is a search for "ordinary people's understanding of legal texts."¹⁸⁶ When presented with a divergence between his corpus-based survey and his concept-based survey, Tobia concludes that the corpus-based survey is in error. The implication is that when conclusions based on an examination of usage evidence diverge from direct questioning about concepts, it is the response to direct questioning that best reflects "ordinary understanding."

¹⁸³ *McBoyle v. United States*, 283 U.S. 25, 27 (1931).

¹⁸⁴ NOAM CHOMSKY, *ASPECTS OF THE THEORY OF SYNTAX* 4 (1965).

¹⁸⁵ *Id.* at 18.

¹⁸⁶ Tobia, *supra* note 17, at 736 (emphasis omitted).

This is problematic. A significant body of research calls into question the reliability of self-reported linguistic judgments and behaviors and the proposition that natural linguistic behaviors can ever be elicited in artificial linguistic settings like interviews and surveys. In circumstances where linguists employ interview or survey methods to elicit linguistic behavior, they take great pains in study design to avoid observer effects and create naturalistic linguistic settings. Tobia does not appear to have taken any such steps in his survey design. He has not cited anything from a large body of social science literature on the use of survey responses to elicit judgments about meaning and usage. And his concept condition results are questionable due to his failure to consider this literature.

The literature identifies at least three categories of problems with Tobia's reliance on his survey evidence: (a) survey respondents cannot be expected to accurately self-report their own linguistic performance and perceptions, (b) natural linguistic behavior cannot reliably be observed in artificial linguistic settings, (c) the survey methods described by Tobia are not designed to elicit responses that are consistent with natural language use.

a) Self-reporting problems. Tobia's concept condition survey results closely resemble experiments measuring respondents' judgments about the grammaticality or acceptability of a given utterance. Indeed, Tobia often uses the phrase "ordinary judgments" to describe the results of his concept-based survey.¹⁸⁷ The use of grammaticality judgment surveys was once widespread in linguistics. Linguists interested in trying to examine the Chomskyan notion of linguistic competence would seek to elicit such judgments from test subjects and survey respondents.

In recent decades, however, this approach to the study of language has been called into question.¹⁸⁸ Linguists have grown

¹⁸⁷ *E.g., id.* at 752.

¹⁸⁸ *See, e.g.,* WAYNE COWART, EXPERIMENTAL SYNTAX: APPLYING OBJECTIVE METHODS TO SENTENCE JUDGMENTS 2 (1997) ("[T]here have been continuing doubts about the empirical reliability and theoretical interpretation of judgment data as well as questions about what constitutes an appropriate technique for gathering judgment data. . . . [A]uthors have alleged that judgments are gravely compromised by instability of several different kinds."); Kathryn Bock, *Language Production: Methods and Methodologies*, 3 PSYCHONOMIC BULL. & REV. 395, 396 (1996):

The assumption is that the properties of a stimulus's mental representation are transparently reflected in the verbal response to the stimulus. This assumption in turn motivates an erroneous supposition that is pervasive in the literature:

concerned with the notion that responding to a survey question about grammaticality may involve entirely different cognitive processes from that of ordinary communication. Put simply, when survey respondents are asked for their judgments about language, they are not performing a natural language task. They are responding to an experiment in an experimental setting.¹⁸⁹ The introspection and inferential reasoning involved in responding to a linguistic prompt, and offering a linguistic judgment, are not the same cognitive task as simply engaging in communication. Thus, “[p]erformance in an experiment, including performance on the standard linguistic task of making grammaticality judgments, cannot be equated with grammatical knowledge. To determine properties of the underlying knowledge system requires *inferential reasoning, sometimes of a highly abstract sort*.”¹⁹⁰ Grammaticality judgments are sometimes said to be unreliable because of their “unavoidable meta-cognitive overtones,”¹⁹¹ that is, they involve cognitive processes that are not part of ordinary communication. For these reasons, “there have been continuing doubts about the empirical reliability and theoretical interpretation of judgment data.”¹⁹²

In fact, one reason given for skepticism about survey evidence of grammaticality judgments is that they do not match up with usage evidence from linguistic corpora:

Theoretical linguistics traditionally relies on linguistic intuitions such as grammaticality judgments for data. But the massive growth of language technologies has made the spontaneous use of language in natural settings a rich and easily

What one says, how one says it, and how long it takes to say it are unsullied reflections of input processing and interpretation.

See also Hiroshi Nagata, *The Relativity of Linguistic Intuition: The Effect of Repetition on Grammaticality Judgments*, 17 *J. PSYCHOLINGUISTIC RSCH.* 1, 3 (1988) (“[S]tudies suggest that grammaticality judgments of sentences are not always invariant but are variable depending on the conditions according to which the sentences are judged.”).

¹⁸⁹ See CARSON T. SCHÜTZE, *THE EMPIRICAL BASE OF LINGUISTICS: GRAMMATICALITY JUDGMENTS AND LINGUISTIC METHODOLOGY* 13 (Language Science Press, 2d ed. 2016) (“[C]ountless studies [] have demonstrated that grammaticality judgments are susceptible to order and context effects, handedness differences, etc., and have then concluded . . . that the grammar itself must have these properties. . . . Such conclusions are not justified.”).

¹⁹⁰ Jane Grimshaw & Sara Thomas Rosen, *Knowledge and Obedience: The Developmental Status of the Binding Theory*, 21 *LINGUISTIC INQUIRY* 187, 188 (1990) (emphasis added).

¹⁹¹ Shimon Edelman & Morten H. Christiansen, *How Seriously Should We Take Minimalist Syntax?*, 7 *TRENDS COGNITIVE SCIS.* 60, 60 (2003) (citing CARSON T. SCHÜTZE, *THE EMPIRICAL BASE OF LINGUISTICS: GRAMMATICALITY JUDGMENTS AND LINGUISTIC METHODOLOGY* (1996)).

¹⁹² COWART, *supra* note 188, at 2.

accessible alternative source of data. Moreover, studies of usage as well as intuitive judgments have shown that linguistic intuitions of grammaticality are deeply flawed.¹⁹³

That is, where corpus evidence and survey evidence diverge, many linguists express skepticism of the survey, rather than the corpus, because the corpus is a record of natural language, while the survey is a record of “highly abstract” “inferential reasoning”¹⁹⁴ with “meta-cognitive overtones.”¹⁹⁵

We do not mean to suggest that asking for linguistic judgments of survey respondents cannot be useful. Survey evidence is a measure of linguistic performance. But an examination of performance evidence is just one *indirect* method of trying to understand *actual* linguistic perception and usage. Yet surveys and elicited language judgments are not a privileged window into the workings of the mind. Where survey evidence and corpus evidence conflict, at a minimum it is not obvious that survey evidence should be preferred.

b) *Natural language versus unnatural language.* A second problem arises with Tobia’s assumption that surveys report natural linguistic behavior. Survey respondents are not engaged in natural linguistic communication, and a survey response is not a natural linguistic behavior. As foreshadowed above, responding to an artificial linguistic prompt in a survey may involve a different type of cognition than the type involved in ordinary communication. Survey respondents know that they are taking part in a survey, and that their responses are being observed and subjected to analysis. This fact alone may cause respondents to deviate from their natural linguistic behavior.

Corpus linguistics is one method of gathering evidence about language. One of the advantages of corpus linguistics is that linguistic corpora are designed to represent natural language use in natural linguistic settings. That is, linguistic corpora are large electronic collections of natural language—language that was produced in a natural linguistic setting and during ordinary communication (like reading and writing, or speaking and hearing).

¹⁹³ Joan Bresnan, *Is Syntactic Knowledge Probabilistic? Experiments with the English Dative Alternation*, in *ROOTS: LINGUISTICS IN SEARCH OF ITS EVIDENTIAL BASE, SERIES 75, 75* (Sam Featherston & Wolfgang Sternefeld eds., 2007).

¹⁹⁴ Grimshaw & Rosen, *supra* note 190, at 188.

¹⁹⁵ Edelman & Christiansen, *supra* note 191, at 60 (citing CARSON T. SCHÜTZE, *THE EMPIRICAL BASE OF LINGUISTICS: GRAMMATICALITY JUDGMENTS AND LINGUISTIC METHODOLOGY* (1996)).

Most corpora are not composed of language that was elicited or recorded specifically for the purpose of study or where the speaker, hearer, reader, or writer would have been aware that they were being observed. Corpus linguistics cannot give us access to the picture evoked in the mind of an individual during communication. But corpus linguistics can give us a lot of evidence of the linguistic output of that mind in natural linguistic settings.¹⁹⁶ The black box of the mind cannot be directly observed, and we can learn a lot about language and meaning through indirect observation—by making observations about the mind’s linguistic output.

Tobia’s article may be taken then as an attempt to validate or invalidate language claims based on evidence of natural language (i.e., corpus evidence) with evidence from unnatural language (i.e., language produced in the artificial linguistic environment of a survey response). The point of relying on usage evidence from linguistic corpora (indeed, one of the key tenets of corpus linguistics) is that the researcher has access to natural, unobserved (and therefore less tainted) language use from natural linguistic environments. We think that relying on survey responses as a proxy for ordinary understanding ignores a whole host of challenges that Tobia fails to address.

Survey responses may not reflect the actual linguistic perceptions or performance of survey respondents. “Even the best designed elicitation tasks are removed from how people use (and think about) language in everyday life, and people’s reports of their linguistic usage may or may not match up with what they actually do.”¹⁹⁷ Surveys “can induce respondents to claim knowledge and use of features they have never heard prior to the research situation.”¹⁹⁸ Survey respondents can very easily confound grammaticality with acceptability or correctness.¹⁹⁹ And survey respondents may be influenced by “ordering effects, participants’ possible discomfort with the test-like nature of the elicitation task, and their resulting desire either to do ‘well’ on the test” by providing the answers they perceive the test

¹⁹⁶ See Svenja Adolphs & Phoebe M.S. Lin, *Corpus Linguistics*, in THE ROUTLEDGE HANDBOOK OF APPLIED LINGUISTICS 597, 597 (James Simpson ed., 2011) (“At the heart of empirically based linguistics and data-driven description of language, corpus linguistics is concerned with language use in real contexts.”).

¹⁹⁷ Natalie Schilling, *Surveys and Interviews*, in RESEARCH METHODS IN LINGUISTICS 96, 102 (Robert J. Podesva & Devyani Sharma eds., 2013).

¹⁹⁸ *Id.*

¹⁹⁹ *See id.* at 102–03.

administrator “expects” or “to get the test over with as quickly as possible.”²⁰⁰ In addition,

The traditional method of eliciting language attitudes is plagued by the same problems as elicitations of speech production. The tasks are unnatural, and there is no guarantee that the results are reflective of listeners’ genuine attitudes. . . . This may be because listeners do not have free access to their attitudes or the ability to accurately convey them, or because they do not wish to express negative attitudes they might really hold.²⁰¹

If the test for ordinary meaning is the picture that is elicited in the common mind by the words of the legal text, it is not at all clear that survey evidence will give us a higher resolution of that picture than corpus linguistics will.

c) *Survey design and observer effects.* One obvious response to these objections to survey evidence is that linguists, like other social scientists, frequently rely on surveys to study language and meaning. Yet linguists have recognized the problems inherent in relying on survey evidence to study linguistic behavior. And they have devised a number of methods for mitigating observer effects and eliciting natural language. Tobia does not appear to have employed any of these techniques.

Survey methods suffer from what sociolinguist William Labov referred to as “the *Observer’s Paradox*: the aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed; yet we can only obtain these data by systematic observation.”²⁰² If we are aiming to understand natural language use or natural language perception (i.e., ordinary understanding), then unnatural language produced in an artificial setting may not be satisfactory.²⁰³

Labov proposed that “[o]ne way of overcoming the paradox is to break through the constraints of the interview situation by

²⁰⁰ *Id.* at 103.

²⁰¹ *Id.* at 106.

²⁰² WILLIAM LABOV, *SOCIOLINGUISTIC PATTERNS* 209 (1972) (emphasis in original).

²⁰³ See William Labov, *Some Principles of Linguistic Methodology*, 1 *LANGUAGE SOC’Y* 97, 108 (1972) (“In the gathering of elicitations and intuitions, there is no obvious sense in which the work can be described as *valid*.” (emphasis in original)); SIDNEY GREENBAUM & RANDOLPH QUIRK, *ELICITATION EXPERIMENTS IN ENGLISH: LINGUISTIC STUDIES IN USE AND ATTITUDE* 7 (1970) (“We are well aware that we cannot escape from the artificiality of the test situation, though with continuing refinement we can hope to remove some of the worst effects of bias that the test situation introduces.”).

various devices which divert attention away from speech,” so that “the subject unconsciously assumes that he is not at that moment being interviewed.”²⁰⁴ Or, the interviewer can “involve the subject in questions and topics which recreate strong emotions he has felt in the past, or involve him in other contexts.”²⁰⁵ A surveyor may administer a rapid and anonymous survey, in which relevant speech is elicited through questions administered before respondents are aware that they are participating in the survey.²⁰⁶

Even where such mitigating techniques are employed, respondents are typically aware that they are in an unnatural linguistic setting and that their language is being evaluated.²⁰⁷ Moreover, there is some evidence that when responding to survey questions, respondents are speaking in an entirely different register from ordinary speech.²⁰⁸

Whatever the advantages or disadvantages of these mitigation techniques, it is clear that Tobia made no attempt to mitigate observer effects. Instead, he merely directed responses to questions that were obviously directed at the respondent’s language judgments, like “Is a car an ailac? [Yes / No]” or “An airplane is a prototypical vehicle. 1 (strongly disagree) to 7 (strongly agree).”²⁰⁹ In such a case, it is not clear that Tobia’s survey responses are representative of the natural language use or perception of his respondents.

