


11-1-2010

The Dictionary Is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning

Stephen C. Mouritsen

Follow this and additional works at: <https://digitalcommons.law.byu.edu/lawreview>

 Part of the [Judges Commons](#), and the [Legal History Commons](#)

Recommended Citation

Stephen C. Mouritsen, *The Dictionary Is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning*, 2010 BYU L. Rev. 1915 (2010).

Available at: <https://digitalcommons.law.byu.edu/lawreview/vol2010/iss5/10>

This Note is brought to you for free and open access by the Brigham Young University Law Review at BYU Law Digital Commons. It has been accepted for inclusion in BYU Law Review by an authorized editor of BYU Law Digital Commons. For more information, please contact hunterlawlibrary@byu.edu.

The Dictionary Is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning

[I]t is one of the surest indexes of a mature and developed jurisprudence not to make a fortress out of the dictionary.

—Judge Learned Hand¹

The judicial conception of lexical meaning—i.e., what judges think about what words mean, or, more importantly, *how* judges arrive at the meaning of contested terms—is often outcome determinative. Vast fortunes or years of confinement may balance precariously on the interpretation of a single word. When faced with hard cases²—cases in which contextual cues or legislative definitions do not decisively favor either party’s asserted meaning—judges, like many speakers of English, will cast about for interpretive tools, often “looking for comfortable reassurance” in one of the language’s “more firmly established and dependably stable institutions”—the English dictionary.³ Such dictionaries, said Justice Jackson, are “the last resort of the baffled judge.”⁴

Baffled or not, judges cannot escape the reverence with which society regards its dictionaries—a reverence that often borders on the devotional.⁵ Indeed, the dictionary is sometimes spoken of as a

1. *Cabell v. Markham*, 148 F.2d 737, 739 (2d Cir. 1945).

2. Ronald Dworkin, *Hard Cases*, 88 HARV. L. REV. 1057, 1060 (1975) (defining “hard cases” as those in which “no settled rule dictates a decision either way”); RICHARD A. POSNER, *HOW JUDGES THINK* 211–12 (2008) (“[D]ifficult cases . . . are so mainly because they are cases in which the orthodox materials of legal decision making cannot produce a satisfactory decision.”).

3. RANDOLPH QUIRK, *STYLE AND COMMUNICATION IN THE ENGLISH LANGUAGE* 86 (1982).

4. *Jordon v. DeGeorge*, 341 U.S. 223, 234 (1951) (Jackson, J., dissenting).

5. JONATHON GREEN, *CHASING THE SUN: DICTIONARY MAKERS AND THE DICTIONARIES THEY MADE* xiii (1997) (“The lexicographer, the interpreter and arbiter of the very language that underpins every aspect of communication, is far more deity than drudge. Or if not a deity, then certainly a priest, charged by society—whether consciously or not—with the revelation of the linguistic verities.”); QUIRK, *supra* note 3, at 86–87 (“The time-hallowed format [of the dictionary] helps to place it mentally with the Bible (alongside which it is likely to find itself physically), and the advertiser’s warning that ‘no home should be without it’ finds a ready response in the natural awe that we rightly have for our language faculty and further contributes to the implicit belief that the dictionary is one’s linguistic bible.”).

“linguistic bible.”⁶ Though such reverence for dictionaries is “deeply embedded in our culture,”⁷ dictionaries are often inadequate objects of our devotion,⁸ and their compilation is a decidedly human endeavor.⁹ “[W]e commonly ignore the fact,” says Professor Lawrence Solan, “that someone sat there and wrote the dictionary, and we speak as though there were only one dictionary, whose lexicographer got all the definitions ‘right’ in some sense that defies analysis.”¹⁰

Perhaps no case in recent memory has better illustrated the Supreme Court’s difficulty in resolving lexical ambiguity and its problematic reliance upon dictionaries as aids to interpretation than that of *Muscarello v. United States*.¹¹ At issue in *Muscarello* was the interpretation of the phrase *carries a firearm* and whether Congress intended by that term to include the notion of conveyance in a vehicle.¹² This language became law as part of the Omnibus Crime Control and Safe Streets Act of 1968 (the “Act”), and was later codified as 18 U.S.C. § 924(c)(1). In reaching the conclusion that Congress did, in fact, intend the conveyance meaning, the Court relied upon an extraordinary panoply of sources, including: the King James Bible, *Robinson Crusoe*, and *Moby Dick*, as well as two electronic newspaper databases and several unabridged dictionaries.¹³

Not to be outdone, the dissent likewise relied upon a legal

6. QUIRK, *supra* note 3, at 86–87.

7. Lawrence Solan, *When Judges Use Dictionaries*, 68 AM. SPEECH 50, 50 (1993) (“[O]ur society’s reverence for dictionaries is not driven by the latest discoveries in psycholinguistic research. Rather, it is deeply embedded in our culture.”).

8. See A. Raymond Randolph, *Dictionaries, Plain Meaning, and Context in Statutory Interpretation*, 17 HARV. J.L. & PUB. POL’Y 71, 72 (1994) (“[C]iting to dictionaries creates a sort of optical illusion, conveying the existence of certainty—or ‘plainness’—when appearance may be all there is.”).

9. GREEN, *supra* note 5, at xiv (“[D]ictionaries do not emerge from some lexicographical Sinai; they are the products of human beings. And human beings, try as they may, bring their prejudices and biases into the dictionaries they make.”).

10. See Solan, *supra* note 7, at 50. Rather than a cultural instinct, a resort to dictionaries in order to ascertain the plain meaning of a text is sometimes mandated by precedent. See, e.g., *Lorillard Tobacco v. Am. Legacy*, 903 A. 2d 728, 738 (Del. 2006) (“We agree that the Vice Chancellor’s abandonment of all dictionaries and his innovative review of how legal writers have used ordinary words in their texts to ascertain the plain meaning of the words are not supported by precedent. Under well-settled case law, Delaware courts look to dictionaries for assistance in determining the plain meaning of terms which are not defined in a contract.”).

11. 524 U.S. 125 (1998).

12. *Id.* at 126–27.

13. *Id.* at 127–30.

dictionary, several alternative translations of the Bible, and works of poetry,¹⁴ as well as a statement from Charles Bronson's character in *The Magnificent Seven*,¹⁵ a quotation from the television show *M*A*S*H*,¹⁶ and one of the Supreme Court's very few references to *Sesame Street*.¹⁷

"As the century nears its end," said one state court judge, "it is apparent that the competing opinions in [*Muscarello*] represent one of the great textualist moments of our period."¹⁸ The case has been variously characterized as "an energetic textual debate"¹⁹ and as more of "a food fight . . . than a serious argument among distinguished jurists."²⁰ The case has likewise been said to show "a need for improvement in judicial reasoning about statutory concepts,"²¹ or, perhaps less generously, "it illustrates how bankrupt courts are when they must actually decide just what makes ordinary meaning ordinary."²²

In the midst of all of the *Muscarello* Court's lexical wrangling, it is easy to lose sight of what is at stake.²³ Frank Muscarello set out to

14. *Id.* at 142–44 (Ginsburg, J., dissenting).

15. *Id.* at 144 n.6 ("You think I am brave because I carry a gun. Well, your fathers are much braver because they carry responsibility, for you, your brothers, your sisters, and your mothers.") (quoting *THE MAGNIFICENT SEVEN* (The Mirisch Company, Inc. and Alpha Productions 1960)).

16. *Id.* ("I will not carry a gun. . . . I'll carry your books, I'll carry a torch, I'll carry a tune, I'll carry on, carry over, carry forward, Cary Grant, cash and carry, carry me back to Old Virginia, I'll even 'hari-kari' if you show me how, but I will not carry a gun!") (quoting *M*A*S*H: Officer of the Day* (CBS television broadcast Sept. 24, 1974)).

17. *Id.* at 147 n.11 ("But as viewers of 'Sesame Street' will quickly recognize, 'one of these things [a statute authorizing conduct] is not like the other [a statute criminalizing conduct].'"). One commentator has noted that these responses by the dissent were intended to be subtly humorous. Laura Krugman Ray, *Justice Ginsburg and the Middle Way*, 68 *BROOK. L. REV.* 629, 671–72 (2003) ("The dissent suggested her sensitivity to the nuances of words and the care with which she chooses them when she is putting herself in opposition to her colleagues, but it also suggested that she is more inclined to moderate her criticism with playfulness than to intensify it with an aggressively hostile tone. It is characteristic of Ginsburg that even her strongest dissents avoid the soaring rhetoric other members of the Court often use.").

18. *Henderson v. State*, 715 N.E.2d 833, 835 n.3 (Ind. 1999).

19. *Id.* at 835.

20. Lawrence M. Solan, *The New Textualist's New Text*, 38 *LOY. L.A. L. REV.* 2027, 2053 (2005).

21. Lawrence M. Solan, *Why Laws Work Pretty Well, but Not Great: Words and Rules in Legal Interpretation*, 26 *LAW & SOC. INQUIRY* 243, 258 (2001).

22. See Solan, *supra* note 20, at 2053.

23. See *Muscarello v. United States*, 524 U.S. 125, 140 (1998) (Ginsburg, J., dissenting).

sell marijuana with a handgun locked in the glove compartment of his truck, and that handgun remained in his glove box until the time of his arrest.²⁴ By interpreting § 924(c)(1) broadly, the majority has extended Mr. Muscarello's likely ten-to-sixteen month sentence—a sentence that included a forty-percent increase for possessing, but not carrying, a firearm—to a mandatory five years, plus the additional six-to-twelve months for his underlying drug offense.²⁵ In our common law system, this means that all criminal defendants similarly situated²⁶ will be brought within the ambit of this statute's heightened penalty.

In their comprehensive article on Supreme Court dictionary use, entitled *The Lexicon Has Become a Fortress*, Judge Samuel Thumma and Jeffrey Kirchmeier cite *Muscarello* as evidence that the Court's dictionary usage "at times appears to lack principled guidance."²⁷ Their discussion of *Muscarello* is framed in the broader context of an overarching trend to rely upon dictionaries to resolve lexical ambiguity²⁸—a trend with important implications for the rule of law.²⁹ Professor Lawrence Solan, referenced above, has characterized *Muscarello* as "an advertisement for the need of more serious judicial approaches to the meanings of statutory words."³⁰ He contends that the discussion of ordinary meaning in *Muscarello* is actually a search for two types of meaning: (1) a rule-based definition of *carry* derived

24. *Id.* at 127 (majority opinion).

25. *Id.* at 141 (Ginsburg, J., dissenting).

26. The precise number of criminal defendants subjected to heightened penalty under the *Muscarello* holding is not entirely clear. One study put the number at "fewer than 200 a year." See Paul J. Hofer, *Federal Sentencing for Violent and Drug Trafficking Crimes Involving Firearms: Recent Changes and Prospects for Improvement*, 37 AM. CRIM. L. REV. 41, 62 (2000). Still, if the *Muscarello* decision is based on fallacious reasoning about the meaning of words and the content of dictionaries, as I argue that it is, it would be small comfort to a federal prisoner spending an extra five years in prison to know that only a few hundred people each year are similarly situated.

27. Samuel A. Thumma & Jeffrey L. Kirchmeier, *The Lexicon Has Become a Fortress: The United States Supreme Court's Use of Dictionaries*, 47 BUFF. L. REV. 227, 270–71 (1999).

28. *Id.* at 248–60 (citing statistical data demonstrating that "the decade of the 1990s [gave] rise to nearly half of all the opinions in the court's two-century history where a Justice has relied on a dictionary"); Note, *Looking It Up: Dictionaries and Statutory Interpretation*, 107 HARV. L. REV. 1437, 1454 (1994) [hereinafter *Looking It Up*] (showing a nearly exponential increase in the Court's reliance upon dictionaries).

29. See Solan, *supra* note 20, at 2053.

30. Solan, *supra* note 21, at 259.

from dictionaries; and (2) a prototypical meaning of *carry* as evidenced by the usage of the term.³¹

This Comment will examine a question not addressed in prior discussions of *Muscarello*, which is that the majority's discussion regarding the ordinary meaning of *carry* rests upon demonstrably incorrect assumptions about the content and structure of dictionaries. I chose the *Muscarello* decision for analysis, not necessarily because I believe it to be illustrative of the interpretive approach favored by the decision's author, Justice Stephen Breyer. Justice Breyer's reticence to resort to what might be viewed as textual pedantry is obvious from the opinion itself.³² Instead, *Muscarello* reveals common but fallacious assumptions about the structure and content of dictionaries—assumptions that are widely shared in the legal community. Having dismissed these assumptions as erroneous, I will then suggest an alternative method for examining questions of ordinary meaning—a method based in a computational approach to linguistic analysis known as corpus linguistics. This approach offers intriguing possibilities for ambiguity resolution, while avoiding some of the pitfalls of the *Muscarello* Court's dictionary-based jurisprudence.

Part I examines the traditional use of dictionaries in statutory interpretation. Part II addresses the majority's argument in *Muscarello*, identifying the definitional fallacies that undermine the Court's reasoning in the case. Part III examines the majority's reliance upon electronic databases in determining the ordinary meaning of *carry*. Part IV will address briefly the majority's purposivist argument and Justice Ginsburg's dissent. Part V proposes a corpus-based approach to resolving questions of lexical ambiguity. Part VI concludes this Comment.

I. DICTIONARIES AND THE COURT

Before proceeding to the arguments at issue in *Muscarello*, it may be helpful to contextualize the dramatic rise in dictionary use by the Court and to examine the historic role of dictionaries in the Court's jurisprudence. As noted in the epigraph above, Judge Learned Hand stated that "it is one of the surest indexes of a mature and developed

31. *Id.*

32. *See infra* notes 148–59 and accompanying text.

jurisprudence not to make a fortress out of the dictionary.”³³ Perhaps a similar sentiment cautioned judicial restraint in appealing to dictionaries in the earliest decades of the Republic. Prior to 1864, the Supreme Court cited dictionaries in only three cases,³⁴ and this paucity of dictionary citation continued until the 1970s when dictionary citation began to grow exponentially.³⁵

The exponential growth in the number of times the Court cites dictionaries corresponds to a similarly dramatic rise in the statistical frequency of the words *ambiguity* and *ambiguous* in Supreme Court decisions.³⁶ From the Court’s inception, through 1969, the term *ambiguous* had a statistical frequency between eleven-to-thirteen words per million.³⁷ From 1970 to 2005, that number increased to over one-hundred words per million.³⁸ During that same timeframe, the Court has increased its invocation of the so-called “Plain Meaning Rule” by an order of magnitude.³⁹ Again from the Court’s inception, through 1969, the phrase *plain meaning* never had a statistical frequency of more than 4.2 words per million, while the phrase *ordinary meaning* never rose above 3.7 words per million.⁴⁰ From 1991 to 2005, the statistical frequency of both phrases nearly quadrupled from their previous highs—*plain meaning* rose to 15.43 words per million and *ordinary meaning* rose to 15.77.⁴¹

Correlation is not causation, but the dramatic and coincident rise in the Court’s reliance upon dictionaries, the Court’s discussion of ambiguity, and the Court’s invocation of the Plain Meaning Rule is striking. Justice Antonin Scalia has suggested that “[w]e live in an age of legislation,”⁴² and that “[b]y far the greatest part of what . . .

33. *Cabell v. Markham*, 148 F.2d 737, 739 (2d Cir. 1945).

34. *See* Thumma & Kirchmeier, *supra* note 27, at 248.

35. *See id.* at 252–60; *Looking It Up*, *supra* note 28, at 1454.

36. *See* Appendix I.

37. *Id.*

38. *Id.*

39. *See* Appendix II.

40. *Id.*

41. *Id.*

42. ANTONIN SCALIA, A MATTER OF INTERPRETATION: FEDERAL COURTS AND THE LAW 13 (1998); *see also*, Felix Frankfurter, *Some Reflections on the Reading of Statutes*, 47 COLUM. L. REV. 527, 527 (1947) (“Broadly speaking, the number of cases disposed of by opinions has not changed from term to term. But even as late as 1875 more than 40% of the controversies before the Court were common-law litigation, fifty years later only 5%, while today cases not resting on statutes are reduced almost to zero. It is therefore accurate to say that courts have ceased to be the primary makers of law in the sense in which they ‘legislated’

all federal judges do is interpret the meaning of federal statutes and federal agency regulations.”⁴³ If Justice Scalia’s assessment is correct, and if we may properly draw the inference from the data above that courts are increasingly turning to dictionaries to resolve questions of ambiguity and ordinary meaning, then a principled approach to use of dictionaries in general and the resolution of lexical ambiguity in particular is needed now more than ever.

A. The Defining and Instantiating Role of Dictionaries

Prior to this comparatively recent escalation in dictionary use by the Court, dictionaries were employed for their definitional function—they were used to aid judges in understanding the meaning of a forgotten or unknown term. This practice of appealing to dictionaries simply as memory aids was deemed a function of judicial notice.⁴⁴

[Words] must receive their ordinary meaning. Of that meaning the court is bound to take judicial notice, as it does in regard to all words in our own tongue; and upon such a question dictionaries are admitted, not as evidence, but only as aids to the memory and understanding of the court.⁴⁵

This definitional use of dictionaries by judges is at once the most obvious and the least controversial;⁴⁶ surprisingly, it is likewise the least common.⁴⁷

the common law. It is certainly true of the Supreme Court, that almost every case has a statute at its heart or close to it.”).

43. SCALIA, *supra* note 42, at 13–14.

44. Thumma & Kirchmeier, *supra* note 27, at 246–48; *see also* *Werk v. Parker*, 249 U.S. 130, 132–33 (1918) (“[W]e deem it clear, beyond question—that the court [is] justified in taking judicial notice of facts that appeared so abundantly from standard works accessible in every considerable library.”); *Brown v. Piper*, 91 U.S. 37, 42 (1875) (“Of private and special facts, in trials in equity and at law, the court or jury, as the case may be, is bound carefully to exclude the influence of all previous knowledge. But there are many things of which judicial cognizance may be taken. . . . Facts of universal notoriety need not be proved. . . . Among the things of which judicial notice is taken are the laws of nations . . . [and] *the meaning of words in the vernacular language . . .*”) (emphasis added).

45. Thumma & Kirchmeier, *supra* note 27, at 246–48; *Nix v. Hedden*, 149 U.S. 304, 306–07 (1893) (holding at the nadir of the Court’s botanical competence that the tomato is a vegetable).

46. *See* Solan, *supra* note 7, at 55 (“In other cases, the dictionary is used to give the reader a general sense of the word, which seems to me an appropriate use of the dictionary, whether or not it is necessary.”); Craig Hoffman, *Parse the Sentence First: Curbing the Urge to Resort to the Dictionary when Interpreting Legal Texts*, 6 N.Y.U. J. LEGIS. & PUB. POL’Y 401,

Dictionaries may also serve an instantiating function, that is, they may be used by the court to confirm that a contested meaning has been employed in either speech or literature, and has thus been recognized as a valid meaning by lexicographers. Of this instantiating function, Professors Hart and Sacks said:

Unabridged dictionaries are historical records (as reliable as the judgment and industry of the editors) of the meanings with which words have in fact been used by writers of good repute. They are often useful in answering hard questions of whether, in an appropriate context, a particular meaning is *linguistically permissible*.⁴⁸

In using a dictionary to instantiate a contested meaning, a judge searches the dictionary to determine what meanings have attained currency in the language at large and are thus linguistically permissible in a given context.⁴⁹ This instantiating function may serve a modest channeling role in disputes involving lexical ambiguity. If a party suggests that an opponent's definition for a given term is entirely made up, or otherwise invalid, a judge may turn to the dictionary to confirm that the meaning is instantiated there. In this

416 (2002–2003) (“Definition is also used when the Court simply does not know (or believes that the reader may not know) the accurate definition of a word that it is using Needless to say, [this approach] is completely appropriate when ‘definition’ is the Court’s sole objective.”).

47. Hoffman, *supra* note 46, at 416.

48. HENRY M. HART, JR. & ALBERT M. SACKS, *THE LEGAL PROCESS: BASIC PROBLEMS IN THE MAKING AND APPLICATION OF LAW* 1375–76 (William N. Eskridge, Jr. & Phillip P. Frickey eds., 1994).

49. See SCALIA, *supra* note 42, at viii (“Words do have a limited range of meaning, and no interpretation that goes beyond that range is permissible.”). Not every meaning of a term will be represented in a dictionary, and thus, if a contested meaning of a given term is not found there, this is not the end the inquiry. For example, in *United States v. Willfong*, 274 F.3d 1297 (9th Cir. 2001), a man was convicted of *interfering with a forest service officer* after merely failing to comply with an officer’s order. *Id.* at 1298–99. Judge Noonan, writing in dissent, noted that of the seven senses of *interfere* in *Webster’s Third New International Dictionary*, none makes any reference to *failure to comply* and concludes that the court’s definition is anomalous. *Id.* at 1304–05 (Noonan, J., dissenting). This conclusion simply ignores the possibility that the editors of the dictionary failed to include an entirely valid, but rare or highly technical meaning. “An unabridged dictionary is simply an historical record, not necessarily all-inclusive, of the meanings which words in fact *have* borne, in the judgment of the editors, in the writings of reputable authors.” HART & SACKS, *supra* note 48, at 1190. More compelling is Judge Noonan’s argument that the portion of the United States Code at issue in the case has five separate provisions for failing to comply with an officer’s order. *Willfong*, 274 F.3d at 1305–06 (Noonan, J., dissenting).

respect, the dictionary functions as an “impersonal authority.”⁵⁰ If the contested meaning appears in the dictionary, then it may be presumed that the meaning is, at the very least, a possible interpretation of a given word. At this point, the utility of the dictionary is at an end; parties with equally plausible meanings must look elsewhere to determine which contested meaning should control.

This modest role for dictionaries—defining unknown terms and instantiating contested meanings—is consistent with both the design and purpose of dictionaries. No longer the arbiters of what is good or correct in language, most contemporary dictionaries have abandoned any concern about prescribing standards of usage and have instead taken on a descriptive approach.⁵¹ This approach has been characterized as follows:

A language has at a given time a finite inventory of words, the meanings of which are revealed in the course of general usage. Since the lexicographer is as liable to be as deviant as the next man, he must have recourse to the usage of as many people as he can—in print. This last phrase embodies the proud ideal of descriptive objectivity; his citations (and interpretations of them) are publicly verifiable.⁵²

The implicit claim made when a given definition is included in a dictionary is not that a particular meaning is correct or even common, but that its use, in a given context, is verifiable.⁵³

A dictionary, it is vital to observe, never says what meaning a word *must* bear in a particular context. *Nor does it ever purport to say this.*

50. See HART & SACKS, *supra* note 48, at 1190.

51. LEXICOGRAPHY: PRINCIPLES AND PRACTICE 5 (R. R. K. Hartmann ed., 1985) (“[Lexicography] is above all a descriptive activity, recording existing usage rather than laying down prescriptive or normative rules about how words should be used or which words are to be avoided.”); RICHARD CHENEVIX TRENCH, ON SOME DEFICIENCIES IN OUR ENGLISH DICTIONARIES 4–5 (2d ed. 1857) (“A dictionary . . . is an inventory of the language . . . It is not the task of the maker to select the *good* words of the language. . . . The business which he has undertaken is to collect and arrange all the words, whether good or bad . . . which . . . those writing in the language have employed. He is an historian of it, not a critic.”); see Thumma & Kirchmeier, *supra* note 20, at 243 (“Today, most contemporary dictionaries are characterized as descriptive rather than prescriptive.”); see also HOWARD JACKSON, LEXICOGRAPHY: AN INTRODUCTION 66 (2002); SIDNEY I. LANDAU, DICTIONARIES: THE ART AND CRAFT OF LEXICOGRAPHY 254–261 (2d ed. 2001).