2. Tobia’s data are unhelpful given his study design.

Tobia’s survey data also suffer from a second definitional problem. In his concept condition survey, Tobia asked his respondents to determine whether items on a list of nouns were properly classified as a “vehicle.” But in so doing Tobia never gave any indication of how the respondent was to decide on the breadth of the classification of “vehicle.” In other words, he never defined “ordinary meaning” for his survey respondents. This left their answers up to their own judgment about the very legal construct he was trying to measure, robbing the data of any utility.

Tobia’s move ignores the wide range of senses of “ordinary meaning” that we identified in *Judging Ordinary Meaning*. Our

²⁰⁴ LABOV, *supra* note 202, at 209.

²⁰⁵ *Id.*

²⁰⁶ See Schilling, *supra* note 197, at 103–04; GREENBAUM & QUIRK, *supra* note 203, at 7.

²⁰⁷ See Schilling, *supra* note 197, at 109.

²⁰⁸ See *id.*

²⁰⁹ Tobia, *supra* note 17, at 756, 759 (brackets in original).

article highlighted a problematic imprecision in the law's search for ordinary meaning.²¹⁰ Citing classic cases on statutory ambiguity, we noted that the range of senses embraced in legal theory spans a frequency continuum (running from "permissible" senses of a given term to the "most frequent" or even "exclusive" sense) and sometimes also sweeps in "prototype" meaning.²¹¹ We also identified different values or policies that could be viewed to sustain different concepts of ordinary meaning.²¹²

The foregoing principles were a prelude to our demonstration of the utility of corpus linguistic tools. We noted that the imprecisions in the attempt to conceptualize ordinary meaning are wrinkles that need to be ironed out in our legal theory of interpretation.²¹³ And we concluded by demonstrating that corpus tools can yield data of relevance to any of the notions of ordinary meaning embraced in the case law.²¹⁴ For example, we

²¹⁰ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 798 ("[I]ronically, we have no ordinary meaning of 'ordinary meaning.'").

²¹¹ *Id.* at 800–01.

²¹² *See, e.g., id.* at 859 (emphasis in original):

If we accept the most common use of the word as the ordinary meaning, we can conclude that the ordinary meaning of *vehicle* is *automobile*.

We can also make a strong case for crediting the most common meaning as the ordinary one, in that it will best avoid unfair surprise (public meaning) and vindicate the presumed intent of the lawmaker (intended meaning). . . . But, as discussed above, that is a question for legal theory.

See also id. at 817–18:

Intended meaning is an appropriate construct to the extent we are aiming to vindicate the preferences of lawmakers. This is a viable, distinct basis for crediting ordinary meaning. . . .

There is also a case for the public or "reader's" understanding. This sort of meaning makes sense to the extent we are seeking to vindicate the notice rationale for the "standard picture"—the protection of reliance interests and the avoidance of unfair surprise.

²¹³ *See id.* at 874 (emphasis in original) (quoting Baude & Sachs, *supra* note 68, at 1089–90):

This raises the question of whether to credit only the top-of-mind sense or a possibly broader, "reflective" sense as ordinary. But this is not a deficiency in corpus data—or even in linguistic theory. It is a question for law—"we have to decide *which* meaning, produced by which *theory* of meaning, we ought to pick."

See also id. at 858:

What if our sense of public meaning differs from our sense of intended meaning? If that happens we would need to decide which data set to rely on. That is a problem for legal theory—and essentially a choice of which of two sets of justifications for the "standard picture" we seek to vindicate.

²¹⁴ *See, e.g., id.* at 858:

showed how corpus tools could help us determine whether airplanes, bicycles, and automobiles were simply acceptable, common, or prototypical examples of “vehicles.”²¹⁵

Tobia skates past all of these nuances. In crediting results from his concept condition survey, Tobia purports to be establishing a baseline truth about human perception of ordinary meaning. But he nowhere acknowledges the range of possible meanings of ordinary meaning. And because his survey gave no guidance to his concept condition respondents, their classifications of various nouns as falling inside or outside the “vehicle” category are ultimately indecipherable. Without more, we cannot tell whether the survey respondents are thinking of the question whether a given item qualifies as a vehicle in terms of the “permissible” sense of the term or whether they may instead be thinking of it in terms of a “common,” “most frequent,” or even “prototype” sense.

This is fatal to the utility of Tobia’s inquiry. Even if we assume away the above concerns about the probity of evidence of language competence, there is no basis in Tobia’s experiment for his decision to elevate the results of his concept condition survey. The survey is in this sense a black box. The respondents inside it could be telling us that a given item counts (or does not count) as a “vehicle” because they think all permissible senses of vehicle should be included. Or they could have in mind a different standard for inclusion, based on a “most frequent” sense of “vehicle” or even a “prototype” sense. We will never know. And without such knowledge we cannot even begin to have a conversation about whether the data are conclusive. (It would only be the beginning of the conversation in any event, as the answer would still depend on the debatable considerations we highlighted in *Judging Ordinary Meaning*—as to whether or when each of the various

The speech community question, as we have noted, has implications for the selection of a relevant corpus. If we are trying to measure intended meaning, we might want to gather data from a corpus of a community of speakers who look demographically like Congress. Yet if we are interested in public meaning, we would want to turn to a broader corpus.

²¹⁵ See, e.g., Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 859 (emphasis in original):

[W]e can conclude that the most common sense of this term is in reference to automobiles. Airplanes and bicycles appear on our frequency continuum: they are attested in the data as possible examples of *vehicle*. But they are unusual—not the most frequent and not even common. If we accept the most common use of the word as the ordinary meaning, we can conclude that the ordinary meaning of *vehicle* is *automobile*.

concepts of ordinary meaning is preferable as a matter of a legal theory of interpretation.)

For these reasons, Tobia’s analysis is circular. His data show only that his concept condition respondents classified various entities as a “vehicle” in accordance with their own unstated standard for inclusion. And the data therefore tell us nothing of interest. We cannot begin to have a coherent conversation about testing ordinary meaning until we define “ordinary meaning.” And Tobia is in no position to close the debate about the right measurement tool if he is unwilling to start by defining the standard that he is seeking to operationalize.

Our corpus analysis suffers from no such problems. We identified a range of possible meanings of “ordinary meaning” and showed how corpus tools can operationalize each of them. Admittedly, we left open the choice among those concepts of ordinary meaning. But we emphasized that that is a question for our legal theory of interpretation. And we identified a range of policy considerations that might seem to favor any of a range of these concepts.

We do not mean to dismiss the possibility that survey instruments could add meaningfully to the inquiry into the ordinary meaning of legal language.²¹⁶ But we see substantial barriers to the utility of the data brought to bear in the Tobia study.²¹⁷ And we think the proponents of surveys must at least define their terms before they can claim to have identified a baseline truth for assessing the accuracy of corpus methods.

3. Tobia’s data fail to consider syntactic and pragmatic context.

Tobia’s survey data also falter in their failure to account for the range of syntactic and pragmatic context that corpus tools can incorporate. In *Judging Ordinary Meaning*, we presented the linguistic basis for the need to assess the ordinary meaning of words or phrases in light of all their relevant context.²¹⁸ We noted that

²¹⁶ See *id.* at 861–62 (discussing the potential of surveys in this enterprise).

²¹⁷ See *id.* at 861 (discussing some “significant barriers to using survey data to address questions of ordinary meaning”).

²¹⁸ See *id.* at 816–17 (citations omitted):

Real human beings do not derive meaning from dictionary definitions and rules of grammar alone. Everyone takes nonsemantic context—pragmatics—into account in deriving meaning from language. And for that reason we see no basis to credit semantic meaning without consideration of pragmatic context. If no lawmaker would read the text that is voted into law purely semantically—devoid of pragmatic context—then there is no reason to credit that kind of meaning as

words are not understood in isolation but in context. And we highlighted the salience of both syntactic context²¹⁹ (surrounding words and language structures) and pragmatic context²²⁰ (the physical or social setting in which the words appear).

These points can be illustrated in the “vehicle” examples listed in Tobia’s article. Tobia’s survey asks respondents to decide only whether items in a list of nouns can qualify as a “vehicle.” No syntactic context or pragmatic context is included—respondents have no sense of whether they are being asked to decide whether an item counts as the sort of “vehicle” that would be subject, say, to a state vehicle registration law (one of Tobia’s examples), fall within an insurance policy covering all vehicles owned by an insured (another), or count as a vehicle prohibited by a city ordinance banning “vehicles in the park” (the classic H.L.A. Hart problem).

This is problematic. If our human understanding of language is informed by syntax and pragmatics, we may have a different understanding of the ordinary meaning of “vehicle” in each of the three settings noted above. And Tobia’s concept condition survey makes no attempt to consider this context.

Corpus analysis could easily do so. We could construct a corpus search not just for the noun “vehicle” but for the noun “vehicle” as an object of a human subject who is seeking to “register” it with a state, or a human subject who is asserting “ownership” in the context of “insurance.” If the ordinary meaning of “vehicle” varies depending on this context, we could measure it using the tools of corpus linguistics.²²¹ The flexibility of corpus tools, in fact, is a key selling point for corpus analysis—a point we developed at great length in *Judging Ordinary Meaning*. With that in mind, it is puzzling that Tobia insists on the superiority of his simplistic survey, which makes no attempt to capture this context or to engage with our defense of corpus methods on this basis.

A survey instrument could of course be devised in a manner incorporating these elements of context. Instead of just asking a respondent to determine whether or not a given item counts as a

a means of vindicating the intent of a lawmaker. The same goes for the public governed by the law. If no one reads laws literally by pure semantics, we have no reason to protect reliance interests or notice concerns rooted in that kind of understanding.

²¹⁹ See *id.* at 821–22.

²²⁰ See Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 823–24.

²²¹ See *id.* at 874–75.

“vehicle” in the abstract, the survey could tell the respondent that it is inquiring into the scope of a state vehicle registration law or the coverage of an insurance policy. In expanding the survey in that way, the context deficiency highlighted here could be addressed.

But that would introduce another set of problems—problems associated with the difficulty of assembling reliable evidence of language competence (due to self-reporting problems, unnatural language issues, and observer effects). These problems can be magnified when we extend the context given to survey respondents to incorporate pragmatic context. The more context we provide, the less certain we can be that the evidence we are gathering is telling us anything useful about ordinary language usage. Because survey respondents are humans (with biases and prejudices), their answers at some point will inevitably be more about preferences for outcomes in anticipated disputes than ordinary understanding of language usage.

The insurance coverage example highlights the problem. Our understanding of “vehicle” may well be influenced by the syntactic and pragmatic context of the insurance policy. But if we tell respondents that they are being asked to decide whether an insurance company should be required to extend insurance coverage to protect a policyholder from harm, that will certainly sway their answers. For example, as many respondents will be policyholders themselves, they may feel an increased desire to answer in the affirmative. Such a swayed answer seems unlikely to be based on perception of language and more likely to be about an implicit bias against insurance companies.

We are not suggesting that this problem is intractable. Survey instruments could perhaps be designed in a manner that could control for or otherwise deal with these concerns. But Tobia’s instrument does not do that. And for now we can at least say that his study falls far short in its attempts to discredit the accuracy of corpus-based analysis. If anything, it highlights the advantages that our proposed tools have over the method he touts as yielding baseline truth.

* * *

While neither corpus evidence nor survey evidence provides a direct window into the natural linguistic behavior or perceptions of language users, both methods can provide useful

information about language and meaning, even if that information must be dimly viewed through “the performance filter.”²²²

We do not mean to suggest that survey methods cannot provide useful evidence of language use and perception. Surveys have been used since at least the 1800s to gather evidence of language usage.²²³ And survey methods have been brought to bear on topics ranging from language perception and attitudes²²⁴ to language variation.²²⁵ Survey methods also have the advantage that they are already familiar to lawyers and judges. As noted by Professors Omri Ben-Shahar and Lior Strahilevitz in advocating for the use of survey methods in contract interpretation: “a large background of law is already in place making survey wording and technique in trademark disputes reasonably consistent across cases,” and “[a] substantial amount of case law exists which provides insight into how to conduct and prepare a trademark survey that will be admissible in court.”²²⁶ And to the extent that acceptability judgments are relevant to the inquiry of ordinary meaning, there is at least some evidence that survey methods using the Amazon Mechanical Turk are at least as good as live surveys in measuring acceptability judgments of respondents.²²⁷

Yet there are reasons to prefer corpus evidence in some contexts. Linguistic corpora provide evidence of natural language—language that occurs during ordinary communication and in natural linguistic settings. Linguistic corpora thus avoid observer effects. The language in the corpus occurred naturally and was not elicited in an experimental setting. No one was around to make metalinguistic judgments about the language being used, and the language in the corpus typically did not involve a subconscious performance for an observer. The corpus allows us to

²²² Grimshaw & Rosen, *supra* note 190, at 217.

²²³ See Schilling, *supra* note 197, at 97.

²²⁴ See generally Kathryn Campbell-Kibler, *Sociolinguistics and Perception*, in 4 LANGUAGE AND LINGUISTICS COMPASS 377 (2010); Jennifer Hay, Paul Warren & Katie Drager, *Factors Influencing Speech Perception in the Context of a Merger-in-Progress*, 34 J. PHONETICS 458 (2006).