52. See QUIRK, *supra* note 3, at 87.

53. LEXICOGRAPHY: PRINCIPLES AND PRACTICE, *supra* note 51, at 19–20.

An unabridged dictionary is simply an historical record, not necessarily all-inclusive, of the meanings which words in fact have borne The editors make up this record by collecting examples of uses of the word to be defined, studying each use in context, and then forming a judgment about the meaning in that context.⁵⁴

The defining and instantiating role of dictionaries is also consistent with the acontextual nature of the lexical information presented in each dictionary entry. In defining a given term, a dictionary merely presents a range of possible meanings and a record of several ways in which the term has been employed in the past. One of the roles of the lexicographer is to collect records of these uses in a citation file,⁵⁵ and to select from among them the most illustrative examples. A dictionary cannot tell us precisely what meaning a word must bear in a particular context, because the lexicographer cannot know *a priori* every context in which the term will be found.

In contrast to this modest role for dictionaries, judges have increasingly sought to employ dictionaries for persuasive ends. They have done so by arguing that because multiple dictionaries define a term in a given way, a particular definition ought somehow to be controlling in a given case. Judges have also maintained that because a definition has been placed in a particular position in what the judge perceives as the dictionary's structural hierarchy, or because the derivation of a term reveals that its original use was similar to the meaning the judge favors, the judge's particular meaning should be

54. HART & SACKS, *supra* note 48, at 1190 (emphasis added); *see also* ARTHUR L. CORBIN, CORBIN ON CONTRACTS § 539 at 511, n.59 (2d ed. 1982) ("The better and more complete the Dictionary the more numerous and varied are the usages that it records and the less dogmatic are its assertions as to their relative merits."); Thumma & Kirchmeier, *supra* note 27, at 293 ("A descriptive dictionary sets forth definitions showing what a word may mean generally, not what a word does mean in context.").

55. *See* LANDAU, *supra* note 51, 190 ("A citation file is a selection of potential lexical units in the context of actual usage, drawn from a variety of written sources and often some spoken sources, chiefly because the context illuminates an aspect of meaning. Citations are clipped from the source and pasted on a card or a slip of paper, or retyped, and nowadays are routinely converted into a computer file by optical scanning or rekeyboarding. Citations are also collected to provide illustrative quotations that will be printed in the dictionary."). Perhaps the most famous of these collections is the Scriptorium at Oxford University, which housed the slips of sample usage sent in from the *Oxford English Dictionary's* more than 800 contributors. SIMON WINCHESTER, THE MEANING OF EVERYTHING: THE STORY OF THE OXFORD ENGLISH DICTIONARY 109–16 (2003). These slips would arrive at the Scriptorium at the rate of nearly 1,000 per day. *Id.* at 112; *see also* LANDAU, *supra* note 51, at 189–206 (describing the role of the citation file in the creation of modern dictionaries).

preferred. These conclusions are erroneous, not simply because they are at variance with the descriptive purpose for which most contemporary dictionaries are created, but because they rely upon deeply flawed assumptions about the structure and content of the information presented in dictionaries. Mounting judicial reliance upon such arguments undoubtedly led Thumma and Kirchmeier to conclude that—contrary to Judge Hand’s imperative—jurists are increasingly making a fortress out of the dictionary.⁵⁶

But the dictionary is not a fortress—its design and structure are no defense for jurists’ fallacious assumptions about language and meaning. Several of these assumptions are addressed below.

II. DEFINITIONAL FALLACIES

Legal proscription of the combination of illicit drug trafficking and firearms is hardly controversial. Few would march on the nation’s capital demanding the right of criminals to keep and bear arms while buying and selling drugs. Consequently, it comes as no surprise that the federal criminal code imposes heightened sanctions upon those who bring their guns to drug deals. The policy implications are fairly straightforward: “[W]hat we are trying to do . . . is to persuade the criminal to leave his gun at home.”⁵⁷

With this fairly uncontroversial goal in mind, the federal criminal code imposes “not less than five years” upon “any person who, during and in relation to any crime of . . . drug trafficking . . . uses or carries a firearm.”⁵⁸ This penalty accrues “in addition to the punishment” for the original drug-trafficking offence.⁵⁹

At issue in *Muscarello* is the question of whether this heightened penalty for carrying a firearm applies to those who convey their firearms in a vehicle as opposed to those who carry their firearms on their person.⁶⁰ Writing for the majority, Justice Breyer frames the dispute by stating that “[a]lthough the word ‘carry’ has many different meanings, only two are relevant here.”⁶¹ He then articulates

56. See generally Thumma & Kirchmeier, *supra* note 27.

57. 114 CONG. REC. 22,244 (1968) (statement of Rep. Randall).

58. 18 U.S.C. § 924(c)(1)(A) (2006).

59. *Id.* Significantly, this penalty may be imposed even if a jury fails to convict the defendant for the predicate offence, that is, even if the defendant is not convicted for buying and selling drugs. *United States v. Carter*, 300 F.3d 415, 425 (4th Cir. 2002).

60. *Muscarello v. United States*, 524 U.S. 125, 126 (1997).

61. *Id.* at 128.

his view that *carry*'s "first, or primary, meaning" reflects the notion of carrying as conveyance (hereafter *carry*₁); while only a "different, rather special" meaning of *carry* reflects the notion of carrying upon one's person (hereafter *carry*₂).⁶²

In order to establish that their characterization of the meaning of *carry* is the correct one, the *Muscarello* Court makes numerous appeals to dictionaries. These dictionaries are cited both for the way in which they define the term *carry* and also for the place accorded a given definition of *carry* in the overall structure of the dictionary.⁶³ Dictionaries are likewise employed to demonstrate that the *carry*₁ interpretation of *carry* is consistent, not just with the early usage of the term, but also with the term's original usage.⁶⁴ The Court insists that the sum of all "[t]he *relevant linguistic facts* are that the word 'carry' in its ordinary sense includes carrying in a car."⁶⁵

These arguments by the *Muscarello* Court reveal tacit assumptions about the structure and content of dictionaries that are at once deeply erroneous and easily dismissed when explicitly stated. Because the majority's argument relies upon these unstated premises that, while superficially persuasive⁶⁶ and, perhaps, commonly believed,⁶⁷ are nevertheless incorrect, I have characterized these assumptions as fallacies.

A. The Sense-Ranking Fallacy

The Court begins the analysis in a way that appears to embrace the modest dictionary use discussed above. The Court notes that the parties have stipulated "that Congress intended the phrase to convey its ordinary, and not some special legal, meaning."⁶⁸ The opinion then states that the *carry*₁ meaning is one that "*can*" be employed

62. *Id.*

63. *Id.* at 127–28, 130–31.

64. *Id.* at 128.

65. *Id.* at 131 (emphasis added).

66. IRVING M. COPI & CARL COHEN, INTRODUCTION TO LOGIC 138 (11th ed. 2001) ("It is customary to reserve the term 'fallacy' for arguments that, although incorrect, are psychologically persuasive.").

67. *Id.* at 137–38 ("The word 'fallacy,' however, as logicians use it, designates not any mistaken inference or false belief, but *typical* errors, that is, mistakes that arise commonly in ordinary discourse and that devastate the arguments in which they appear. . . . [F]allacies are dangerous because most of us are fooled by some of them on occasion.") (emphasis in original).

68. *Muscarello*, 524 U.S. at 128.

“as a matter of ordinary English,”⁶⁹ not necessarily one that *must* be employed. This use of *can* suggests that the majority is attempting to demonstrate that *carry*₁ is simply “linguistically permissible.”⁷⁰

Justice Breyer soon makes clear that such definitional liberality is unintended. He characterizes *carry*₁ as the “first,” “primary,” or “ordinary” meaning of *carry*, while referring to *carry*₂ as a “different, special way” of using *carry*.⁷¹ Then, in his first sweeping appeal to several dictionaries, Justice Breyer says the following:

Consider first the word’s primary meaning. The *Oxford English Dictionary* gives as its *first* definition “convey, originally by cart or wagon, hence in any vehicle, by ship, on horseback, etc.” 2 OXFORD ENGLISH DICTIONARY 919 (2d ed. 1989); see also WEBSTER’S THIRD NEW INTERNATIONAL DICTIONARY 343 (1986) (*first* definition: “move while supporting (as in a vehicle or in one’s hands or arms)”); RANDOM HOUSE DICTIONARY OF THE ENGLISH LANGUAGE UNABRIDGED 319 (2d ed. 1987) (*first* definition: “to take or support from one place to another; convey; transport”).⁷²

In the quotation above, Justice Breyer cites three of the most revered English dictionaries, noting that in each *carry*₁ is the “*first* definition.” For each citation, he prints “*first*” in italics. With all of these *firsts*, it is perhaps revealing that Justice Breyer chooses the word *primary* to sum up what these quotations demonstrate. “Consider first the word’s primary meaning,” he says.⁷³

Why say *primary* when he has elsewhere said *first*? The answer may be that Justice Breyer is equivocating,⁷⁴ that is, he is implicitly using *primary* to mean both first in sequence and first in importance. In the passage above, he calls attention to the fact that his suggested meaning of *carry* is listed first among several meanings in three dictionaries, but he elsewhere contrasts *carry*’s “primary meaning,”

69. *Id.* (emphasis added).

70. See HART & SACKS, *supra* note 48.

71. *Muscarello*, 524 U.S. at 128.

72. *Id.*

73. *Id.*

74. See COPI & COHEN, *supra* note 66, at 163 (“[W]hen we confuse the several meanings of a word or phrase—accidentally or deliberately—we are using the word equivocally. If we do that in the context of an argument, we commit the fallacy of equivocation.”).

with what he believes to be the term's "different" or "special meaning," i.e., *carry*.⁷⁵

Consequently, Justice Breyer uses the term *primary* in two senses: First, to suggest that a particular meaning of *carry* "[o]ccur[s] or exist[s] first in a sequence,"⁷⁶ and, second, to suggest that his particular sense of *carry* is of "the highest rank or importance."⁷⁷ The implication of the passage above is clear: Because *carry*₁ occurs first in the sequence of senses it is the meaning "[o]f the highest rank or importance."

This notion that a given sense of a term may be considered somehow primary or ordinary or more likely to be legally operative than another sense simply because it is listed first in a dictionary is what I have called the Sense-Ranking Fallacy. The *Muscarello* Court is by no means alone in its reliance upon this fallacy. Advocates have credited the Sense-Ranking Fallacy,⁷⁸ as have appellate judges.⁷⁹ In

75. *Muscarello*, 524 U.S. at 128.