²²⁵ See generally Isabelle Buchstaller & Karen Corrigan, *How to Make Intuitions Succeed: Testing Methods for Analysing Syntactic Microvariation*, in ANALYSING VARIATION IN ENGLISH 30 (Warren Maguire & April McMahon eds., 2011).

²²⁶ Omri Ben-Shahar & Lior Strahilevitz, *Interpreting Contracts via Surveys and Experiments*, 92 N.Y.U. L. REV. 1753, 1774 (2017) (citing Jerre B. Swann, *Judge Richard Posner and Consumer Surveys*, 104 TRADEMARK REP. 918, 921 (2014)) (quoting Robert H. Thornburg, *Trademark Surveys: Development of Computer-Based Survey Methods*, 4 J. MARSHALL REV. INTELL. PROP. L. 91, 91 (2004)).

²²⁷ See Jon Sprouse, *A Validation of Amazon Mechanical Turk for the Collection of Acceptability Judgments in Linguistic Theory*, 43 BEHAV. RSCH. METHODS 155, 164–66 (2011).

examine linguistic production only. Thus, it is difficult to obtain the “picture” “evoked in the common mind” by the words of a legal text.²²⁸ But the corpus allows us to see how those words are used in a similar context, by a similar speech community, and at a similar time. The corpus may still provide indirect evidence of natural language, but it is untarnished by an experimental setting.

Surveys also provide indirect evidence of language use and perception, but responding to a survey is not a natural linguistic task—it is an unnatural one. Responses to surveys require a sort of metalinguistic, inferential cognition that does not form part of ordinary communication.

Thus, where corpus evidence and survey evidence diverge in interpretive contexts, it is not obvious that survey evidence should be preferred. Indeed, if the task is to understand natural linguistic behavior in natural linguistic settings, then we have reason to prefer corpus evidence to survey evidence.

B. Tobia’s Survey Does Not Show That Corpus Tools Are Biased

Tobia claims that dictionary tools establish what he calls “more extensive” meaning and that corpus linguistic tools establish a sort of “prototypical[]” meaning.²²⁹ And because corpus linguistic tools purportedly point only to prototypical meaning, Tobia says that they lead to the “Nonappearance Fallacy”—the notion that “any use” of a studied term “that is not reflected” in a corpus “is therefore not part of the ordinary meaning” of the term.²³⁰ Tobia also invokes a concept that he thinks of as a related concern—an “Uncommon Use Fallacy,” or the view (attributed to proponents of corpus tools) that “the relative rarity of some use in the corpus indicates that this use is outside of the ordinary meaning.”²³¹

The results of Tobia’s dictionary condition survey flow directly from the study design. Tobia gets extensional “vehicle” data from his dictionary condition participants because he gives them only extensional definitions—the “means of carrying or transporting something” and “agent of transmission” definitions.²³² With these definitions of “ailac” in hand, respondents naturally

²²⁸ *McBoyle*, 283 U.S. at 27 (alteration omitted).

²²⁹ Tobia, *supra* note 17, at 766.

²³⁰ *Id.* at 795 (citing *Carpenter v. United States*, 138 S. Ct. 2206, 2238–39 (2018) (Thomas, J., dissenting)).

²³¹ *Id.* at 796.

²³² *Id.* at 755.

conceived of a concept of “vehicle” that is broad and extensional. We agree that that is problematic. But the problem doesn’t stem from reliance on dictionaries. It stems from a flawed focus on only broad, extensional dictionary definitions.

We made this point as to “vehicle” in *Judging Ordinary Meaning*. We noted that dictionaries use both extensional definitions (e.g., “carrier,” “agent of transmission”) and more limited ones that could be examples of a prototype definition (e.g., “automobile”).²³³ And we thus concluded that the dictionary cannot answer the problem of lexical ambiguity presented in the “no vehicles in the park” problem.²³⁴ That is the problem with dictionaries.

Tobia also resorts to the “vehicle” example to support his view that corpus analysis “often neglects nonprototypical uses of a term.”²³⁵ He notes that the “predominant[]” example of “vehicle” in the corpus is a car.²³⁶ And because he views items like “golf carts, airplanes, and horse-drawn carriages” as falling “within the modern ordinary meaning of ‘vehicle,’” he concludes that corpus analysis is systematically flawed.²³⁷

These claims are not supported by the data presented in the Tobia article. Tobia’s survey results do not reveal fundamental flaws in the corpus analysis that we have advocated (and that careful interpreters conduct). They just show that the flawed use of these tools will produce predictably flawed data. The Nonappearance Fallacy and Uncommon Use Fallacy, moreover, misstate our position. We have not advocated a use of corpus tools susceptible to Tobia’s criticisms. In fact, corpus tools, unlike Tobia’s survey, are quite capable of avoiding the fallacies that Tobia complains about.

That said, Tobia’s critiques help highlight some points of imprecision in our previous articulations of the role of corpus linguistic analysis in legal interpretation. A full response to Tobia’s analysis requires us to refine our approach—to map out a more careful framework for assessing corpus data, and to explain how the data might fit into a broader framework of interpretation.

We elaborate on all of these points below. First, we show that the supposedly skewed nature of Tobia’s corpus condition results flows not from a flaw in corpus tools but in the design of Tobia’s

²³³ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 801–02, 807.

²³⁴ *See id.* at 807.

²³⁵ Tobia, *supra* note 17, at 795.

²³⁶ *Id.*

²³⁷ *Id.*

experiment. Then, we show that Tobia’s purported fallacies speak to uses of corpus linguistics that we did not embrace in *Judging Ordinary Meaning*. And finally, we clarify the use of corpus tools that we propose and explain how we see those tools fitting into an overall framework of legal interpretation.

1. Tobia’s results come from his flawed corpus condition design.

Tobia’s critiques in his corpus condition analysis also flow from methodological defects in his study design. Again, the data don’t highlight flaws in corpus analysis generally, but rather in Tobia’s flawed survey design.

Tobia gets prototype data from his corpus condition participants because he gives them only a list of collocates followed by nine concordance lines.²³⁸ That isn’t how a careful practitioner uses corpus tools. A prudent, informed use of the tools can avoid the problems that Tobia is concerned with.

Tobia claims to be giving his study participants “*precisely*” what practitioners of corpus analysis would give them.²³⁹ But that is wrong. Tobia gives his corpus condition participants only a list of common collocates of “vehicle” and nine bare concordance lines. By asking them to assess the dummy term “ailac,” he then robs participants of any access to their linguistic intuition about the actual term in question (“vehicle”).

All of these moves run contrary to the methods we have advocated. And all of them predictably skew the numbers in Tobia’s survey.

In our corpus analyses (in judicial opinions and academic articles), we have always started with linguistic intuition regarding the range of meanings of a studied term. And we have used dictionaries to establish a starting set of possible senses of the studied term that we may then seek to measure in a chosen corpus.

We did that in our “vehicles” analysis in *Judging Ordinary Meaning*.²⁴⁰ That is an important first step. Without it, and especially when looking for meaning of a dummy term, the empirical design effectively dictates “prototype” results. If study participants have no range of senses of the studied term in mind and are

²³⁸ *Id.* at 755–56, 756 n.137.

²³⁹ *Id.* at 756 (emphasis in original).

²⁴⁰ See Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 800–02 (coming up with different possible senses of the word “vehicle” by resorting to dictionaries and linguistic intuition).

given only a list of collocates and examples, they are left with nothing more to do than to establish a mental picture of attributes of the examples they are given. That is prototype analysis. And Tobia's prototype results flow from this study design.

Tobia's results are skewed further by the dumbed-down corpus evidence that he gives his study participants. We have never used a list of collocates as a determinative basis for assessing ordinary meaning. We have always paired such a list with a starting set of definitions or senses of a studied term.²⁴¹ In assessing the relative frequency of competing senses, moreover, we have never examined a list of concordance lines that is anywhere near the bare nine lines that Tobia employed.²⁴² In *Judging Ordinary Meaning*, one of our key selling points was the ability of corpus analysis to allow us to use relatively large numbers of concordance lines to improve the statistical reliability of our analysis.²⁴³ And we acknowledged that a limited evidence set could produce the same kind of bias we might find in a panel of judges.²⁴⁴

One important reason to examine a large number of concordance lines is to overcome the risk that a small number of people (like the nine people who sit on the U.S. Supreme Court) will understand a term differently than the speech community in which we're interested. By presenting only nine concordance lines and depriving study participants of any background knowledge of lexical information about the studied term, Tobia is eliminating the possibility of assembling any useful evidence. His results tell us nothing useful about the effect of corpus evidence. They tell us

²⁴¹ See, e.g., *id.*

²⁴² See, e.g., *id.* at 841 ("In order to examine the sense distribution of *vehicle*, we reviewed one hundred randomized concordance lines of *vehicle* in the NOW Corpus." (emphasis in original)).

²⁴³ See, e.g., *id.* at 795 (advocating the use of corpus linguistics because it employs "large bodies—*corpora*—of naturally occurring language" that allow us to "look for patterns in meaning and usage in large databases of actual written language" and "allow us to conceptualize and measure the 'standard picture' in a much more careful way" (emphasis in original)); *id.* at 820 (explaining that the "large bodies or databases of naturally occurring language" means that "[c]orpus analysis has allowed lexicographers to address the problem of sense division with greater granularity" and "view a more complete range of potential uses of a given word and collect statistical information about the likelihood of a given word appearing in a particular semantic environment").

²⁴⁴ See, e.g., *id.* at 878 ("Moving forward, judges, lawyers, and linguists will need to collaborate to settle on some best practices in this emerging field" such as establishing "standards for the appropriate sample size for a given search."); see also *id.* at 866 ("Corpus data can be gathered and analyzed properly only with care and a little background and training in the underlying methodology. A judge who proceeds willy-nilly may, either consciously or unwittingly, proffer data that has only the appearance of careful empiricism.").

only what people are likely to do with hopelessly scant and flawed corpus evidence.

2. Tobia's fallacies misstate our views.

The two "fallacies" invoked by Tobia are straw men. We have not adopted and do not endorse the notion that any "use that is not reflected" in a corpus²⁴⁵ (or is even only uncommonly reflected) cannot fall within the "ordinary meaning" of a studied term. As to whether an airplane is a vehicle, we did "ask" in *Judging Ordinary Meaning* "if *airplane* is even a possible sense of *vehicle*" given that we found no example of anyone speaking of an airplane as a vehicle in the corpus.²⁴⁶ But the quoted sentence was phrased as a question to be considered. And in forming the question, we were just highlighting a larger point about the kinds of questions that corpus evidence can answer.

Because one sense of ordinary meaning is a possible or previously attested sense of a term, we were noting that an interpreter who believes that an unattested example of a given use of a term is not covered by the ordinary meaning of a statute might reject "airplane" as falling outside the statute.²⁴⁷

Elsewhere in our article, moreover, we went to great lengths to stop short of committing to a correct sense of "ordinary meaning." We emphasized that the choice of a concept of ordinary meaning is a matter for our legal theory of interpretation.²⁴⁸ And we highlighted the range of policy considerations and social values that might favor each of a range of senses of "ordinary meaning."²⁴⁹ Tobia's critiques miss these nuances. And by missing them

²⁴⁵ Tobia, *supra* note 17, at 795.

²⁴⁶ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 844 (emphasis in original).

²⁴⁷ *See id.* at 844–45 (emphasis in original):

To the extent that *airplane* fits what some lexicographers have regarded as the necessary and sufficient conditions for inclusion in the class of *vehicles* (i.e., anything that is a "means of carriage, conveyance, or transport"), all that can be said of *airplane* is that it may be a possible meaning of *vehicle*, but it is unattested in the corpus data.

²⁴⁸ *See, e.g., id.* at 874 ("It is a question for law—'we have to decide *which* meaning, produced by which *theory* of meaning, we ought to pick.'" (emphasis in original) (quoting Baude & Sachs, *supra* note 68, at 1089–90)).

²⁴⁹ *See, e.g., id.* ("We think the answers to these questions are dictated in part by the rationales that drive us to consider ordinary meaning. A concern for fair notice and protection of reliance interests may well direct us to stop at the top-of-mind sense of a statutory term.").

he misattributes the cited fallacies to our proposed corpus linguistic analysis.

Tobia's critique also misses another important nuance in our analysis. In discussing a range of examples of lexical ambiguity in *Judging Ordinary Meaning*, we acknowledged the complications that arise in circumstances in which each of two competing senses of a term is "closely related" to the other.²⁵⁰ And, citing an insightful article by Professors Larry Solan and Tammy Gales, we noted the question whether "corpus data may reflect only the fact that a given sense of a certain term is a more factually common iteration of that term in the real world."²⁵¹ These are questions implicated by Tobia's critique. But Tobia does not engage with them. He does not even acknowledge the discussion of these problems in *Judging Ordinary Meaning*.

Our discussion of the related senses problem analyzed both a hypothetical and *Taniguchi*. The question in *Taniguchi* was straightforward—whether a statute authorizing an award of costs for the use of an "interpreter" encompassed the services of a written translator or was limited only to face-to-face oral translation services. We noted that the corpus evidence shows that the vast majority of uses of the term "interpreter" are in reference to a face-to-face translator. But we noted that "[t]he notion of *oral translator* could simply be perceived as a more common 'prototype' of the more general notion of 'one who translates,'" while the "*written translator* idea could [] be viewed as an *atypical* example."²⁵² Far from committing to the Nonappearance or Uncommon Use Fallacies, we emphasized that the nonappearance or limited appearance of a closely related sense of a term "would not tell us that an ordinary person would not understand text providing for compensation for an *interpreter* to cover a *written translator*."²⁵³

We highlighted this point by reference to a hypothetical rental agreement prohibiting tenants from keeping "dogs, cats,

²⁵⁰ *Id.* at 852 (citation omitted):

One possible limitation [of corpus linguistics] stems from the vagaries of word sense division. Sense division is subjective. . . . [Linguists] concede that distinctions among senses may be "more of a descriptive device rather than a claim about psycholinguistic reality." This seems particularly true as regards closely related or fine-grained sense distinctions.