76. 7 OXFORD ENGLISH DICTIONARY 472 (2d ed. 1989).

77. *Id.*

78. In a recent Supreme Court case, the respondents faulted the Solicitor General for employing not "primary, secondary, or even tertiary definitions of the relevant terms but . . . octonary ones." Brief of Respondents Riverkeeper, Inc. at 23, *Entergy Corp. v. Riverkeeper, Inc.*, 129 S. Ct. 1498 (2009) (Nos. 07-588, 07-589, 07-597); *see also* *Hasbro, Inc. v. MGA Entm't, Inc.*, C.A. No. 06-262 S, n.7 (Dist. R.I. 2007) ("Hasbro argues that these definitions should be considered 'uncommon' because they occur toward the end of the [dictionary] entry's definitional list . . . Hasbro's . . . claim is unpersuasive . . . No case suggests that the placement of a definition in the entry list is dispositive, or even particularly relevant, to whether the term is generic or not."); *Fireman's Fund Ins. Cos. v. Ex-Cell-O Corp.*, 702 F. Supp. 1317, 1324 (E.D. Mich. 1988) ("Policyholders argue that the 'first, and thus preferred' definition of 'sudden' is 'happening or coming unexpectedly.' The separation of the senses of a word by numbers and letters, however, does not 'evaluate or establish an enduring hierarchy of importance among them. The best sense is the one that most aptly fits the context of an actual genuine utterance.'") (citations omitted).

79. *Anderson v. Malloy*, 700 F.2d 1208, 1215 n.1 (8th Cir. 1983) (Gibson, J., dissenting) ("Not only the dictionaries cited by the majority but also The American Heritage Dictionary of the English Language, Funk & Wagnall's, Standard Encyclopedic Dictionary, Webster's New World Dictionary, and The Oxford English Dictionary *all give as the first definition that recognized by the Supreme Court* . . . and as the second definition that recognized by the majority in its opinion.") (emphasis added); *State v. Paul*, 860 P.2d 992, 995 (Utah Ct. App. 1993) (Garff, J., dissenting) ("The majority relies on the 'first definitions' of several dictionaries to determine the definition that is 'the most commonly accepted.' However, a coordinate sub-sense of the primary definition is the following: 'to propel through the air in any manner.' Webster's Third New International Dictionary (unabridged) 2385 (1986); *see also* Webster's Explanatory Note 12.2 and 12.4 at 17a (explaining that the various sub-senses do not represent 'an enduring hierarchy,' and that the 'best sense is the one that most aptly fits the context.')).

fact, the Sense-Ranking Fallacy led one exasperated Fifth Circuit judge to say that

In support of its intimation that “same” must mean “identical,” the majority draws a distinction between “primary” and “secondary” dictionary definitions. Where does this distinction come from and what does it mean? The majority does not base this nonsensical distinction on any authority. I cannot imagine that the majority favors interpreting statutes by choosing the first definition that appears in a dictionary.⁸⁰

What is the origin of this fallacy? The Sense-Ranking Fallacy is perhaps consistent with the basic human presumption that the most important things should go first in a sequence: National champions are listed at the top of the rankings, CEO’s and managers are listed at the top of the organizational chart, and gold medalists stand at the top of the podium. More to the point, the Sense-Ranking Fallacy is consistent with the simple fact that virtually no one reads the front-matter of a dictionary,⁸¹ and consequently, most are left unaware of the basic content, structure, purpose, and typology of the dictionaries they use.⁸²

A careful examination of each of the dictionaries cited by the majority reveals that the Court’s reliance upon these sense rankings is fallacious.

1. The Webster’s Third New International Dictionary

We begin with the *Webster’s Third New International Dictionary* (“Webster’s Third”), the most cited dictionary in Supreme Court jurisprudence.⁸³ That Webster’s Third has come to occupy this position on the Court is something of a surprise, considering its rocky beginning. The dictionary was revolutionary when it was published in 1961, being among the first to adopt a purely descriptivist stance, including among its reference material colloquialisms and examples from common speech.⁸⁴ The dictionary

80. *Miss. Poultry Ass’n v. Madigan*, 992 F.2d 1359, 1368–69 (5th Cir. 1993) (Reavley, J., dissenting) (citation omitted).

81. See JACKSON, *supra* note 51, at 76–77.

82. P. G. J. VAN STERKENBURG, *A PRACTICAL GUIDE TO LEXICOGRAPHY* 233 (2003) (“Dictionary editors . . . claim that it is only the reviewers who read the front matter.”); R.R.K. HARTMANN, *TEACHING AND RESEARCHING LEXICOGRAPHY* 83–86 (2001).

83. See Thumma & Kirchmeier, *supra* note 27, at 262–63.

84. See LANDAU, *supra* note 51, at 207.

was “roundly denounced for doing so,” even though the changes in the dictionary were modest and the dictionary’s editors still tended to favor educated writing.⁸⁵

The controversy surrounding the publication of Webster’s Third was not lost on the Supreme Court. In *MCI Telecommunications Corp. v. American Tel. & Tel. Co.*,⁸⁶ Justice Scalia refused to accept that *modify* could mean “to make a basic or important change,” as suggested in the lone definition by Webster’s Third. He stated that “[s]uch intentional distortions, or simply careless or ignorant misuse, must have formed the basis for the usage that Webster’s Third, and Webster’s Third alone, reported.”⁸⁷ Justice Scalia continued that this characterization “is not an unlikely hypothesis. Upon its long-awaited appearance in 1961, Webster’s Third was widely criticized for its portrayal of common error as proper usage.”⁸⁸

Despite this one-time criticism, Webster’s Third is a widely respected authority both in academia and among the Supreme Court justices who cite this dictionary more than any other.⁸⁹

What then can the content and structure of Webster’s Third tell us about the Sense-Ranking Fallacy? Perhaps anticipating their readers’ cognitive predisposition to ascribe meaning to the dictionary’s ranking of senses, the editors of Webster’s Third provide this word of caution:

The system of separating by numbers and letters reflects something of the semantic relationship between various senses of a word. It is only a lexical convenience. *It does not evaluate senses or establish an enduring hierarchy of importance among them. The best sense is the one that most aptly fits the context of an actual genuine utterance.*⁹⁰

85. *Id.* For a discussion of the controversy surrounding the publication of Webster’s Third, see generally JAMES SLEDD & WILMA R. EBBIT, *DICTIONARIES AND THAT DICTIONARY* (1962); see also HERBERT C. MORTON, *THE STORY OF WEBSTER’S THIRD: PHILIP GOVE’S CONTROVERSIAL DICTIONARY AND ITS CRITICS* (1994).

86. 512 U.S. 218 (1994).

87. 512 U.S. 218, 225–26 (1994) (quoting WEBSTER’S THIRD NEW INTERNATIONAL DICTIONARY 1452 (1981)).

88. *Id.* at 228 n.3.

89. See Thumma & Kirchmeier, *supra* note 27, at 262–63.

90. WEBSTER’S THIRD NEW INTERNATIONAL DICTIONARY 17a (1971) (emphasis added). A Westlaw search reveals over 10,000 state and federal cases (the maximum search return) that rely upon Webster’s Third for definitions of terms. A search for the last line of the cautionary note above reveals five citations only.

This word of caution represents an express disavowal of any intent to rank meanings according to their semantic importance. The editors have here undermined any claim that a given meaning is primary or ordinary merely because it is ranked first. Instead, they have stated emphatically that the ranking of senses is “only a lexical convenience” and that they do not “evaluate senses” with the intent of “establish[ing] an enduring hierarchy.”⁹¹ The editors have also provided perhaps the best advice imaginable for jurists struggling with the interpretation of a given word in a statute, namely “[t]he best sense is the one that most aptly fits the context of an actual genuine utterance.”⁹² Thus, those seeking “to make a fortress” out of the sense ranking in Webster’s Third are sent back to the statute to locate contextual cues to aid in selecting the appropriate meaning.

Setting aside this disavowal, how does Webster’s Third rank its senses? After all, the dictionary is alone among those cited by the *Muscarello* majority in cautioning against finding meaning in its ranking of senses. One might assume that these rankings are somehow anomalous, or follow a more haphazard course than other dictionaries. This is not the case. Webster’s Third ranks its senses historically.

The order of senses is historical: the one known to have been first used in English is entered first. This reordering does not imply that each sense has developed from the immediately preceding sense. . . . Sometimes an arbitrary arrangement or rearrangement is the only reasonable and expedient solution to the problems of ordering senses.⁹³

The explanatory note concedes that “[s]ometimes an arbitrary arrangement” is necessary, but does not tell the reader how to distinguish arbitrary rankings from principled ones. The editors are here simply acknowledging that there are times when no meaningful or conclusive method exists for determining which meanings occurred first in the language. Consequently, we must assume that at least some of the first-place sense rankings in Webster’s Third are listed first for *no reason at all*.

More importantly, even assuming that the instances of arbitrary ranking are rare, the explanatory note reveals the futility of seeking

91. *Id.*

92. *Id.*

93. *Id.*

to establish a primary or ordinary meaning by looking to see which sense is ranked first. Such a system tells us little about current, ordinary usage. For example, the first sense of *carry*, the one favored by the majority, dates back to the fourteenth century CE. Does the majority seriously contend that we must assume that Congress intended this centuries-old meaning? Such intent is, of course, possible; but the first-place ranking of that sense in Webster's Third does nothing to signal such Congressional intent.

Imagine a statute that prohibits the public from "using or carrying pipes" within twenty feet of a federal courthouse. What kind of pipes might Congress have intended to keep away? Perhaps the statute was passed in response to drug paraphernalia being smuggled into jury rooms.⁹⁴ Or perhaps it was passed to address the very real threat of pipe bombs.⁹⁵ Maybe the statute seeks to prevent federal judges from routinely being upstaged by the "elementary deductions" of pipe-smoking, Victorian-era sleuths.

Almost certainly the last pipe proscription that would come to mind is a statute designed to prevent the federal courthouse from being constantly regaled with Renaissance-style music, but this is precisely the direction in which the majority's First-Definition-Jurisprudence would point. Webster's Third gives the following as its first definition for *pipe*.

1a: a tubular wind instrument; *specifically*: a small fipple [sic] flute held in and played by the left hand b: one of the tubes of a pipe organ.⁹⁶

This venerable definition of *pipe*, the dictionary tells us, dates from "before the 12th century."⁹⁷

The example illustrates the absurdity of appealing to sense ranking in a dictionary that ranks its definitions historically. It is entirely plausible that Congress sought to regulate the meaning of *carry* supported by the *Muscarello* majority, but it is entirely irrelevant that this meaning happens to correspond to the first definition in Webster's Third.

94. See, e.g., *Tanner v. United States*, 483 U.S. 107 (1987).

95. *Pipe Bomb Damages San Diego Courthouse*, S.F. CHRON., May 5, 2008.