²⁵¹ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 873–74 (summarizing a concern raised by Solan and Gales, since published in Solan & Gales, *supra* note 13, at 1351).

²⁵² *Id.* at 821 (emphasis in original).

²⁵³ *Id.* (emphasis in original).

birds, or other pets.”²⁵⁴ We noted that a tenant “found in possession of a caged dodo” would “not likely [] escape the wrath of the landlord by insisting that a dodo is an ‘obsolete’ sort of a bird” not attested in a corpus.²⁵⁵

A further elaboration of this point appeared in our discussion of the Solan-Gales question noted above. We conceded that “the fact that a given sense of a certain term is a more factually common iteration of that term in the real world” could give “reason to doubt the probity of the data in establishing the semantic meaning” of a statutory term.²⁵⁶ We emphasized the importance of this concern and warned that “anyone turning to corpus analysis would do well to consider [this] limitation[] before jumping too quickly to an inference about ordinary meaning.”²⁵⁷ And we proceeded to highlight a key element of our thesis: that corpus tools can help operationalize a wide range of senses of ordinary meaning seemingly embedded in the law of interpretation, but that the choice among those senses is a matter for our legal theory of interpretation.²⁵⁸

On the question whether to credit a “top-of-mind” sense that might match up closely with linguistic prototype or a “possibly broader, ‘reflective’ sense,”²⁵⁹ we discussed the question presented in *Muscarello*—whether transporting a gun in the locked glove box of a car counts as “carrying” it in relation to a drug crime. After noting that the overwhelming majority of corpus uses of a person carrying a firearm involve personally *bearing* or “packing” a gun, we questioned “whether to credit only the top-of-mind sense” in line with corpus evidence or include a broader sense that might be viewed to be closely related and included in a “reflective” sense of the term (e.g., *transporting* a gun in a car).²⁶⁰ Instead of committing to the former, we noted that this choice is “a question for law—we have to decide *which* meaning, produced by which *theory* of meaning, we ought to pick.”²⁶¹

This looks nothing like the Nonappearance or Uncommon Use Fallacy. It is a nuanced clarification of a problem with which

²⁵⁴ *Id.*

²⁵⁵ *Id.*

²⁵⁶ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 873–74.

²⁵⁷ *Id.* at 874.

²⁵⁸ *See id.* at 874–76.

²⁵⁹ *Id.* at 874.

²⁶⁰ *Id.*

²⁶¹ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 874 (emphasis in original) (quoting Baude & Sachs, *supra* note 68, at 1089–90).

all theories of interpretation must grapple. The difficulty does not reveal any “deficiency in corpus data—or even in linguistic theory.”²⁶² It just highlights an unresolved problem for legal interpretive theory.

In our article, we made clear that the choice of a concept of ordinary meaning is “dictated in part by the rationales that drive us to consider ordinary meaning.”²⁶³ “A concern for fair notice and protection of reliance interests,” for example, “may well direct us to stop at the top-of-mind sense of a statutory term.”²⁶⁴

If the personally *bear* sense of *carry* is the first one that comes to mind, then that may be the sense that the public will have in mind upon reading the terms of a statute, and if we are interested in protecting reliance interests and avoiding unfair surprise, we may want to stop short of including the broader *transport* sense that the public might concede to be covered upon reflection.²⁶⁵

If, on the other hand, we are convinced that the two senses of “carry” are indistinguishable, or that our sense of the pragmatic context of a statute convinces us that the lawmaker could not have intended a difference, then we may wish to sweep in the broader, “reflective” sense of “carry.” But that, again, is a matter for legal theory. If we did go with the “reflective” sense, we would be crediting an alternative set of policies or values—a “faithful agent” theory of interpretation.²⁶⁶ This credits the presumed intentions of the lawmaker. That is the point of our dodo bird example. A dodo may be so long forgotten that it is unattested in a modern corpus. But because there is no sense of “birdness” that includes a common example like a parakeet but excludes an uncommon example like a dodo, there is no basis for excluding the dodo.

3. Tobia’s critiques highlight the need to refine our methodology.

The above should close the door on the notion that corpus tools can reveal only prototypical meaning. It should also highlight the fact that there is no Nonappearance or Uncommon Use

²⁶² *Id.*

²⁶³ *Id.*

²⁶⁴ *Id.*

²⁶⁵ *Id.* at 874 (emphasis in original).

²⁶⁶ See Sunstein, *supra* note 93, at 435.

Fallacy in our article. Tobia's responses, however, highlight the need for some clarification and increased refinement of our proposed methodology. Some of the refinement appears in *Judging Ordinary Meaning*—in sections dealing specifically with at least one of the examples (golf carts) that Tobia criticizes us for excluding. We restate and amplify our analysis of the golf cart example here, with an eye toward refining the methodology that we propose (and that Tobia has either missed or chosen to ignore).

Tobia criticizes corpus analysis for purportedly excluding golf carts, airplanes, and bicycles, which he deems to fall within the ordinary meaning of vehicle.²⁶⁷ Yet he presents no defensible basis for his conception of ordinary meaning.²⁶⁸ And his critique ignores not only the nuances in our theory but our specific analysis of golf carts. In *Judging Ordinary Meaning*, we noted that “[w]e found no examples of golf carts as vehicles in the corpus” but immediately suggested that that did not “mean they do not qualify under the ordinary meaning of *vehicle*.”²⁶⁹ Instead, we explained that the absence of golf cart examples in the corpus requires more careful thinking about (a) the necessary and sufficient components of “vehicleness” as informed by the example senses in the corpus’s concordance lines and (b) whether the specific example in question (golf cart) falls within that concept of “vehicle.”²⁷⁰ Thus, we noted that “a golf cart shares a number of features with the most common vehicles” (automobiles) but that it also differs in some respects.²⁷¹ And we stated that “[t]he question whether a golf cart fits into the ordinary meaning of vehicle . . . is [] a difficult one” that “turns on the viability of the sense divisions at work—on whether the golf cart is an unusual example [of the ordinary sense of vehicle] or perceived as a distinct linguistic construct.”²⁷²

²⁶⁷ Tobia, *supra* note 17, at 757.

²⁶⁸ Tobia’s “concept condition” data do not establish such a basis for all the reasons explained in Part IV.A. And the only other basis for Tobia’s conclusion is the invocation of “common sense,” *id.* at 795—a move that ought to trouble him (and troubles us) in light of the concerns he raises elsewhere about motivated reasoning, *id.* at 776, 778. There is an irony in Tobia’s analysis of this point. For all his focus on “testing ordinary meaning,” Tobia never commits to a tool for assessing the ordinary meaning of language. He simply critiques the tools that others use and falls back on the idea that the flaws in those tools suggest the need to rely more on things like context, history, legislative purpose, and “other interpretive commitments.” *Id.* at 778.

²⁶⁹ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 859.

²⁷⁰ *See id.* at 859–60.

²⁷¹ *Id.* at 859.

²⁷² *Id.* at 859–60 (emphasis omitted).

To answer this question in a reliable way, we noted that it could be helpful to develop “further corpus analysis” that could help us “assemble a list of criteria for things we speak of as an automobile,” which would then allow us to “ask whether a golf cart has those criteria.”²⁷³ To illustrate our approach, we suggested that “[p]ossible criteria” for “vehicleness” could “include a steering wheel, motor, wheels for passage on land, and seats for passengers.”²⁷⁴ We conceded that “a golf cart might count” as a vehicle “[i]f those are the criteria.”²⁷⁵ But we also suggested the possibility of other criteria, “like usual usage on paved roads or highways, or licensure by the state motor vehicle division.”²⁷⁶ And we noted that “if those are the criteria, then a golf cart might not count.”²⁷⁷

We also conceded that a complete analysis of the golf cart question could extend to the use of other experimental tools.²⁷⁸ We noted that such tools are “costly to design and implement” and are “notoriously susceptible to context effects and response bias.”²⁷⁹ But we opened the door to the possibility that “these alternative empirical linguistic methods [may] provide possible approaches to addressing questions of ordinary meaning beyond the use of corpus linguistics.”²⁸⁰ We are still of that view. We hold hope of a future in which questions about ordinary meaning are informed not just by linguistic intuition, dictionaries, and corpus tools, but also by reliable survey experiments. We see barriers and limitations to the utility of each of these tools, but see no reason to exclude any of them a priori.

In wrapping up our discussion of the “vehicle” problem in *Judging Ordinary Meaning*, we conceded the possibility that “[t]he limitations of the empirical methods” we had identified could lead to a dead end—to the conclusion “that we cannot give a conclusive answer to the question of whether the ordinary meaning of *vehicle* extends to the *golf cart*.”²⁸¹ In so doing, we emphasized the possibility that at that point the interpretive inquiry could “fall back” to other interpretive inquiries that seek to

²⁷³ *Id.* at 860.

²⁷⁴ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 860 n.257.

²⁷⁵ *Id.*

²⁷⁶ *Id.*

²⁷⁷ *Id.*

²⁷⁸ *Id.* at 860–62.

²⁷⁹ See Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 860–61.

²⁸⁰ *Id.* at 862.

²⁸¹ *Id.* (emphasis in original).

establish the “legal content” of law that may not necessarily be “in line with its communicative content.”²⁸²

We stand by that position. But we hasten to emphasize that there are ample reasons to seek first to inquire into the law’s communicative content. And we think that refined corpus methods will be an important element of this first-order inquiry. With that in mind, we close with some points of refinement, as applied to another of Tobia’s examples: the cement mixer.²⁸³

Assume the cement mixer is in fact an example of an item that does not appear in the corpus as a “vehicle.” That would not mean that it does not fall within the ordinary meaning of vehicle. It surely would count, as it would meet any of the necessary and sufficient criteria for “vehicleness.” A cement mixer has each and every one of the conditions of “vehicleness” hypothesized above. It has a steering wheel, a motor, wheels for passage on land, and seats for passengers; and it is a vehicle that is used on highways and required to be registered by a state department of motor vehicles. For those reasons it seems apparent that the nonappearance of cement mixer in the corpus would just be a “dodo bird”—or, to use a parallel example raised by Solan and Gales, a “blue pitta” (a bird, which may appear rarely, if at all, in corpora of American English or be thought of as an example of a “bird” because it is found in Asia and not North America, but has all the necessary and sufficient conditions of “birdness”).²⁸⁴ And because it seems impossible to imagine any criteria of “vehicleness” that would not sweep in the cement mixer, we can conclude that its nonappearance is beside the point and does not undermine the conclusion that it falls within the ordinary meaning of vehicle.

This suggests some needed refinements to our approach. One refinement can build on interpretive principles presented by Professor Larry Solum in his work on constitutional interpretation. Solum draws an important distinction between “original expected application[s]” of a constitutional provision and the legal principle embedded in the “[o]riginal public meaning” of the text.²⁸⁵ He clarifies that what is fixed in the law is the principle embedded in its communicative content. Expected applications, at most, have

²⁸² *Id.*

²⁸³ Tobia, *supra* note 17, at 796.

²⁸⁴ Solan & Gales, *supra* note 13, at 1315.

²⁸⁵ Lawrence B. Solum, *Triangulating Public Meaning: Corpus Linguistics, Immersion, and the Constitutional Record*, 2017 BYU L. REV. 1621, 1637 (first alteration in original) (quotation marks omitted) (“The meaning of a text is one thing; expectations about how the text will or should be applied to particular cases or issues is another.”).

evidential significance—they can help *inform* ambiguities in our interpretation of the communicative content, but they do not define or *limit* the reach of the language of law.²⁸⁶

These principles can refine our analysis of the golf cart and cement mixer questions. Through corpus analysis and otherwise, we can define the reach of the legal concept embedded in the communicative content of “vehicle”—in identifying the necessary and sufficient conditions of “vehicleness.” The corpus evidence will also help identify some expected applications of a law covering vehicles. If a golf cart or cement mixer doesn’t appear in the corpus, that may tell us that the legislative body may not have been thinking of these examples in enacting the law in question (or, perhaps, that people subject to the law would not think of these examples at first blush). But that doesn’t tell us that these examples are excluded. Expected applications have only evidential significance. They can help us define the necessary and sufficient conditions of “vehicleness,” but they do not necessarily define the full reach of the legal concept embedded in the language of the law.

Tobia seems to miss this nuance in his criticism of corpus analysis. In suggesting that corpus tools would exclude the golf cart or the cement mixer, he assumes that ordinary meaning is limited to its expected applications. That has never been our claim, nor the point of corpus analysis. But we hope this refinement will help avoid further misunderstandings of this nature.

V. ORDINARY MEANING AND FREQUENCY OF OCCURRENCE

Perhaps the most persistent criticism of the use of corpus tools in legal interpretation has been referred to as the “frequency fallacy.”²⁸⁷ This critique posits that corpus advocates merely use linguistic corpora to determine the most commonly used sense of a word, and then label that the ordinary meaning.²⁸⁸

²⁸⁶ See *id.* at 1637–38 (emphasis in original):

Thus, the framers and ratifiers of the Second Amendment may have expected that the “right to . . . bear Arms” would be applied to muskets and flintlocks, but the meaning of *arms* is more general and would encompass modern weapons. . . .

Although original expected applications do not constitute the original meaning of the constitutional text, they are nonetheless relevant to constitutional interpretation because they can provide *evidence* of the original public meaning.

²⁸⁷ See generally, e.g., Shlomo Klapper, *(Mis)Judging Ordinary Meaning: Corpus Linguistics, the Frequency Fallacy, and the Extension-Abstraction Distinction in “Ordinary Meaning” Textualism*, 8 BRIT. J. AM. LEGAL STUD. 327 (2019).

²⁸⁸ See Tobia, *supra* note 17, at 796 (defining the “Comparative Use Fallacy” as “when considering two possible senses, the comparatively greater support for one sense in the

The frequency fallacy is another straw man. It has no foundation in our writing on law and corpus linguistics. Indeed, both of us have expressly disavowed an approach that merely seeks to determine the most common sense of a word and then labels that sense the ordinary meaning.²⁸⁹ As we have said, such an approach would be arbitrary.²⁹⁰

Proponents of the frequency fallacy assert that corpus linguistics represents a “new theory about how statutes ought to be interpreted”—one that reframes the question of ordinary meaning as an empirical question that “ought to be answered by how frequently a term is used in a particular way.”²⁹¹ But we have never advocated the blind acceptance of the most common sense of a word, and the notion that courts consider frequency is not a reframing of the question of ordinary meaning. It is just one of the ways (certainly not the only way) that courts have talked about ordinary meaning. As we point out in *Judging Ordinary Meaning*, courts often frame the question of ordinary meaning in

corpus indicates that this sense is a better candidate for ordinary meaning”); Ethan J. Herenstein, *The Faulty Frequency Hypothesis: Difficulties in Operationalizing Ordinary Meaning Through Corpus Linguistics*, 70 STAN. L. REV. ONLINE 112, 113 (2017) (ascribing to us the view that “where an ambiguous term retains two plausible meanings, the ordinary meaning of the term (and the one that ought to control) is the more frequently used meaning of the term”); James A. Macleod, *Ordinary Causation: A Study in Experimental Statutory Interpretation*, 94 IND. L.J. 957, 988 (2019) (“Legal scholars drawing on corpus linguists search corpora for a given word or phrase to ascertain the frequency with which it is used in a given manner.”).

²⁸⁹ See *J.M.W. v. T.I.Z. (In re Adoption of Baby E.Z.)*, 266 P.3d 702, 726 (Utah 2011) (Lee, J., concurring in part and concurring in the judgment) (citation omitted):

I share the view that we should not blindly attribute to every statutory term its most frequent meaning. . . . Such an approach would be arbitrary and would lead to statutory incoherence. This is not the approach I have articulated, and not the one I have followed in my consideration of corpus linguistic data.

See also Mouritsen, *Definitional Fallacies*, *supra* note 1, at 1962:

My contention is not that because [carry on your person] is far more common than [carry in a car], § 924(c) ought to be interpreted with the [carry on your person] meaning. Such a reading would be arbitrary. There are undoubtedly circumstances in which Congress employs the less frequent of two senses of a word.

²⁹⁰ See *In re Adoption of Baby E.Z.*, 266 P.3d at 726 (Lee, J., concurring in part and concurring in the judgment); Mouritsen, *Definitional Fallacies*, *supra* note 1, at 1962. Here we feel some solidarity with Professors Alan Schwartz and Robert Scott. See Alan Schwartz & Robert E. Scott, *Contract Interpretation Redux*, 119 YALE L.J. 926, 932–33 (2010) (“Initially, though our article was the first cut at a difficult subject, and so was less clear than it could have been, the mistakes in representing our view are hard to explain as resulting only from a lack of clarity on our part.”).

²⁹¹ Hessick, *supra* note 16, at 1506.

terms of frequency, sometimes referring to possible, common, most common, or exclusive ways in which words are used.²⁹²

Judicial consideration of frequency is not surprising and not problematic in itself. The concept of ordinariness evokes a notion of things that more commonly occur or things that are more commonly experienced.²⁹³ Still, we agree that the “more common” formulation is not the only way that judges talk about ordinary meaning. Sometimes courts speak of “ordinary meaning” to refer to whether a meaning is “permitted,” “obvious,” or something a “reasonable person” would say or understand.²⁹⁴ Courts also “focus on whether or not it feels ‘natural’ or ‘appropriate’ to apply a given

²⁹² See Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 800; see also, e.g., *Muscarello*, 524 U.S. at 149 (Ginsburg, J., dissenting) (using a “possible” sense of ordinary when arguing that a word sense is ordinary because it is “hardly implausible nor at odds with an accepted meaning” of the statutory terms); *Taniguchi*, 566 U.S. at 569 (rejecting a sense as ordinary when it was merely a “possible” meaning but “hardly a common or ordinary meaning”); *Muscarello*, 524 U.S. at 128–29, 131 (employing a “common” sense of ordinary when asserting that the transport-in-a-vehicle sense of “carry” is ordinary because “many”—“perhaps more than one-third”—of the instances of “carrying a firearm” in the *New York Times* and *U.S. News* databases reflect that sense, and because the ordinary sense of “‘carry’ includes conveyance in a vehicle”) (emphasis added); *id.* at 128 (referring to a “most common” sense of ordinary when reasoning that “we believe Congress intended to use the word in its primary sense and not in this latter, special way”); *id.* at 143 (Ginsburg, J., dissenting) (speaking in terms of the “most common” sense of ordinary by referring to “what meaning showed up some two-thirds of the time” as opposed to the alternative sense which showed up “‘more than one-third’ of the time”).

²⁹³ See WEBSTER’S THIRD NEW INTERNATIONAL DICTIONARY 1589 (3d ed. 2002) (defining “ordinary” as “being of frequent occurrence”); THE AMERICAN HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE 1241 (5th ed. 2011) (defining “ordinary” as “[c]ommonly encountered; usual”); THE RANDOM HOUSE DICTIONARY OF THE ENGLISH LANGUAGE 1363 (2nd ed. 1987) (defining “ordinary” as “the commonplace or average condition, degree, etc.”). Professor Hessick insists that use of dictionaries by corpus advocates is somehow “almost ironic,” Hessick, *supra* note 16, at 1508 n.16, but as we have repeatedly made clear, we have no objections to the use of dictionaries for defining terms or highlighting a range of possible uses. See Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 826 (explaining that dictionaries “can be useful for defining unknown terms and attesting contested uses”); *In re Adoption of Baby E.Z.*, 266 P.3d at 729 (Lee, J., concurring in part and concurring in the judgment) (“I have no problem citing dictionaries for the information that they *do* contain. Dictionaries may help the court by defining unknown terms or presenting a range of *possible* meanings that a term may bear in a given context.” (emphasis in original)); *State v. Rasabout*, 356 P.3d 1258, 1272 (Utah 2015) (Lee, A.C.J., concurring in part and concurring in the judgment) (noting that dictionaries are “useful in cataloging a range of possible meanings that a statutory term may bear” (quoting *Hi-County Prop. Rts. Grp. v. Emmer*, 304 P.3d 851, 856 (Utah 2013))). Our objection is the reliance on dictionaries to determine ordinary meaning.

²⁹⁴ Hessick, *supra* note 16, at 1507. Though we note that we disagree with Hessick’s assertion that “[n]one of these usages is empirical—or at least they are not readily quantifiable.” *Id.* We believe that there are a number of ways that linguists might test whether a meaning is obvious or how it would be understood by a reasonable person.

term or phrase to a particular example.”²⁹⁵ And we agree that judges sometimes appear to be invoking a concept of ordinary meaning that is similar to the linguistic notion of a prototype.²⁹⁶

We disagree, however, with any suggestion that judges (or even any given judge) use any of these notions of ordinariness exclusively. None of these conceptions of ordinary meaning should be taken as *the* definitive notion of ordinary meaning. The point of our discussion of different conceptions of ordinary meaning is to highlight the fact that, “ironically, we have no ordinary meaning of ‘ordinary meaning.’”²⁹⁷ Judges and lawyers lack a shared, coherent, and workable understanding of this concept. Sometimes individual judges contradict themselves within a single opinion.²⁹⁸ And often it is not obvious that two closely related uses of a word are in fact different senses, as judges often assume.

While we do not endorse the view that blindly attributes to each word its most frequent sense, we think the frequency assessment should play a role in the interpretation of legal texts. We have said 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 [including the syntactic, semantic, and pragmatic context in which an utterance occurs], its historical usage and the speech community in which it was uttered.²⁹⁹

And if all of these factors are to be taken into account, we will need a method to gather evidence of such usage.

This view of the role of context permeates *Judging Ordinary Meaning*, and its application is illustrated in several of the examples in the paper. In assessing the “carries a firearm” question in *Muscarello*, for example, we spoke at length of the need to assess not just the frequency of different senses of the verb “carry,” but also to consider the use of the verb “carry” in the context of

²⁹⁵ Macleod, *supra* note 288, at 990.

²⁹⁶ See Lawrence M. Solan, *Why Laws Work Pretty Well, but Not Great: Words and Rules in Legal Interpretation*, 26 LAW & SOC. INQUIRY 243, 258 (2001) (“Some Supreme Court cases concerning statutory interpretation can be seen as battles among the justices over definitions versus prototypes.”).

²⁹⁷ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 798.

²⁹⁸ See, e.g., *Muscarello*, 524 U.S. at 128–31 (togglng between the “common” and “most common” senses of ordinary in the majority opinion); *id.* at 143–44, 149 (Ginsburg, J., dissenting) (togglng between the “possible” and “most common” sense of ordinary in the dissent); *supra* text accompanying note 292.

²⁹⁹ Lee & Mouritsen, *Corpus Linguistics*, *supra* note 4 (emphasis added).

“firearm” and its synonyms by looking at concordance lines in which the carrying is being done by a human subject, and to make this examination at a given time and with respect to an appropriate language community.³⁰⁰ Corpus tools allow us to do all of that. They provide evidence of the syntactic argument structure in which the word occurs, the semantic features of a given utterance, the linguistic conventions prevailing at the time that the words of a legal text were first executed, and the linguistic conventions prevailing in the relevant speech community, genre, and register in which the words are used. This is important because words can take on different meanings when used with different inflections, in different parts of speech, or when they merely co-occur with different collocates. And existing tools (dictionaries, intuition, etc.) do not allow for any of this kind of assessment. So unless we are going to abandon the search for ordinary meaning altogether, we need the new tools.

Responding to this more nuanced framework, Professor Hessick has argued that “everyone agrees that context is important.”³⁰¹ But mutual agreement about the importance of context is irrelevant where courts lack a shared understanding of what context means or a viable means of examining language usage in those contexts.

Our contention is not that corpus linguistics will provide push-button answers to difficult questions of legal interpretation simply by highlighting the most common sense of a word. Instead, language evidence from linguistic corpora can help us give content to an otherwise vague and empty term like ordinary meaning, by providing evidence of the way words and phrases are used in particularized contexts, in particular speech communities or linguistic registers, and at particular times. Once jurists have a more accurate picture of prevailing linguistic conventions, they will have to make difficult jurisprudential decisions about what language evidence is useful and how such evidence can inform (and perhaps modify) interpretive practices.

Corpus tools do not supplant a judge’s experience, training, or professional judgment. They are a check against the judge’s linguistic intuition. Hessick offered a puzzling retort. She said that “[t]he idea that the traditional ordinary meaning inquiry is somehow inferior to a frequency analysis because judges might rely on

³⁰⁰ See Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 845–48.

³⁰¹ Hessick, *supra* note 16, at 1514 n.43.

their intuition—that is, their professional judgment—rejects the very foundation of the judicial role.”³⁰² There is a lot to unpack in this sentence. To begin with, Hessick used the phrase “traditional ordinary meaning inquiry” as if judges and lawyers have some sort of shared, coherent understanding of what “ordinary meaning” actually means. They don’t.³⁰³ Hessick and others have implicitly acknowledged this very problem by highlighting ways in which courts describe their search for ordinary meaning.³⁰⁴

More importantly, Hessick’s statement conflated (and incorrectly suggested that we conflate) a judge’s “intuition” with a judge’s “professional judgment.” There is no foundation in any of our writing for this move. We are in no way against the judge’s reliance on her professional judgment—informed by her training, education, or experience—in judging ordinary meaning. Our concerns about “intuition” are directed to the judge’s “linguistic intuition,” not her “professional judgment.”³⁰⁵ Judges, like all human

³⁰² *Id.* at 1511.

³⁰³ See Clark D. Cunningham, Judith N. Levi, Georgia M. Green & Jeffrey P. Kaplan, *Plain Meaning and Hard Cases*, 103 YALE L.J. 1561, 1563–65 (1994) (explaining that “the phrase ‘plain meaning’ itself presents interpretive difficulties” and discussing various, conflicting senses of the phrase); see also Sandra F. Sperino, *Flying Without a Statutory Basis: Why McDonnell Douglas Is Not Justified by Any Statutory Construction Methodology*, 43 HOUS. L. REV. 743, 764–65 (2006) (“While the term ‘plain meaning’ exudes a sense of simplicity, such an assumption would be misplaced because the exact contours of plain meaning interpretation are debated.”); Richard J. Lazarus & Claudia M. Newman, *City of Chicago v. Environmental Defense Fund: Searching for Plain Meaning in Unambiguous Ambiguity*, 4 N.Y.U. ENVTL. L.J. 1, 15 (1995) (“[T]he definition of ‘plain meaning’ is itself anything but plain. How much ambiguity is required before the meaning of a provision becomes ambiguous? Words are hardly ever entirely free of ambiguity and there is almost always room for disagreement based on at least plausible readings.”).