96. See WEBSTER'S, *supra* note 90, at 976.

97. *Id.*

2. *The Oxford English Dictionary*

Perhaps out of nostalgia for his *alma mater*,⁹⁸ Justice Breyer pays special attention to the sense-ranking in the *Oxford English Dictionary* (“OED”). Not only does he note that *carry*₁ is ranked first among the term’s fifty-six definitions in the OED,⁹⁹ but he takes special care to observe that *carry*₂ is ranked “twenty-sixth.”¹⁰⁰ This is the only instance in which Justice Breyer notes the ranking given for *carry*₂ as well as that for *carry*₁. The OED, like Webster’s Third, ranks its senses historically.¹⁰¹

A separation from the first definition to the twenty-sixth definition seems like a compelling lexicographical distance. What accounts for this wide disparity? The answer may be found in the explanatory note, ignored by the majority opinion, which appears at the top of the entry for the verb-form of *carry*. It says:

From the radical meaning which includes at once ‘to remove or transport’, and ‘to support or bear up’ arise two main divisions, in one of which (I.) ‘removal’ is the chief notion, and ‘support may be eliminated’ . . . ; while in the other (II.) ‘support’ is the prominent notion, and ‘motion’ (though usually retained) may entirely disappear. *For the former take is now largely substituted.*¹⁰²

This explanatory note sheds very important light on the tacit assumptions of the majority opinion. There are forty-two definitions of the non-phrasal verb form of *carry* in the OED,¹⁰³ but only two senses. The first sense (represented in the text by Roman numeral I) is that of *carry*₁, and is defined consistently with the definitions employed by the *Muscarello* majority:

98. JEFFREY TOOBIN, *THE NINE: INSIDE THE SECRET WORLD OF THE SUPREME COURT* 94 (2007).

99. *Muscarello v. United States*, 524 U.S. 125, 128 (1997) (citing 2 OXFORD ENGLISH DICTIONARY 919 (2d ed. 1989)).

100. *Id.* (citing 2 OXFORD ENGLISH DICTIONARY 921 (2d ed. 1989)).

101. 1 OXFORD ENGLISH DICTIONARY xxix (“[T]hat sense is placed first which was actually the earliest in the language: the others follow in order in which they have arisen”).

102. 2 OXFORD ENGLISH DICTIONARY 919 (2d ed. 1989) (emphasis added).

103. *Id.* The remaining definitions are classified under a third Roman-numeral, not mentioned in the explanatory note, which includes the phrasal-verb uses of *carry*, such as *carry on* or *carry away*. *Id.* at 921–22.

I. To transport, convey while bearing up.

**Of literal motion or transference in space.*¹⁰⁴

The first twenty-four definitions are classified under this single sense. The second sense (represented in the text by Roman numeral II) is that of *carry*₂, and is defined consistently with the meaning favored by the *Muscarello* dissent:

II. To support, sustain.

**With more reference to motion.*¹⁰⁵

The next eighteen definitions of *carry*, including the twenty-sixth definition referenced by the Court, are listed under this second sense.¹⁰⁶ Thus, in the broadest terms there are not twenty-six or forty-two senses of *carry* in the OED, there are only two. And why does the OED list one sense before the other? Simply because the first recorded use of *carry*₁ was written down circa 1320 CE,¹⁰⁷ while the first recorded use of *carry*₂ was drafted circa 1380 CE.¹⁰⁸

Not only does this explanatory note reveal that there are only two broad senses of *carry* defined in the OED—senses ranked by an accident of history—the note likewise suggests that the first sense, the one favored by the majority as the “ordinary English meaning” is waning into obsolescence. It says that “[f]or the former (i.e. *carry*₁), *take* is now largely substituted.” That is, the editors of the OED, perhaps the most venerable of all English dictionaries, suggest that *carry*₁ is on its way out of the lexicon entirely.

The majority does not define its use of the term *ordinary meaning*, and indeed, reference to the OED’s *first* definition (the *legal* definition) of the term *ordinary* would be profoundly unhelpful.¹⁰⁹ Still, the OED’s entry and explanatory note on *carry*

104. *Id.* at 919.

105. *Id.* at 920.

106. *Id.* at 921–22.

107. *Id.* at 919 (“c1320 Dugdale, *Monast.* (1661) II. 102 De libero transitu cum plaustis carectis & equis . . . carandi decimas suas et alia bona sua.”).

108. *Id.* at 921. (“c1380 WYCLIF Sel. Wks. III. 266 Carie a swerd in a scaberge.”). One might argue that the Wyclif reference is more germane to the statute at issue in *Muscarello* because it refers to those who “Carie a swerd in a scaberge.” *Id.* But there is no suggestion that these *swerds* were carried to a drug deal.

109. 10 OXFORD ENGLISH DICTIONARY 912 (“1. *Law.* Of a judge: having regular jurisdiction, i.e. exercising authority by virtue of office and not by special deputation; esp. empowered *ex officio* to take jurisdiction of ecclesiastical or spiritual cases.”).

are best cited, not to support the majority's conclusions, but to show that in contemporary usage, *carry*₁ is anything but ordinary.

3. *The Random House Dictionary of the English Language*

The *Random House Dictionary of the English Language, Unabridged* (the "Random House") employs an entirely different method for ranking its senses than those dictionaries cited above. Instead of ranking its senses historically, it does so by an appeal to the senses' statistical frequency:

[T]he most frequently encountered meaning generally comes before less common ones. Specialized senses follow those in the common vocabulary, and rare, archaic, and obsolete senses are listed last.¹¹⁰

This mode of sense-ranking ought to have important implications for the meaning of *carry* in *Muscarello*. Though the majority does not define the phrase "ordinary meaning," that phrase could easily be interpreted as implying that a given meaning is the "most frequently encountered."¹¹¹ Does this sense-ranking system reveal something about the ordinary meaning of *carry* that the other sense-ranking regimes do not? Unfortunately, no.

The problem is that the sense-ranking method employed by the editors of the Random House is impressionistic and unreliable. This is hardly a fatal criticism of the dictionary. It seems unlikely that a dictionary's editors would assume that their ordering of senses would be deemed legally significant. Whatever the case may be, these impressionistic sense-rankings do not necessarily correlate with the statistical frequency of a given sense.

For example, by far the most commonly employed sense of the word *deal* is that sense which suggests a particular amount: i.e., *great deal* or *good deal*.¹¹² This is true both in spoken English and in

110. RANDOM HOUSE DICTIONARY OF THE ENGLISH LANGUAGE, UNABRIDGED xxxii (2d ed. 1987).

111. Indeed, the OED provides the following as one of its definitions—though, admittedly, not its first definition—of *ordinary*: "2. d. Of language, usage, discourse, etc.: *that most commonly found or attested.*" 10 OXFORD ENGLISH DICTIONARY 912 (emphasis added).

112. DOUGLAS BIBER, SUSAN CONRAD & RANDI REPPEN, CORPUS LINGUISTICS: INVESTIGATING LANGUAGE STRUCTURE AND USE 37 (1998); see RANDOM HOUSE DICTIONARY, *supra* note 110, at 512 ("21. an indefinite but large quantity, amount, extent, or degree (usually prec. by *good* or *great*).").

written texts.¹¹³ Thus, if the Random House ranks its senses by statistical frequency as it claims, we would expect to find this sense listed first. In fact, this sense is listed twenty-first.¹¹⁴

What accounts for this dramatic difference between statistical frequency and sense ranking? The answer is that human intuition about the frequency of lexical items is often unreliable.¹¹⁵ As one commentator has put the issue, information about lexical frequency “is not susceptible to recovery via introspection.”¹¹⁶ This is because “humans tend to notice unusual occurrences more than typical occurrences.”¹¹⁷ Nor are professional linguists immune from this human deficiency.¹¹⁸

Perhaps, in light of the growing recognition of this limitation on human intuition, the qualifier *generally* was added to the Random House dictionary’s explanatory note referenced above.¹¹⁹ In the first edition, this note read merely “the most frequently encountered meaning appears as the first definition.”¹²⁰ *Generally* was not added until the publication of the second edition. Further, like the “sometimes . . . arbitrary arrangement” of senses in the Webster’s Third, the Random House fails to convey in what instances the dictionary departs from this “general” practice. Thus, even if we were to accept the Random House sense-rankings as accurate, we would not be able to tell, based on this qualifier, when the sense ranking is frequency-based and when it is not.

113. See BIBER ET AL., *supra* note 112, at 38.

114. *Id.* at 40.

115. See SUSAN HUNSTON, *CORPORA IN APPLIED LINGUISTICS* 20 (2002) (“Although a native speaker has experience of very much more language than is contained in even the largest [database], much of that experience remains hidden from introspection.”).

116. See TONY MCENERY & ANDREW WILSON, *CORPUS LINGUISTICS* 12–14 (2d ed. 2003) (“[H]uman beings have only the vaguest notion of the frequency of a construct or a word. Natural observation of data seems the only reliable source of evidence for such features as frequency. . . . There are certain types of language data which can only be gathered accurately from a corpus. . . . [Frequency information] is not susceptible to recovery via introspection.”).

117. See BIBER ET AL., *supra* note 112, at 3.

118. J. Charles Alderson, *Judging the Frequency of English Words*, 28 *APPLIED LINGUISTICS* 383, 383 (2007) (examining empirically the frequency judgments of professional linguists and noting that “judgments by professional linguists do not correlate highly with [objective measures of word frequency].”).

119. See *supra* note 72.

120. See *RANDOM HOUSE DICTIONARY OF THE ENGLISH LANGUAGE, UNABRIDGED I* (1968).

More importantly, the explanatory note makes no claim to a sense-ranking based on the most frequently *occurring* meaning, i.e., the actual sense frequency in the language at large. Rather, the note claims to present “the most frequently *encountered* meaning.”¹²¹ This raises an important question: Just where were these senses encountered? The answer is that the senses were found in the “large citation file” compiled by the dictionary’s editorial staff.¹²² For the purposes of statistical frequency, such a citation file is problematic in several respects. First, the file is compiled by hand, almost exclusively of written sources.¹²³ When scanning these sources for citations, editors often gravitate toward the more prestigious authors, thus creating a skewed sample that may enshrine archaic usage.¹²⁴ Second, subject to the same cognitive limitations discussed above, the dictionary’s editors are more likely to notice unusual occurrences than typical ones.¹²⁵ The editors will be looking for sample sentences that are either particularly illustrative of a common meaning, or helpful in exemplifying a rare use.¹²⁶ Finally, nothing in the explanatory note suggests that the Random House editors intended

121. See RANDOM HOUSE, *supra* note 110 (emphasis added).

122. See RANDOM HOUSE, *supra* note 120, at v; LANDAU, *supra* note 51.

123. See LANDAU, *supra* note 51, at 104 (2d ed. 2001) (“Even very large citation files, because they are collected by a process of selection, cannot be used reliably for statistical studies of frequency because they are apt to be unrepresentative of the language as a whole. . . . [C]itation readers all too often ignore common usages and give disproportionate attention to uncommon ones, as the seasoned birder thrills at a glimpse in the distance of a rare bird while the grass about him teems with ordinary domestic varieties that escape his notice. By contrast, a corpus that is sensibly developed will, by design, be representative, at least to a much greater degree than any citation file.”).

124. See QUIRK, *supra* note 3, at 88 (“Given . . . the tendency to take citations from the more prestigious authors, it is not difficult to see the danger of a highly skewed lexicon emerging from principles designed precisely in the interests of objective generality.”); HART & SACKS, *supra* note 48.

125. See BIBER ET AL., *supra* note 112, at 3.

126. *Id.* at 26 (“[C]itation slips represent only those contexts that a human reader happens to notice (in some cases representing only the more unusual uses.”). In compiling citation files, the selection of the “more unusual uses” is not only uncontroversial, it is something of an industry standard. See, e.g., GEOFF BARNBROOK, *DEFINING LANGUAGE: A LOCAL GRAMMAR OF DEFINITION SENTENCES* 46 (“Even the OED, despite its comprehensively descriptive aims, suffers from the lack of a properly representative corpus. . . . Detailed instructions were given to the [citation compilers] in the later stages, but these make it clear that the basis of selection would not produce a fully representative sample. [They were told], ‘Make a quotation for every word that strikes you as rare, obsolete, old-fashioned, new, peculiar, or used in a particular way. . . . Make as many quotations as you can for ordinary words, especially when they are used significantly, and tend to by context to explain or suggest their own meaning.”).

to establish a statistically reliable sense-ranking regime from which one could determine whether a given meaning is more or less ordinary. Thus, in appealing to the Random House sense-ranking regime in order to determine ordinary meaning, we are asking the dictionary to perform a task it was simply not designed to perform.

This is not to suggest that it is impossible to correctly determine the statistical frequency of a given word or word sense. As will be demonstrated below, advances in computational databases and a linguistic methodology referred to as corpus linguistics provide tools for accurately determining the statistical frequency of words and word senses.¹²⁷ Though such quantitative methods make frequency-based sense rankings a theoretical possibility, efforts to produce such rankings are not without their challenges,¹²⁸ and they have played no part in the development of any of the dictionaries cited by the majority in *Muscarello*.¹²⁹

4. *Sense-ranking: a conclusion*

After examining each of the dictionaries cited by the majority, and after having taken note of their structure and design, it becomes manifestly clear that the majority's implicit assumption that sense-ranking should be legally significant is fallacious.

127. Indeed, several English dictionaries have been compiled with the aid of frequency data obtained from sophisticated electronic corpora. *See, e.g.*, MARK DAVIES & DEE GARDNER, A FREQUENCY DICTIONARY OF AMERICAN ENGLISH (2010); LONGMAN DICTIONARY OF CONTEMPORARY ENGLISH (4th ed. 2008). Likewise, several foreign language frequency dictionaries have been compiled to serve, among other things, as pedagogical resources. *See, e.g.*, TIM BUCKWALTER & DILWORTH PARKINSON, A FREQUENCY DICTIONARY OF ARABIC (2010); DERYLE LONSDALE & YVON LE BRAS, A FREQUENCY DICTIONARY OF FRENCH (2009); RANDALL JONES & ERWIN TSCHIRNER, A FREQUENCY DICTIONARY OF GERMAN (2006); MARK DAVIES, A FREQUENCY DICTIONARY OF SPANISH (2006); *see also*, Stephen C. Mouritsen, A Frequency Dictionary of Arabic Newsprint (March 8, 2007) (unpublished M.A. thesis, Brigham Young University) (on file with author).

128. *See* LANDAU, *supra* note 55, at 303 (“To be honest, it is still not always possible to distinguish frequencies of meanings with any great certitude because so much judgment is involved in distinguishing between closely allied meanings in the first place.”).

129. American dictionaries such as the Webster's Third and the Random House have been particularly reluctant to incorporate information from electronic corpora into their design. *Id.* at 289–91 (“It is now apparent that some time in the 1970s the focus of innovative work in commercial lexicography shifted from the United States to Great Britain. . . . [I]n spite of showy graphics and ballyhooed usage notes, there have been very few meaningful advances in commercial American lexicography in the past twenty years. . . . [I]s it any wonder that American dictionary houses are reluctant to make the investment to build their own American corpus?”).

B. The Etymological Fallacy

Immediately following the Court's appeal to the sense-ranking of *carry*, Justice Breyer seeks to strengthen his claim by examining the etymology of *carry*.¹³⁰ To show this, he cites several dictionaries showing that

The origin of the word "carries" explains why the first, or basic, meaning of the word "carry" includes conveyance in a vehicle. See BARNHART DICTIONARY OF ETYMOLOGY 146 (1988) (tracing the word from Latin "carum," which means "car" or "cart"); 2 OXFORD ENGLISH DICTIONARY, *supra*, at 919 (tracing the word from Old French "carier" and the late Latin "carricare," which meant to "convey in a car").¹³¹

As with the Sense-Ranking Fallacy discussed above, appeals to a term's original use in order to determine its contemporary meaning are as fallacious as they are common. Sometimes courts look to a term's Latin antecedents,¹³² sometimes to the Greek,¹³³ and sometimes to both.¹³⁴ But these historical antecedents are not necessarily related to contemporary usage. The great philologist Henry Sweet stated the matter as follows:

The meaning of a word in a given period of a given language is a matter of usage, and the fact of its having had a certain meaning at some earlier period or in some cognate language does not necessarily afford any help in determining . . . its present meaning.¹³⁵

130. See Solan, *supra* note 20, at 2051 ("To bolster its choice of dictionary, the Court turned to the etymology of the word 'carry' and found that it shares its Latin origin with the word 'car'—good news for the government, *assuming (without presenting a reason) that word origins should make a legal difference.*") (emphasis added).

131. *Muscarello v. United States*, 524 U.S. 125, 128 (1997).

132. See, e.g., *United States v. Lombardo*, 241 U.S. 73, 76 (1916) ("The word 'file' was not defined by Congress. No definition having been given, the etymology of the word must be considered and ordinary meaning applied. The word 'file' is derived from the Latin word 'filum,' and relates to the ancient practice of placing papers on a thread or wire for safe keeping and ready reference").

133. See, e.g., *Acuff-Rose Music, Inc. v. Campbell*, 972 F. 2d 1429, 1440 (6th Cir. 1992) ("The etymology of the word has direct relevance to this case. The term comes from the Greek *parodeia*, meaning 'a song sung alongside another.' The musical parody is thus the very archetype of the genre.")

134. See, e.g., *Miller v. California*, 413 U.S. 15, 19 n.2 (1973) (citing the Latin antecedents of *obscene* and the Greek antecedents of *pornography*).

135. HENRY SWEET, *THE PRACTICAL STUDY OF LANGUAGES: A GUIDE FOR TEACHERS*

Characterized as “one of the most pernicious of popular *idées fixes*” about language,¹³⁶ the Etymological Fallacy has fallen under the condemnation of both linguists¹³⁷ and logicians.¹³⁸ “[P]recisely because meaning changes over time,” says one commentator, “the meaning of a word cannot be established from its etymology.”¹³⁹

The notion that we may accept a given meaning as valid simply because its etymology is consistent with our proffered meaning is unsustainable because it would lead to absurd results: *December* would quite literally mean *October*, *anthology* would mean a *bouquet of flowers*.¹⁴⁰

The Court’s appeal to etymology, like its references to the usage of the King James Bible, *Robinson Crusoe*, and *Moby Dick*, all serve to undermine the claim that *carry*₁ is the ordinary meaning of the term. Imagine that I were able to mimic perfectly the English of Defoe, of Melville, or of the Bible; or that I could speak fluently the Latin or Old French of those who coined the term *carry*. In what context would such speech be considered *ordinary*?

AND LEARNERS 88 (1906).

136. See QUIRK, *supra* note 3, at 87.

137. ANDREW L. SIHLER, LANGUAGE HISTORY: AN INTRODUCTION 131–32 (“[I]t cannot be overemphasized that a word’s etymology is not its meaning. And emphasis is truly called for, here: *the etymological fallacy* is immemorial, and shows no signs of exhaustion . . . it is enshrined in the very word *etymology*, from a Greek term that means ‘the study of true [meaning].’”); JOHN LYONS, LANGUAGE AND LINGUISTICS: AN INTRODUCTION 55 (1981) (“[T]he etymological fallacy is the assumption that the original form or meaning of a word is, necessarily and by virtue of that very fact, its correct form or meaning. This assumption is widely held. How often do we meet the argument that because such and such a word comes from Greek, Latin, Arabic, or whatever . . . the correct meaning of the word must be what it was in the language of origin! The argument is fallacious, because the tacit assumption of an originally true or appropriate correspondence between form and meaning, upon which the argument rests, cannot be substantiated.”); THOMAS PYLES & JOHN ALGEO, THE ORIGINS AND DEVELOPMENT OF THE ENGLISH LANGUAGE 239–40 (1993) (“The belief is widespread, even among some quite learned people, that the way to find out what a word means is to find out what it previously meant—or, preferably . . . what it originally meant. . . . [S]uch an appeal to etymology . . . is as unreliable as would be an appeal to spelling to determine modern pronunciation.”).

138. NIGEL WARBURTON, THINKING FROM A TO Z 59 (2000) (“It does not follow that because a word originally had one meaning that it will always continue to have that same meaning or even one directly related to it.”).

139. MICHAEL STUBBS, WORDS AND PHRASES: CORPUS STUDIES OF LEXICAL SEMANTICS 172 (2001).

140. See SIHLER, *supra* note 137, at 133.

Words lose their meaning; often the earliest meanings of a term become obsolete.¹⁴¹ In fact, this is precisely what the comment in the OED's entry for *carry* suggests is happening to *carry*.¹⁴² Because semantic shift may render a word's meaning unrecognizable from—and irreconcilable with—its prior meaning, and because the OED suggests that this sort of semantic shift is taking place with *carry*, the majority's fallacious appeal to the etymology of *carry* is unavailing.

C. Condorcet and Some Problems with Aggregating Definitions

In the end, the *Muscarello* Court relies on seven dictionaries—three unabridged, general-use dictionaries¹⁴³; two etymological dictionaries¹⁴⁴; one legal dictionary¹⁴⁵; and one thesaurus¹⁴⁶—in order to demonstrate that *carry*_i is the “ordinary English meaning” of *carry*. In response, Justice Ginsburg claims that “[u]nlike the Court, I do not think dictionaries, surveys of press reports, or the Bible tell us, dispositively, what ‘carries’ means embedded in § 924(c)(1).”¹⁴⁷ Still, the dissent is unable to resist the impulse to appeal to a dictionary, noting that “the only *legal* dictionary the Court cites . . . defines ‘carry arms or weapons’ restrictively.”¹⁴⁸ In this respect, the *Muscarello* decision is a battle of dictionaries, with the majority prevailing by sheer force of numbers.

In compounding all of these dictionary citations, the majority appears to be operating under the assumption that where one dictionary is good, seven are better, or rather that the combined expertise of the editorial boards of several dictionaries is more likely to reveal the correct ordinary meaning of a given term. Certainly the *Muscarello* Court is not alone in entertaining this assumption,¹⁴⁹ nor

141. See PYLES & ALGEO, *supra* note 137, at 239.

142. See OXFORD ENGLISH DICTIONARY, *supra* note 76.

143. OXFORD ENGLISH DICTIONARY (2d ed. 1989); WEBSTER'S THIRD NEW INTERNATIONAL DICTIONARY (1986); RANDOM HOUSE DICTIONARY OF THE ENGLISH LANGUAGE UNABRIDGED (2d ed. 1987).

144. BARNHART DICTIONARY OF ETYMOLOGY (1988); OXFORD DICTIONARY OF ENGLISH ETYMOLOGY (C. Onions ed., 1966).

145. BLACK'S LAW DICTIONARY (6th ed. 1990).

146. WEBSTER'S DICTIONARY OF SYNONYMS (1942).

147. *Muscarello v. United States*, 524 U.S. 124, 142–43 (1998) (Ginsburg, J., dissenting).

148. *Id.* at 142 n.2 (emphasis added).