³⁰⁴ See Hessick, *supra* note 16, at 1507 (“Sometimes courts use the term ‘ordinary meaning’ to refer to whether a meaning is permitted, sometimes to refer to whether the meaning is obvious, and sometimes to refer to the meaning that the hypothetical reasonable person would give to the statutory language.” (citations omitted)). A “permissible” meaning and an “obvious” meaning are not necessarily the same thing. A meaning can be permissible without being obvious. See also Macleod, *supra* note 288, at 990 (“[C]ourts’ ordinary meaning analysis tends to focus on whether or not it feels ‘natural’ or ‘appropriate’ to apply a given term or phrase to a particular example.”).

³⁰⁵ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 806 (“Typically, this assessment is made at a gut level, on the basis of a judge’s *linguistic intuition*, without recognition of the empirical nature of the question.” (emphasis added)); *id.* at 831 (“Linguistic corpora can perform a variety of tasks that cannot be performed by *human linguistic intuition* alone.” (emphasis added)); *id.* at 857 (“Our *linguistic intuitions* about usage and meaning in our own time and our own speech community can be highly unreliable. But this problem is amplified when we are interpreting a text that dates from a period of which we have no linguistic memory or experience.” (emphasis in original)); 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, 175 (2011) (citation omitted):

beings, have a language faculty that has certain limitations. One of those limitations is that a great deal of information regarding language use is “not susceptible to recovery via introspection.”³⁰⁶ This is in part because, “[i]n many cases, humans tend to notice unusual occurrences more than typical occurrences.”³⁰⁷ Judges and lawyers, like all language users, may not be particularly adept at objectively and predictably identifying and resolving lexical ambiguities when faced with high-frequency, highly polysemous words—words that occur often and that have a lot of different, related senses.³⁰⁸ This is a problem because word frequency is correlated with polysemy—the more commonly a word is used, the more likely it is to have many different senses, and the more senses it has, the more likely two people are to disagree as to its meaning in a given context.³⁰⁹ Judges and lawyers have no special immunity from this and other linguistic limitations. And our argument is simply that linguistic corpora may provide language evidence through which judges and lawyers can test their intuitions about the meaning of a legal text.

Another criticism is that linguistic corpora may not reflect ordinary usage, but instead will “reflect the prevalence or newsworthiness of the underlying phenomenon that the term

With regard to linguistic intuition, a judge is “as liable to be as deviant as the next man.” . . . [J]udges are subject to the same linguistic limitations as the rest of us, which limitations include the inability to intuit which features of the language are common or ordinary and which are unusual.

See also id. at 178 (“This is not to say that a judge’s linguistic intuition is useless. Human experience with language allows the judge—like any other language user—to recognize almost instantly which uses of a given term are grammatically correct and which are not.”); *id.* at 204 (“Against this backdrop, the corpus methodology presents an attractive alternative [] to the judge’s sometimes unreliable linguistic intuition.”).

³⁰⁶ TONY MCENERY & ANDREW WILSON, *CORPUS LINGUISTICS* 15 (2d ed. 2001).

³⁰⁷ BIBER, CONRAD & REPPEN, *supra* note 50, at 3.

³⁰⁸ *See* Scott Crossley, Tom Salsbury & Danielle McNamara, *The Development of Polysemy and Frequency Use in English Second Language Speakers*, 60 *LANG. LEARNING* 573, 575 (2010).

³⁰⁹ *See id.* at 576; *see also* Martha Palmer, Hwee Tou Ng & Hoa Trang Dang, *Evaluation of WSD Systems*, in *WORD SENSE DISAMBIGUATION: ALGORITHMS AND APPLICATIONS* 75, 91 (Eneko Agirre & Philip Edmonds eds., 2007) (noting that “[h]igh polysemy has a detrimental effect” on the performance of disambiguation tasks); George Tsatsaronis, Iraklis Varlamis & Kjetil Nørvåg, *An Experimental Study on Unsupervised Graph-Based Word Sense Disambiguation*, in *COMPUTATIONAL LINGUISTICS AND INTELLIGENT TEXT PROCESSING* 184, 193 (Alexander Gelbukh ed., 2010) (noting that human annotators have higher rates of disagreement when tasked with disambiguating highly polysemous words); *CONCISE ENCYCLOPEDIA OF SEMANTICS* 224 (Keith Brown & Keith Allan eds., 2009) (noting that accuracy on word sense disambiguation tasks declines where finer-grained sense distinctions are required).

denotes.”³¹⁰ Hessick offers the following thought experiment to illustrate this criticism:

Imagine . . . a dispute over the scope of a statute that provides relief for flood victims. The dispute centers around how much water must have accumulated in an area in order for an event to be considered a flood. A database search of how often the word “flood” is used to refer to very large amounts of accumulated water will doubtlessly be skewed by the fact that instances of extensive flooding—such as New Orleans after Hurricane Katrina in 2005 or Houston during Hurricane Harvey in 2017—will receive far more media coverage than other events. Indeed, a corpus analysis may demonstrate that seventy percent of all mentions of the word “flood” occur in the context of these superstorms. But that does not tell us whether the average American would understand the statutory term “flood” to include three inches of water in a homeowner’s basement after a neighboring water main burst.³¹¹

This illustration does more to highlight the utility of corpus tools than it does to undermine that utility.

Implicit in the illustration is the notion that the average American’s understanding of the word “flood” would differ meaningfully from the word’s most salient, newsworthy usage at a given time. This conclusion is possible, but not obvious. After all, average Americans (and average American judges) are people, and people view the world through the lens of a variety of systematic biases. One of these biases is the availability heuristic, which is that people believe that events are more probable if examples are easier to remember.³¹²

It is not difficult to imagine the ways in which the availability bias might color the way that average Americans and judges alike think about statutory terms. When a salient newsworthy event like Hurricane Katrina or Hurricane Harvey occurs, our average American speaker will be bombarded with nightly news

³¹⁰ Herenstein, *supra* note 288, at 122; *see also* Macleod, *supra* note 288, at 988 (“Legal scholars drawing on corpus linguists search corpora for a given word or phrase to ascertain the frequency with which it is used in a given manner.”).

³¹¹ Hessick, *supra* note 16, at 1509.

³¹² NICK WILKINSON & MATTHIAS KLAES, AN INTRODUCTION TO BEHAVIORAL ECONOMICS 120 (3d ed. 2017) (“The main source of error with the availability heuristic is salience; this factor features in other types of bias also, but the main effect here is that events that have been well publicized or are prominent in people’s memories tend to be estimated as having exaggerated probabilities.” (emphasis omitted)).

broadcasts, radio programs, podcasts, and watercooler talk about the event. In the wake of a recent event of this nature, it is not hard to imagine that individuals' judgments about the meaning of a statutory term could be shaped by an event that finds its way into contemporary speech through a variety of vectors. These same individuals may not be consciously aware of the ways that recent salient events have shaped their perceptions of the meaning of "flood," and they will no longer have introspective access to their prior perceptions of meaning.

It is in this context that the examination of language evidence in a corpus can aid the intuition of the interpreter. The interpreter can examine the use of the word "flood" through time to determine whether the use of the word has changed over time (and, in particular, whether the use of the word has changed since the time of the passage of the imaginary statute in question). Analyzing how hydrological events are reflected in historical language usage is precisely one of the things that corpus evidence has been shown to be good at.³¹³

Using comparative corpus-based methods, an interpreter can examine the use of a word across language genres, registers, and speech communities to determine whether the word has taken on some specialized meaning in a relevant speech community, register, or genre. Indeed, using comparative corpora, the interpreter can check against the very problem the illustration is meant to highlight—determining whether a given sense of a word is overrepresented in newsprint, but less likely to be used in other contexts. In addition, though no sample text is provided for the hypothetical statute, examining the use of the text in contexts similar to the statutory context may reveal usage patterns that are not obviously available via introspection. Thus, the corpus can help address precisely one of the problems the illustration is meant to identify—providing a check against the outsized influence that real-world salience may have on judgments about ordinary meaning.

It is difficult to imagine how the "flood" illustration maps onto any real-world interpretive scenario. The illustration posits the ineffectiveness of corpus evidence without providing more than a single word from the imagined statute. Again, this highlights a

³¹³ See generally Tony McEnery, Helen Baker & Carmen Dayrell, *Working at the Interface of Hydrology and Corpus Linguistics: Using Corpora to Identify Droughts in Nineteenth-Century Britain*, in *USING CORPUS METHODS TO TRIANGULATE LINGUISTIC ANALYSIS* 52 (Jesse Egbert & Paul Baker eds., 2019).

main advantage of corpus analysis—that it allows us to examine uses of words and phrases in particular syntactic, semantic, and pragmatic contexts (and at particular times and in particular speech communities).

Finally, as we discussed in *Judging Ordinary Meaning*, there is a difference between addressing ambiguity on the one hand and vagueness on the other in legal interpretation. With ambiguity, two or more senses of a given word may be possible in a given context. Vagueness, by contrast, is a question of scope. We can make the case that “carry on your person” and “carry in a car” are two different senses of “carry.” But a “big flood” and a “little flood” are both floods. Not surprisingly, the use of the word “flood” to describe basement flooding or flooding from a burst pipe is very well attested, even in a corpus of newspapers and magazines.³¹⁴ Without more statutory context, it is impossible to predict how this information could be used in determining how the statute ought to be interpreted. But as discussed above, we have never advocated for the blind application of the most frequent use of a word as the ordinary meaning.

CONCLUSION

Professor Tobia asserts that his critiques “shift the argumentative burden” to those who advocate the use of corpus tools in statutory interpretation.³¹⁵ He also challenges us to show that our methods can produce “a nonarbitrary and reliable method of interpretation.”³¹⁶ This theme runs through each of the other criticisms leveled to date. And it falters on two fundamental grounds.

The first problem is that the starting premise is largely based on a misstatement of our proposed methods. For the most part the prevailing critiques serve only to further underscore basic problems with existing methods of assessing the communicative content of the language of law and to actively highlight core contributions (not shortcomings) of corpus methods. Corpus tools are better able to address concerns about judging the right language community. They also address concerns about fair notice. And they have not been shown to be inaccurate or to falter on a supposed nonappearance or frequency fallacy. On closer review, these

³¹⁴ For example, “flood” appears within 5 words of “basement” in 234 codable concordance lines in the NOW Corpus, which compiles web-based newspapers and magazines. NOW CORPUS (NEWS ON THE WEB), <https://www.english-corpora.org/now/>.

³¹⁵ Tobia, *supra* note 17, at 805.

³¹⁶ *Id.*

criticisms help highlight salient features of corpus analysis. The burden has not been shifted.

A second problem is in our critics' failure to identify a framework or even any methods for a better alternative. No one can credibly contend for an interpretive regime in which we ignore the threshold inquiry into the ordinary communicative content of legal text. The search for such content is too embedded in our very concept of written law—and too central to too many of the policy premises that underlie it—to think that we could or should ever avoid it. So no one seriously proposes that move. And in the absence of such a proposal the onus is on our critics to clarify how they propose to paint the “standard picture”—whether they prefer to retain our current set of underdeterminate tools or, if not, can proffer something that better accounts for the shortcomings of the existing regime. None of them have done that.

We think the criticisms fail for all sorts of reasons explained in detail above. But even if some of the pushback stands as highlighting imperfections in what we have proposed (and there certainly are some wrinkles to iron out), the onus is on those who think that our methods are still not ideal to show how they could perform this function better. It takes a method to beat a method. And one of our core points has been to show that our methods account for shortcomings in existing methods and do a better job of assessing ordinary communicative content in a more transparent, reliable way.

None of our critics offer a framework for assessing the communicative content of difficult cases of lexical ambiguity—cases like *Muscarello*, *Taniguchi*, and *Mont*. And without a replacement method for resolving these kinds of cases, the utility of such criticism is limited.

Tobia comes closest to offering a substitute framework—or at least appears at first glance to do so. But even he isn't ultimately suggesting that his survey methods should take the place of corpus methods (or existing tools). He is just claiming that his surveys show that corpus tools are imperfect. And from there his only move is to suggest that we ought to fall back instead on tools that can help us discern statutory purpose.³¹⁷ That move just takes us back to square one, however, since all agree that the best

³¹⁷ See *id.* at 804 (“Legal interpreters will have to look beyond the simple dictionary definition and corpus frequency analysis—to the legal text’s context, history, and purpose; and to their other interpretive commitments.”).

indicator of purpose is in the statutory text³¹⁸ and any other indicator (like legislative history) will still require a framework and toolkit for interpreting language (since legislative history is also comprised of text).

All that said, the pushback advanced by our critics is still constructive. It helps highlight the ongoing need for us to develop and refine the proper use of corpus tools in statutory interpretation. To highlight some possible steps in that direction we will close with some thoughts on refinements in corpus linguistic analysis of the types of problems raised in cases like *Muscarello*, *Taniguchi*, and *Mont*.