149. For example, on several occasions, the Justices of the Supreme Court have used the same three general-use dictionaries discussed above (or versions thereof) to define a single term

does the Court's reliance upon the compounded expert opinions of these lexicographers seem terribly controversial. In fact, the notion that many lexicographical minds are more likely to arrive at a correct determination of the ordinary meaning of *carry* finds theoretical support in a well-known approach to analyzing democratic decision-making known as the Condorcet Jury Theorem.¹⁵⁰ The Jury Theorem was an effort by its creator, Nicholas de Condorcet, to "justify the use of majority rule and to assess the optimal size of a deliberative body."¹⁵¹ The Theorem states that "if it is more probable than not that [a decision-maker] will decide in conformity with the truth, the more the number of [decision-makers] increases, the greater the probability of the truth of the decision."¹⁵² That is, if a certain class of decision-makers is, say, fifty-one percent likely to arrive at a correct result, then the more decision-makers from this class are added to the group, the more the likelihood of a correct group decision approaches certainty.¹⁵³ When applied to deliberative

in a single decision. *See, e.g.*, *Gross v. FBL Fin. Servs. Inc.*, 129 S. Ct. 2343, 2350 (2009) (defining *because of*); *Crawford v. Metro. Gov't Cnty., Tenn.*, 129 S. Ct. 846, 850 (2008) (defining *oppose*); *MCI Telecomm. Corp. v. Am. Tel. & Tel. Co.*, 512 U.S. 218, 225 (1994) (defining *modify*); *Victor v. Nebraska*, 511 U.S. 1, 14 (1994) (defining *moral certainty*); *Pembaur v. Cincinnati*, 475 U.S. 469, 481 n.9 (1986) (defining *policy*). The lower federal courts have similarly employed this multi-dictionary approach. *See, e.g.*, *Fathauer v. United States*, 566 F.3d 1352, 1355–56 (Fed. Cir. 2009) (defining *employee*); *United States v. Loney*, 219 F.3d 281, 284 (3rd Cir. 2000) (defining *connection*); *Commodity Trend Serv. v. Commodity Futures Trading*, 233 F.3d 981, 989 (7th Cir. 2000) (defining *customer*); *Miss. Poultry Ass'n v. Madigan*, 31 F.3d 293, 310 n.4 (5th Cir. 1994) (defining *same*); *United States v. Chambers*, 985 F.2d 1263, 1268 (4th Cir. 1993) (defining *manager*); *Donovan v. Anheuser-Busch, Inc.*, 666 F.2d 315, 327 (8th Cir. 1981) (defining *such as*); *Hartford Nat'l Bank & Trust Co. v. United States*, 467 F.2d 782, 786 n.14 (2d Cir. 1972) (defining *welfare*). Numerous state courts have followed suit. *See, e.g.*, *Riepe v. Riepe*, 91 P.3d 312, 325 (Ariz. Ct. App. 2004) (defining *parent*); *Sy-Lenc of Washington, Inc. v. Starwood Urban Retail*, 829 A.2d 540, 547 (Md. 2003) (defining *limit*); *Emp't Sec. Comm. v. Gen. Motors Corp.*, 189 N.W.2d 74, 77 n.5 (Mich. Ct. App. 1971) (defining *lay off*); *Graev v. Graev*, 11 N.Y.3d 262, 272 (N.Y. 2008) (defining *cohabit*); *State v. Hathman*, 65 Ohio St. 3d 403, 412 (Ohio 1992) (defining *accessible*); *Lyon v. State*, 766 S.W.2d 879, 881 (Tex. Ct. App. 1989) (defining *occasion*); *Servais v. Port of Bellingham*, 904 P.2d 1124, 1130 (Wash. 1995) (defining *research* and *data*).

150. *See* ADRIAN HADDOCK, ALAN MILLAR, DUNCAN PRITCHARD, *SOCIAL EPISTEMOLOGY* 23–25 (2010).

151. Paul H. Edelman, *On Legal Interpretations of Condorcet Jury Theorem*, 31 J. LEGAL STUD. 327, 327 (2002).

152. NICOLAS DE CONDORCET, *CONDORCET: SELECTED WRITINGS* 48–49 (Keith Michael Baker ed., 1976).

153. CASS R. SUNSTEIN, *INFOTOPIA: HOW MANY MINDS PRODUCE KNOWLEDGE* 25 (2006); DAVID M. ESTLUND, *DEMOCRATIC AUTHORITY: A PHILOSOPHIC FRAMEWORK* 223

bodies with some expertise—i.e., deliberative bodies whose members, because they are experts, have greater than fifty-one percent likelihood of reaching a correct answer—we might expect the Condorcet result to be amplified.¹⁵⁴

While the mathematics of the Condorcet Jury Theorem have been characterized as “unassailable,”¹⁵⁵ the theorem’s utility is dependent upon several assumptions, not the least of which is the assumption that decision-makers will arrive at their results independently.¹⁵⁶ And of course, the Theorem only has predictive power when its first condition is satisfied—i.e., when the likelihood that the decision-makers in the group will reach correct result is better (or worse) than average. These limitations upon the utility of the Condorcet Jury Theorem have important implications for the question of deriving ordinary meaning from aggregated dictionary definitions.

1. Independence

With regard to independence, it has been said that “[t]he history of English lexicography usually consists of a recital of successive and often successful acts of piracy.”¹⁵⁷ While it is true that “no reputable dictionary today would take over entire sections of another work and print them verbatim,” it is likewise true that a comparison of common definitions across dictionaries will yield “few sharp discontinuities.”¹⁵⁸ This is because “[d]ictionary editors look at each

(2008) (“Suppose there are two options, and suppose each voter is independently 51 percent likely to choose the correct option (and 49 percent likely to choose the incorrect option): then among a group of 1,000 voters, the probability that the majority will vote for the correct option is approximately 69 percent. If the number of voters is increased to 10,000, then that probability rises to virtual certainty: 99.97. The probability that the majority will support the correct option tends toward certainty as the number of voters approaches infinity.”).

154. See SUNSTEIN, *supra* note 153, at 39 (“If experts are likely to be right, a statistical group of experts should have exactly the same advantage over individual experts as a statistical group of ordinary people has over ordinary individuals. Many expert minds are likely to be better than a few. A great deal of evidence supports this claim.”) (citations omitted).

155. See Edelman, *supra* note 151.

156. See ESTLUND, *supra* note 153, at 225–27.

157. See LANDAU, *supra* note 51, at 43.

158. *Id.* at 402; see GREEN, *supra* note 9, at 23 (“In the end any lexicographer has to face a basic fact: how many ways can there be of defining a given word? No reputable dictionary would simply reproduce material stolen from a predecessor, but it is inevitable, and generally accepted, that just a language demands that the word list must, while gradually expanding, remain relatively stable, so too must the definitions included in that list.”).

other's books, and though editors form their own opinions about what ground should be covered, they dare not depart too far from the area laid out by their competitors."¹⁵⁹

This reticence to depart from an established consensus may be explained with reference to conformity effects: the tendency of individuals to yield to the unanimous views of others.¹⁶⁰ Sometimes, individuals conform because "they believe that unanimous others cannot be wrong," at other times, individuals simply believe that public dissent is not worthwhile.¹⁶¹ In the realm of dictionary compilation, it is easy to see how conformity effects might influence a lexicographer's choice of definitions. "The definition quality of the leading American and British dictionaries is high, and in the vast majority of instances any major variation in treatment would be unwise."¹⁶² Even absent unanimity, the occurrence of a given sense of a word in a particularly prestigious source, like the Oxford English Dictionary, may result in the impulse to conform.¹⁶³

The Condorcet Jury Theorem requires "not causal independence, but statistical independence."¹⁶⁴ As long as each editor is as likely to arrive at a correct definition when acting alone as

159. *Id.*

160. CASS R. SUNSTEIN, DAVID SCHKADE, LISA M. ELLMAN & ANDRES SAWICKI, ARE JUDGES POLITICAL? AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY 74 (2006) ("A wide range of social science evidence shows conformity effects: When people are confronted with the views of unanimous others, they tend to yield."); *see also* RICHARD H. THALER & CASS R. SUNSTEIN, NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS 53-73 (2008).

161. *See* SUNSTEIN ET AL., *supra* note 161.

162. *See* LANDAU, *supra* note 51, at 402.

163. *Id.* ("All American and British dictionaries . . . owe a great debt to the OED, and one can find numerous similarities in wording as well as in sense division between the OED and later generations of dictionaries."). Sometimes the lack of independence among dictionary definitions is far more obvious than that produced by conformity effects. Publication houses often publish a single, definitive, unabridged dictionary from which many other ancillary dictionaries are derived. As noted by Justice Scalia in *MCI Telecommunications*, "[t]he Webster's New Collegiate Dictionaries . . . are essentially abridgments of that company's Webster's New International Dictionaries and recite that they are based upon those lengthier works." *MCI Telecommunications Corp. v. American Telephone & Telegraph Co.*, 512 U.S. 218, 226 n.2 (1994). In *Muscarello*, the Court relies on both the OED and the Oxford Dictionary of English Etymology ("ODEE") in order to establish the etymology of *carry*. But the original etymological materials in the OED were prepared by the same editor that provided this information for the ODEE, Dr. C. T. Onion. *See* OXFORD DICTIONARY OF ENGLISH ETYMOLOGY (C. T. Onions ed., 1966). Further, the ODEE was "based in the first instance on the OED," though it included "further research on the etymology of English." *Id.*

164. *See* ESTLUND, *supra* note 153, at 225.

she is when influenced by the decisions of others, the independence requirement is met.¹⁶⁵ The question of lexicographer independence, if it can be answered at all, must be answered empirically. But the tendency of dictionary-makers to rely—for various reasons—on the judgments of their predecessors and the pressure to produce a product that conforms to market expectations counsel against the too hasty conclusion that their judgments are independent.

2. *Accuracy and relevance*

Even accepting that all of the dictionaries cited by the *Muscarello* Court have arrived at their definitions independently, the question remains as to whether the editors of these dictionaries satisfy the first condition of the Condorcet Jury Theorem, i.e., whether they are more likely than not to provide correct answers to the question of ordinary meaning. Here, “everything depends on whether the relevant experts were in a position to offer good answers on the questions at hand.”¹⁶⁶

Two problems present themselves: First, as noted above, the editors of the general-use dictionaries cited by the *Muscarello* Court are not in the business of establishing ordinary usage. Instead, they seek to instantiate the realm of permissible or possible usage.¹⁶⁷ Second, even if these editors affirmatively assert that a given sense of a word in a given context is ordinary, typical, or common, it is not entirely clear that we would have reason to trust their judgment. Information about the statistical frequency of words or word senses is often inaccessible through introspection,¹⁶⁸ and, as noted above, linguistic training does not alleviate this common human deficiency.¹⁶⁹ Consequently dictionary editors, as all human beings, “have only the vaguest notion of the frequency of a construct or a word.”¹⁷⁰ They are likely to “give disproportionate attention to uncommon [uses],”¹⁷¹ and to favor the language use—likely the

165. *Id.* (“Statistical independence means that the probability of one voter, say Joe, getting the right answer is exactly equal to the probability of Joe getting the right answer given that Jane did.”).

166. *See* SUNSTEIN, *supra* note 153, at 41.

167. *See supra* notes 48–55 and accompanying text.

168. *See supra* notes 112–29 and accompanying text.

169. *See supra* note 116.

170. *Id.*

171. *See supra* note 123.

anachronistic language use—of “prestigious authors” (like the *Muscarello* Court’s Defoe and Melville).¹⁷² The editors themselves may have idiosyncratic views of language and language use,¹⁷³ and the methodologies they employ in preparing a dictionary may compound, rather than dampen, these deficiencies.¹⁷⁴

There may be circumstances in which an appeal to multiple dictionaries is helpful. After all, dictionaries differ both in the number of terms they define and in the number of senses they include in those definitions.¹⁷⁵ But even if a particular sense of a given term is defined in every available dictionary, we could not conclude that this sense of a term is *the* ordinary meaning. In the end, if we are seeking to establish whether a given sense is linguistically permissible,¹⁷⁶ then one dictionary may be enough. If we are seeking to determine which sense is the most commonly used in a given context, then one hundred dictionaries of the sort examined above would be insufficient.