One difficulty presented by cases like those discussed above is the “closely related senses” problem—the fact that the competing notions of “carry” (transport in a car versus personally bear), “interpreter” (real-time oral translator versus written-text translator), and “imprisonment in connection with a conviction” (after conviction versus before but later counted toward full term) can either be viewed as two distinct senses or just a general sense that encompasses more specific variations on a theme. An alternative framing of this problem is the idea that “corpus data may reflect only the fact that a given sense of a certain term is a more factually common iteration of that term in the real world.”³¹⁹ This is a serious problem. And uncareful use of corpus evidence in cases like these can gloss over the problem in a manner that may undermine its key selling points. We should instead be careful. And we close with some ideas on how best to do so.

Before doing so, it is important to start by noting that this problem is not just a problem for those who turn to corpus tools. It is a problem for any theory or approach to interpretation. Corpus tools don’t create this problem. They just help refine our perception of ordinary meaning in a way that helps us see a complexity or nuance that we had previously missed. And the Court did miss the complexity in each of the cited cases. The justices’ competing views on “carry,” “interpreter,” and “imprisonment” skate

³¹⁸ See Abbe R. Gluck, *The States as Laboratories of Statutory Interpretation: Methodological Consensus and the New Modified Textualism*, 119 YALE L.J. 1750, 1756–58 (2010) (concluding, based on a comprehensive study of state court approaches to statutory interpretation, that state courts are engaged in an “effort[] to increase predictability in statutory interpretation,” and that they give primacy to text and decline to look to external sources of meaning if they find the text “plain”).

³¹⁹ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 873–74 (summarizing a concern by Solan and Gales, since published in Solan & Gales, *supra* note 13, at 1351); see also Klapper, *supra* note 287, at 349.

past the question whether they are really talking about separate senses of these terms. In insisting on competing views of “imprisonment” in *Mont*, for example, the majority and dissent don’t just have two different senses of ordinary meaning in mind (“permissible” versus “natural” sense of the term).³²⁰ They each assume that these are separate, competing senses. To the extent that is debatable, the debate is one that all interpreters must confront.

We think corpus tools can best facilitate a meaningful analysis of that question. They can do so in two ways. First, corpus analysis can provide a transparent view of the range of options available to the interpreter—options that implicate a choice for the chosen theory of interpretation. By analyzing a random set of concordance lines from selected corpora, the interpreter can gain access to important information that an interpreter would otherwise have to guess at. We highlight this contribution by outlining proposed steps for corpus analysis of the interpretive question in *Mont*—as to the ordinary communicative content of “imprisonment in connection with a conviction.”

Step one is to choose an appropriate corpus—a corpus that can help us test the relevant language community. This is a threshold question for our legal theory of interpretation. And it is implicit in the conflict between the majority and dissent in *Mont*. One difference between the two opinions is in the language community they have in mind. The majority seems to be seeking to assess public meaning, citing dictionary definitions of “imprison” in support of its conclusion that this verb “encompass[es] pretrial detention.”³²¹ And the dissent’s disagreement stems, at least in part, on its focus on Congress’s meaning. It cites provisions of the U.S. Code in support of its view that “Congress regularly uses the word ‘imprisoned’ (or ‘imprisonment’) to refer to a prison term following a conviction.”³²²

Corpus analysis can help refine this debate. Instead of just resorting to selected dictionary definitions we can assemble a randomized set of concordance lines that can help us assemble transparent evidence on how the term “imprison” is used by the general public. And we can also systematize the inquiry that the

³²⁰ *Compare Mont*, 139 S. Ct. at 1832 (“[T]he definition of ‘is imprisoned’ may well include pretrial detention.” (emphasis added)), *with id.* at 1838 (Sotomayor, J., dissenting) (“[I]mprisoned’ [] is most naturally understood in context to mean postconviction incarceration.”).

³²¹ *Id.* at 1832.

³²² *Id.* at 1838 (Sotomayor, J., dissenting).

dissent has in mind. By searching a corpus of statutory language we can access systematic evidence to inform our inquiry into how Congress uses “imprison” and “imprisonment.” The choice between these two types of corpora is a question for our theory of interpretation. We take no position on the matter here, except to emphasize that corpus tools can help refine this inquiry and make it more transparent.

A second step is to consider and code the randomized concordance lines we assemble from our chosen corpora. In so doing we can further inform the debate that is otherwise taking place at the “take my word for it” level. Instead of just insisting that “imprison” “encompass[es] pretrial detention,”³²³ on one hand, or that the phrase “is imprisoned” “is most naturally understood in context to mean postconviction incarceration,”³²⁴ we can assemble transparent evidence from natural language usage to gauge the reliability of these assertions.

Our analysis of “is imprisoned” shows that this term is most often associated with a term of incarceration imposed after entry of a conviction. Often this use of the statutory phrase isn’t apparent from the context of the concordance line. But where there is enough context to tell, the reference is almost always in connection with incarceration after conviction.

The last step is to determine the significance of the corpus evidence. Here we return to the “closely related senses” problem. This is a key, unresolved problem in law and corpus linguistics. And it is implicated in *Mont*, as it is possible to view the competing senses of “imprison” as either two separate senses or just two alternative examples of a single sense.

In *Judging Ordinary Meaning*, we offered a starting point for a response to this problem. We suggested that even where two senses of a statutory term are closely related, the fact that one of them is overrepresented in a corpus may tell us something important—that that sense is “likely to be the one that first comes to mind when we think of this term.”³²⁵ We conceded that this “top-of-mind sense . . . may not exhaust the breadth of human perception of th[e] term,” since on reflection, “some people might concede that the term encompasses” other, less common examples.³²⁶ And we noted that the choice between the two senses (“top-

³²³ *Id.* at 1832 (majority opinion).

³²⁴ *Id.* at 1838 (Sotomayor, J., dissenting).

³²⁵ Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 874.

³²⁶ *Id.*

of-mind” versus broader, “reflective” sense) “is not a deficiency in corpus data” but a problem for our legal theory of interpretation.³²⁷ The choice will be “dictated in part by the rationales that drive us to consider ordinary meaning”—“[a] concern for fair notice and protection of reliance interests may [] direct us to stop at the top-of-mind sense of a statutory term,” while other rationales could press us to endorse a broader, reflective sense.³²⁸

These principles are a good starting point for a corpus-based analysis of the *Mont* case. Corpus evidence suggests that the top-of-mind sense of “imprisonment” is in connection with post-conviction incarceration. Because that is the most frequently attested sense of this term, there are notice-based and reliance-based reasons to limit the term to that application. But there is nothing in corpus analysis that requires that outcome. The term “imprisonment” can certainly be used to refer to any form of detention. And it is fairly easy to find a rationale for extending the statutory term more broadly—a rationale, for example, based on presumed congressional intent.³²⁹

We stand by these points but also wish to offer some further refinements. Although we still maintain that the threshold question for the closely related senses problem is a matter for legal theory, we think linguistic methods can further enhance the debate. We highlight two possibilities here. One is a means of analysis proposed by Professors Solan and Gales, which they refer to as “double dissociation.”³³⁰ The starting point is the notion that “the strongest cases for using corpus analysis are ones in which not only does one meaning predominate over an alternative meaning in an appropriate corpus, but the second, less common meaning is generally expressed using language other than the language in the disputed statute.” Solan and Gales propose to test the inference from corpus analysis by using corpus tools to assess the language we use when expressing the alternative sense of a statutory term.³³¹ If the alternative sense “is generally expressed using language other than the language in the disputed statute,” they say

³²⁷ *Id.*

³²⁸ *Id.*

³²⁹ *See Mont*, 139 S. Ct. at 1832 (noting that “Congress, like most States, instructs courts calculating a term of imprisonment to credit pretrial detention as time served on a subsequent conviction,” and concluding that it thus “makes sense that the phrase ‘imprison[ment] in connection with a conviction’ would include pretrial detention later credited as time served” (alteration in original)).

³³⁰ Solan & Gales, *supra* note 13, at 1315.

³³¹ *See id.* at 1353.

this helps dissociate the alternative sense from the language of the statute.³³²

We think this is an important next step in corpus analysis. And we can use it to help reinforce a possible inference from the corpus evidence of “imprisonment.” Corpus analysis shows that when we speak of pretrial detention, we rarely use the terms “imprison” or “imprisonment.” Most often we use other terms like detain, detention, jail, or jailed.

We have performed some corpus analysis to support this conclusion. First we compiled a list of synonyms of “imprison,” based on lists from various thesauruses. Our list was “incarcerate,” “confine,” “detain,” and “jail.” Next we reviewed two hundred concordance lines for “imprison” and each of the synonyms. We used this analysis to determine how often each term refers to pretrial detention. We multiplied the frequency with which each term refers to pretrial detention by the term’s frequency in the corpus: the resulting frequency for “imprison,” for example, would show how commonly the corpus includes the term “imprison” when it communicates the meaning of “pretrial detention” (e.g., if 10% of words in the corpus were “imprison,” and imprison refers to pretrial detention 50% of the time, then 5% of the words in the corpus are “imprison” referring to pretrial detention). By adding the resulting product for all the synonyms, we predicted roughly how often any term in the corpus refers to “pretrial detention” (the prediction is a little less than the real figure, because some non-synonym terms might refer to pretrial detention a few times). We then divided the percent of the time that the term “imprison” refers to pretrial detention by the percent of the time that any term refers to pretrial detention. The resulting calculation indicates what percent of the time the term “imprison” is used when a term in the corpus refers to pretrial detention. If the corpus shows that “imprison” is rarely used when referring to pretrial detention, then it follows that the speaker (here, the legislative drafter) probably wouldn’t have chosen “imprison” if meaning to refer to pretrial detention.

We performed this analysis using both COCA and a corpus comprised of all statutes in the U.S. Code.³³³ Our corpus analysis of COCA showed that at most 28% of references to detention without conviction used the term “imprison.” And in the U.S. Code, we

³³² *Id.* at 1315, 1353–54.

³³³ See *Corpus of the Current U.S. Code*, BYU L. & CORPUS LINGUISTICS <https://lnc12.byu.edu/>.

found that “imprisoned” referred to detention without conviction 0% of the time—none of the references to detention without conviction in the U.S. Code used the term “imprison.”

This evidence lends support to the view that “imprison” is not frequently used to refer to detention without conviction in the U.S. Code. But in COCA, reflecting ordinary language usage, “imprison” refers to detention without conviction 28% of the time—essentially just as often as the other two words (“confined” and “detained”) that commonly refer to detention without conviction. Because other terms are much more often used to refer to this sort of detention, the corpus evidence also suggests that the public may not understand “imprison” to refer to detention without conviction. We may thus expect the general public to sometimes use “imprison” when they intend to refer to pretrial detention.³³⁴

This corpus evidence is helpful for assessing the question presented in *Mont*. But even this kind of evidence would not necessarily be conclusive. There still remain underlying questions for our legal theory of interpretation—questions as to the proper language community, and for what to make of this usage evidence. Remember that the *Mont* majority seemingly was asking only what “imprison” can permissibly mean in the context of this statute. And the corpus evidence doesn’t ultimately disprove that; in fact, it confirms it. What the corpus evidence does tell us, however, is that “imprison” is a very unusual word choice for someone who is seeking to refer to pre-conviction detainment—and perhaps a word choice that a drafter would not have used if the intent was to cover that sort of detainment, or a word choice that the general public would not be thinking of in relying on the language of the law. And transparent analysis of this kind of evidence can help refine questions for our legal theory to resolve.

Double-dissociation analysis thus gets us closer to painting a full picture of the questions that lie at the threshold of our ordinary meaning inquiry. And we think that further development of this point can help make corpus analysis even more useful.

A final refinement could be in integrating survey-based inquiries with corpus analysis. Above we identified a range of concerns with survey instruments. We stand by those concerns. But we also find room for optimism about refinements that could be made to survey methods that could help address the problems that we have identified. As noted above, corpus tools are a good

³³⁴ See Appendix for tables presenting some details of our analysis.

way to gather evidence of natural language use—language use that takes place in natural linguistics settings and is not elicited for the purposes of study through a survey or interview. There are reasons to prefer such evidence, but also some limitations. A corpus may not have a sufficient number of examples of the language use in question. And a corpus may not allow us to gather information about the pragmatic context of the communication—the social or spatial context in which the communication occurred. An advantage of survey methods is that the survey prompts may be designed to more precisely take into account the syntactic and pragmatic context of the communication under analysis. This was demonstrated in a recent paper by Shlomo Klapper, Soren Schmidt, and Tor Tarantola, in which survey respondents were presented not only with prompts concerning the meaning of a handful of words that are relevant to important Supreme Court cases but also the factual background of those cases and precise language of the statute in question.³³⁵ We take exception to some of the claims in existing work on using surveys in legal interpretation. The purported strength of survey methods can be a key weakness—when we give survey respondents extensive detail about the nature of a legal dispute (a civil case involving an insurance company, or a criminal case involving allegations of a violent crime), the results may not give us linguistic information at all—just a reaction to respondents’ priors about the litigants or preference for a perceived “fair” outcome.³³⁶ That said, we agree that there may be ways in which surveys can be useful in evaluating claims about the meaning of legal texts, and there may be information about language use and perception that a survey can provide and that can’t be obtained from a corpus.

We are open to these and other possibilities. Although we see inherent shortcomings of survey results, we concede that our corpus methods are also imperfect. Going forward, we envision a framework in which the search for ordinary meaning is informed by a sort of triangulation, in which corpus evidence, survey results, and purpose-based evidence are all brought to bear.

Corpus analysis, by the way, could also help refine the inquiry into legislative purpose. Our methods acknowledge the

³³⁵ See generally Shlomo Klapper, Soren Schmidt & Tor Tarantola, *Ordinary Meaning from Ordinary People* (July 17, 2020) (unpublished manuscript) (<https://perma.cc/ULB4-N5C5>).

³³⁶ See Ben-Shahar & Strahilevitz, *supra* note 226, at 1780 (advancing the case for the use of surveys in contract interpretation but acknowledging this concern).

need to consider pragmatic context in assessing ordinary meaning.³³⁷ Such context clearly encompasses legislative purpose. One of the biggest problems with the inquiry into such purpose is the cherry-picking concern—the “looking over a crowd and picking out your friends” problem.³³⁸ And corpus tools can help deal with that problem. A legislative history corpus is already under construction. By searching such a corpus, an analyst could assemble systematic evidence of language usage in the legislative body in a more transparent attempt to assess legislative purpose.

Such an approach could help refine the *Mont* debate even further. Perhaps survey methods, properly controlled, could help further inform our understanding of the ordinary meaning of “imprisonment.” And a corpus analysis of the U.S Code Corpus and a legislative history corpus could help systematize the inquiry into legislative purpose.

We think this is the future of statutory interpretation. We think corpus tools will be an important part of that future.

³³⁷ See Lee & Mouritsen, *Judging Ordinary Meaning*, *supra* note 7, at 823–24.

³³⁸ Wald, *supra* note 100, at 214.

APPENDIX

In this Appendix, we describe in more detail the corpus analysis for *United States v. Mont* (described more briefly on pages 64–66). Notably, this analysis went beyond the concordance line analysis we showcased in *Judging Ordinary Meaning*; we also engaged in double-dissociation analysis. We hope that this Appendix can show how others can utilize double-dissociation analysis (and can invite criticism to improve our methods).

A.I Corpus Analysis for *Mont*

Our first step was to code randomized concordance lines from our chosen corpora for the operative term, “imprisoned.” We reviewed two hundred randomized concordance lines from COCA and the Corpus of the Current U.S. Code.³³⁹ The following meanings emerged:

TABLE I: MEANING FOR “IMPRISONED” IN THE U.S. CODE AND COCA

Meaning	U.S. Code	COCA
Detainment post-conviction	195 (97.5%)	30 (15%)
Detainment without conviction	0	55 (27.5%)
Detainment generally (unclear if the term implies conviction)	1 (0.5%)	33 (16.5%)
Detainment in a non-U.S. legal system ³⁴⁰	4 (2%)	57 (28.5%)
Unclear	0	4 (2%)
Other (including metaphorical uses)	0	21 (10.5%)
Total lines	200	200

³³⁹ For COCA, we limited our search to the 1990s. Although Congress enacted the statute at issue in *Mont* in 1986, see Criminal Law and Procedure Technical Amendments Act of 1986, 99 Pub. L. 99-646, 100 Stat. 3592, we deemed it unlikely that a term like “imprisoned” would change in meaning between 1986 and the turn of the millennium.

³⁴⁰ We decided to distinguish the uses of “imprisoned” referring to a foreign legal system. While we could have researched foreign legal systems to determine whether and when they “convicted” (or similar), that inquiry was beyond the scope of this Article. Moreover, it’s unclear whether “imprisoned” referring to other countries’ penal systems necessarily informs what “imprisoned” means referring to the American system. Therefore, we have set aside non-U.S. lines in our corpus analysis.

In nearly all cases, “imprisoned” in the U.S. Code means punishment after conviction. In COCA “imprisoned” refers to “detainment post-conviction,” “detainment without conviction,” and detainment generally at similar (though not equal) rates.

A.II Double-Dissociation Analysis

Following the lead of Professors Solan and Gales,³⁴¹ we also pursued a double-dissociation analysis. That analysis asks how often someone intending to communicate a particular meaning (e.g., “detained post-conviction”) will choose to use a particular word (e.g., “imprisoned”). If we construct a large domain of instances in which someone uses a term to refer to “detained without conviction,” and “imprisoned” constitutes only a small percentage, this would indicate that people usually don’t use the term “imprisoned” when they refer to “detained without conviction.” Not only would language users not understand “imprisoned” to commonly mean “detained without conviction,” they would also assume that “imprisoned” did not communicate that meaning because someone communicating “detained without conviction” would use another term.

To operationalize this approach, two steps are necessary: (1) constructing a domain of near-synonyms³⁴² for “imprisoned;” and (2) observing these near-synonyms in context to see how often the meaning “detained without conviction” arises.

We assembled a list of near-synonyms from seven thesauruses³⁴³ and then reviewed four common synonyms of “imprisoned” for further analysis: incarcerated, confined, detained, and

³⁴¹ See Tammy Gales & Lawrence M. Solan, *Revisiting a Classic Problem in Statutory Interpretation: Is a Minister a Laborer?*, 36 GA. ST. U. L. REV. 481, 512 (2012) (“Following the idea of double dissociation utilized in our earlier work, we examined instances of ‘work’ as a variation of ‘labor’ to see whether statutory language indeed covered a broader concept of employment-related activity using a term other than ‘labor.’”).

³⁴² See Philip Edmonds & Graeme Hirst, *Near Synonymy and Lexical Choice*, 28 COMPUTATIONAL LINGUISTICS 105, 107 (2002) (“Usually, words that are close in meaning are near-synonyms . . . almost synonyms, but not quite; very similar, but not identical, in meaning; not fully intersubstitutable, but instead varying in their shades of denotation, connotation, implicature, emphasis, or register.”).

³⁴³ Several sources support using dictionaries and thesauruses to find near-synonyms. Professor Dilin Liu of the University of Alabama used dictionaries, thesauruses, and synonym dictionaries to unearth near-synonyms. See generally Dilin Liu, *Is It a Chief, Main, Major, Primary, or Principal Concern?: A Corpus-Based Behavioral Profile Study of the Near-Synonyms*, 15 INT’L J. CORPUS LINGUISTICS 56 (2010).

jailed. We could have included more,³⁴⁴ but we included only the more common synonyms to simplify our analysis.

To set up the double-dissociation analysis, we first noted how frequently each term occurred in the corpora:³⁴⁵

TABLE II: RELATIVE FREQUENCY OF TERMS, RAW NUMBERS

Term	COCA, 1990s	U.S. Code
imprisoned	1220	6381
incarcerated	502	637
confined	2376	1191
detained	800	1497
jailed	689	274

Then we converted these raw numbers into decimal frequencies. Because it's only necessary to understand the terms' frequencies relative to each other, we used estimates of the actual numbers of words in each database. For COCA, we assumed that the corpus had 330 million words—one-third of the "billion words" that COCA advertises (we narrowed our searched to one-third of the database—the 1990s—to review materials from closer to the passage of the relevant act). For the Corpus of the Current U.S. Code, we assumed that our corpus had 22 million words—a number identified by a scholarly mathematical paper.³⁴⁶ It doesn't matter if the corpora have more or less words than we expect; we could have assumed that the corpora had only ten thousand words to make our math a little more easily explained. In more formal terms, we eventually divide this unit of measurement by the summed frequencies based on the same unit of measurement—eliminating the unit of measurement and showing the relative frequency as opposed to an actual measure of frequency.

³⁴⁴ Indeed, we originally reviewed two other terms—interned and impounded—and discarded them from our analysis when we found essentially no results that referred to someone detained without conviction in a U.S. jurisdiction.

³⁴⁵ To get enough results for all the terms in the U.S. Code corpus, we included lemmas (i.e., other forms of the words).

³⁴⁶ Michael J. Bommarito & Daniel Katz, *A Mathematical Approach to the Study of the United States Code* 1 (Oct. 22, 2018) (unpublished manuscript) (<https://perma.cc/DB4C-DDPA>).

TABLE III: RELATIVE FREQUENCY OF TERMS, DECIMAL

Term	COCA, 1990s	U.S. Code
Imprisoned	0.00000369696	0.00029004545
Incarcerated	0.00000152121	0.00002895454
Confined	0.0000072	0.00005413636
Detained	0.00000242424	0.00006804545
Jailed	0.00000208787	0.00001245454

We then analyzed each of the near-synonyms for the previously identified meanings of “imprisoned.” For the sake of our analysis, we could have just looked for results referring to one secondary meaning of “imprisoned” to construct a domain of terms referring to that meaning. But we reviewed for all the possible meanings so we can double dissociate for each meaning. We coded two hundred lines of each term from each corpus:

TABLE IV: MEANINGS IN COCA

Meaning	Impris- oned	Incar- cerated	Con- fined	De- tained	Jailed
Detained after conviction	30 (15%)	98 (44%)	1 (0.5%)	1 (0.5%)	36 (18%)
Detained without conviction	55 (27.5%)	22 (11%)	32 (16%)	81 (40.5%)	29 (14.5%)
Detained generally	33 (16.5%)	58 (24%)	1 (0.5%)	12 (6%)	58 (29%)
Non-U.S. legal system	57 (28.5%)	21 (10.5%)	5 (2.5%)	90 (45%)	76 (38%)
Unclear	4 (2%)	0	2 (1%)	3 (1.5%)	0 (0%)
Other (including metaphorical)	21 (10.5%)	1 (0.5%)	159 (79.5%)	13 (6.5%)	1 (0.5%)
Total lines	200	200	200	200	200

TABLE V: MEANINGS IN THE U.S. CODE

Meaning	Impris- oned	Incar- cerated	Con- fined	De- tained	Jailed
Detained after conviction	195 (97.5%)	126 (63%)	52 (26%)	16 (8%)	4 (2%)
Detained without conviction	0	7 (3.5%)	23 (11.5%)	126 (63%)	0
Detained generally	1 (0.5%)	65 (32.5%)	28 (14%)	30 (15%)	0
Non-U.S. legal system	4 (2%)	2 (1%)	2 (1%)	27 (13.5%)	11 (5.5%)
Unclear	0	0	1 (0.5%)	0	0
Other (including metaphorical)	0	0	94 (47%)	1 (0.5%)	185 (92.5%)
Total lines	200	200	200	200	200

We then multiplied the frequency with which a term appears in the corpus (expressed as a decimal) by the percentage of the time that the term communicates a particular meaning. The product is that percent of the time that any randomly chosen term in the corpus is both (1) the relevant word itself (e.g., “confined”) and (2) communicating a particular meaning (e.g., “detained without conviction”). Next, we added the values in each row to estimate the percent of words in the corpus that communicate a particular meaning.

TABLE VI: DOMAIN PERCENTAGES IN COCA

Meaning	Impris- oned	Incar- cerated	Con- fined	De- tained	Jailed	Total
Detained after conviction	5.54544e-7 (33.6%)	6.69332 4e-7 (40.5%)	3.6e-8 (2.2%)	1.21212 e-8 (0.7%)	3.7581 66e-7 (22.8%))	1.65 e-6
Detained without conviction	0.00000101 666 (28%)	1.67333 1e-7 (4.6%)	0.00000 1152 (31.8%)	9.81817 2e-7 (27.1%)	3.0274 115e-7 (8.4%)	3.620 55e-6
Detained generally	6.099984 e-7 (35.7%)	3.65090 4e-7 (20.7%)	3.6e-8 (2%)	1.45454 4e-7 (8.7%)	6.0548 23e-7 (34.4%))	1.76 e-6
Non-U.S. legal system	0.00000105 363 (32.1%)	1.59727 05e-7 (4.9%)	1.8e-7 (5.5%)	0.00000 10909 (33.3%)	7.9339 06e-7 (24.2%))	3.277 65e-6
Unclear	7.39392e-8 (40.6%)	0	7.2e-8 (39.7%)	3.63636 e-8 (20%)	0	1.82 e-7
Other (including metaphori- cal)	3.881808 e-7 (6.2%)	7.60605 e-9 (0.1%)	0.00000 5724 (91%)	1.57575 6e-7 (2.5%)	1.0439 35e-8 (.2%)	6.29 e-6

TABLE VII: DOMAIN PERCENTAGES IN U.S. CODE

Meaning	Imprisoned	Incarcerated	Confined	Detained	Jailed	Total
Detained after conviction	0.000282 79431 (88.1%)	0.00001 824136 (5.7%)	0.00001 407545 (4.4%)	0.000005 44363 (1.7%)	2.49090 8e-7 (0.1%)	0.0003 20804
Detained without conviction	0	0.00000 10134 (2%)	0.00000 622568 (12.4%)	0.000042 86863 (85.6%)	0	5.0107 7e-5
Detained generally	0.000001 45022 (5.1%)	0.00000 941022 (32.8%)	0.00000 757909 (26.5%)	0.000010 20681 (35.6%)	0	2.8646 3e-5
Non-U.S. legal system	0.000005 8009 (35.2%)	2.89545 4e-7 (1.8%)	5.41363 6e-7 (3.3%)	0.000009 18613 (55.7%)	6.84999 7e-7 (4.2%)	1.6502 9e-5
Unclear	0	0	1.65029 e-5 (100%)	0	0	1.6502 9e-5
Other (including metaphorical)	0	0	0.00002 544408 (68.2%)	3.402272 5e-7 (0.9%)	0.00001 152044 (30.9%)	3.7304 7e-5

Ultimately, we can divide the value in each cell by the summed value in the “Total” column—that result is expressed as a percentage in each cell. By dividing (a) the frequency at which a term in the corpus is both a particular word and refers to a particular meaning, by (b) the frequency at which a term (any word) in the corpus refers to a particular meaning, we measure (c) what percent of words communicating a meaning are each particular word.

A lower percentage indicates that someone communicating a particular meaning will rarely choose to use that word. A higher percentage indicates that people will often choose that word.