III. AN APPEAL TO A DATABASE

Perhaps sensing that the argument for *carry*₁ has dwelt too long in references to historic usage, the majority seeks to modernize the discussion by appealing to a pair of electronic databases. “[T]o make certain,” says Justice Breyer, “that there is no special ordinary English restriction (unmentioned in dictionaries) upon the use of ‘carry’ . . . we have surveyed modern press usage, albeit crudely, by searching computerized newspaper data bases.”¹⁷⁷ These searches were conducted in a New York Times database found in Lexis/Nexis, and a U.S. News database found in Westlaw.¹⁷⁸ Justice Breyer then describes the search parameters and results as follows:

We looked for sentences in which the words “carry,” “vehicle,” and “weapon” (or variations thereof) all appear. We found thousands of such sentences, and random sampling suggests that many, perhaps

172. *See supra* note 124.

173. *See* QUIRK, *supra* note 3.

174. *See* BIBER ET AL., *supra* note 113.

175. *See supra* note 49.

176. *See supra* note 48.

177. *Muscarello v. United States*, 524 U.S. 125, 129 (1997).

178. *Id.*

more than one-third, are sentences used to convey the meaning at issue here, i.e., the carrying of guns in a car.¹⁷⁹

The majority's appeal to an electronic database in order to determine the ordinary meaning of *carry* has been characterized as unique¹⁸⁰ and promising.¹⁸¹ However, the structure of Justice Breyer's search is fatally flawed, and his conclusion about the search's results reveals an important weakness in the majority's argument.

To begin with, we must bear in mind the majority's chief contention. The majority has not simply asserted that *carry*₁ is a linguistically permissible reading of *carry*. Rather they have claimed that *carry*₁ is the "ordinary" and "primary" meaning of *carry*, while *carry*₂ is a "different, rather special" meaning.¹⁸² The entire case turns on whether *carry* ordinarily means *in a vehicle*. Thus, when Justice Breyer conducts his search for sentences containing the words *carry*, *weapon*, and *vehicle*, he is simply phrasing his question in the form of his desired answer.

Imagine that I am persuaded that people named Larry are more likely to be convicted of fraud. In an attempt to give my suspicion an air of statistical certainty, I enter the following search terms into a newspaper database: *Larry*, *convicted*, and *fraud*. What am I going to find? I will have undoubtedly uncovered numerous instances of guys named Larry being sent up for fraud. Have I proved my hypothesis?

In light of this question-begging data-entry, the majority's satisfaction with the returns is striking. "We found thousands of such sentences," says Justice Breyer, "and random sampling suggests that many, perhaps more than one-third, are sentences used to convey the meaning at issue here, i.e., the carrying of guns in a car."¹⁸³ Justice Ginsburg rightly pounces on this conclusion, noting that

179. *Id.*

180. See Hoffman, *supra* note 46, at 427 n.94 ("I have found no case besides *Muscarello* where the Court has used this method."). Hoffman's assessment of the Court's methodology is not favorable. *Id.*

181. See Solan, *supra* note 20, at 2059 ("[T]he use of this method is both creative and promising."). Note, *The Supreme Court 1997 Term, Leading Cases*, 112 HARV. L. REV. 355, 365 (1998) ("Empirical analysis of common usage will not by itself solve thorny questions of statutory interpretation, but it can provide a tool that will prove of use to . . . judicial interpreters.").

182. *Muscarello*, 524 U.S. at 128.

183. *Id.* at 129.

At issue here is not “carries” at large but “carries a firearm.” The Court’s computer search of newspapers is revealing in this light. Carrying guns in a car showed up as the meaning “perhaps more than one-third” of the time. One is left to wonder what meaning showed up some two-thirds of the time.¹⁸⁴

Still, Justice Ginsburg’s criticism misses an important point. Even if one hundred percent of Justice Breyer’s search returns contained sentences instantiating *carry*₁, the results would still be inconclusive. This is because Justice Breyer is searching for only for sentences containing *carry*, *firearm*, and *vehicle*. In these circumstances, we might expect to see only sentences referring to the *carry*₁. That only one-third of the sentences returned contained this usage is actually surprising.

If Justice Breyer’s search had included only *carry* and *firearm*, then a comparison of the number of sentences instantiating uses of *carry*₁ against the number of sentences instantiating *carry*₂ may have been illuminating. As it stands, Justice Breyer has merely confirmed that *carry*₁ is *sometimes* used (a fact that neither the majority, nor the dissent would find controversial), but he has certainly not confirmed that *carry*₁ is the *ordinary* meaning. In fact, because the search terms were skewed in favor of *carry*₁, the paucity of instances of *carry*₁ suggests that *carry*₁ is somewhat unusual.

The majority’s satisfaction with these results leaves one to wonder not just what meaning was employed in the other two-thirds of the returned sentences, but also what the majority now means by “ordinary” and “primary” meaning.¹⁸⁵ Certainly these returns cannot mean that *carry*₁ is the most frequent or common meaning. Rather, they suggest that *carry*₁ is the “different, rather special” meaning, as the OED comment suggests.

IV. PURPOSIVISM AND THE DISSENT

Leaving behind Justice Breyer’s textual argument, one feels compelled to pause for a moment’s reflection. If indeed we have just witnessed “one of the great textualist moments of our period,”¹⁸⁶ it

184. *Id.* at 143 (Ginsburg, J., dissenting).

185. *Smith v. United States*, 508 U.S. 223, 242 (1993) (Scalia, J., dissenting) (“The Court does not appear to grasp the distinction between how a word *can be* used and how it *ordinarily* is used.”).

186. *Henderson v. State*, 715 N.E.2d 833, 835 n.3 (Ind. 1999).

is difficult to suppress the sense of irony that this textualist manifesto was written by an avowed purposivist. In his 2006 book, *Active Liberty*,¹⁸⁷ Justice Breyer articulates a purpose-driven view of statutory interpretation¹⁸⁸—a view that “places more emphasis on statutory purpose and congressional intent.”¹⁸⁹ This interpretive regime encourages judges to “pay primary attention to a statute’s purpose in difficult cases of interpretation in which [statutory] language is not clear.”¹⁹⁰ The book likewise criticizes textualism’s reliance on history, tradition, uncertain linguistic norms, and canons of interpretation.¹⁹¹ “Near-exclusive reliance upon canons and other linguistic interpretive aids in close cases,” says Justice Breyer, “can undermine the Constitution’s democratic objective.”¹⁹²

Why then does Justice Breyer go to such great lengths to make his textualist case? The decision in *Muscarello* has been cited for the proposition that because purposivist and textualist judges are experimenting with one another’s interpretive tools, the divide between them is narrowing.¹⁹³ A further rationale for Justice Breyer’s foray into textualism is Justice Brennan’s famous Rule of Five, which states that “with five votes you can do anything around here!”¹⁹⁴ In

187. STEPHEN BREYER, *ACTIVE LIBERTY: INTERPRETING OUR DEMOCRATIC CONSTITUTION* (2006).

188. JEFFREY ROSEN, *THE SUPREME COURT: THE PERSONALITIES AND RIVALRIES THAT DEFINED AMERICA* 216 (2007) (“In . . . *Active Liberty*, Breyer argued that judges in constitutional cases should focus on the practical consequences of their decisions—in particular on the likelihood that a particular decision will promote political participation, rather than discourage it.”); JEFFREY TOOBIN, *THE NINE: INSIDE THE SECRET WORLD OF THE SUPREME COURT* 303 (2007) (“[I]n 2005, [Justice Breyer] did something that no justice had attempted in several generations—to write his own manifesto on the meaning of the Constitution. Characteristically, Breyer’s book . . . was hardly an airy philosophical treatise but a practical book by a practical man.”).

189. See BREYER, *supra* note 188, at 85.

190. *Id.*

191. *Id.* at 124–27.

192. *Id.* at 98–99.

193. John F. Manning, *What Divides Textualists from Purposivists?* 106 COLUM. L. REV. 70, 86 (2006) (“[There is] potent evidence of an increasingly text-based approach in the opinions of contemporary purposivists on the Supreme Court.”). *Id.* at 88 n.64 (“[P]urposivists on the present Court presume that a word’s ordinary or dictionary meaning supplies at least a useful starting point for discerning the purposes they seek.” (citing *Muscarello v. United States*, 524 U.S. 125, 127–32 (1998))). For further discussion of this narrowing divide, see Jonathan T. Molot, *The Rise and Fall of Textualism*, 106 COLUM. L. REV. 1 (2006).

194. H. JEFFERSON POWELL, *CONSTITUTIONAL CONSCIENCE: THE MORAL DIMENSION OF JUDICIAL DECISION* 16 (2008).

the five-to-four decision of *Muscarello*, that fifth vote came from committed textualist, Justice Clarence Thomas.¹⁹⁵

Was the majority's textualist experiment merely an exercise in consensus building? It seems clear that Justice Breyer would have preferred to keep the textual analysis to a minimum. He states that "because [the parties] argue the linguistic point at length, we too have looked into the matter *in more than usual depth*."¹⁹⁶ Then, when he arrives at his purposive argument, in which he examines both the statute's legislative purpose and history, one gets the palpable sense of Justice Breyer's relief. "We now explore more deeply the *purely legal question* of [what] Congress intended," he says.¹⁹⁷

Because this examination has focused on the text-based reasoning of the *Muscarello* Court, I will not dwell on the purposive section of the majority's argument in great detail, except to note that in addressing the statute's purpose, the opinion cites the very broad statements from the floor debates, all to the effect that § 924(c)(1) was passed "to persuade the man who is tempted to commit a federal felony to leave his gun at home."¹⁹⁸ The majority then compares these statements to those which suggest some members of Congress had a very narrow interpretation of *carry* in mind.¹⁹⁹ These are summarily dismissed by stating that none of them "purports to define the scope of the term *carries*."²⁰⁰ In this case, characterized by one commentator as a "food fight," and in which references to the King James Bible are answered by references to M*A*S*H, it is this interpretive move that is perhaps the most difficult to justify. The Congressional Record suggests that some members of Congress had

195. See Molot, *supra* note 193, at 29 ("Justices Scalia and Thomas are the only self-identified textualists on the Supreme Court.").

196. *Muscarello v. United States*, 524 U.S. 125, 128 (1997) (emphasis added).

197. *Id.* at 129 (emphasis added).

198. *Id.* at 132 (citing 114 CONG. REC. 22231 (1968) (statement of Rep. Poff); 114 CONG. REC. 22,243–22,244 (1968) (statutes would apply to "the man who goes out taking a gun to commit a crime") (statement of Rep. Hunt); *id.* at 22244 ("Of course, what we are trying to do by these penalties is to persuade the criminal to leave his gun at home.") (statement of Rep. Randall); *id.* at 22236 ("We are concerned . . . with having the criminal leave his gun at home.") (statement of Rep. Meskill)).

199. *Muscarello*, 524 U.S. at 133 ("We have found an instance in which a legislator referred to the statute as applicable when an individual 'has a firearm on his person'; . . . an instance in which a legislator speaks of 'a criminal who takes a gun in his hand'; and a reference . . . to a 'gun carried in a pocket.'" (citations omitted)).

200. *Id.*

a broad version of *carry* in mind (i.e. *carry*₁) while other members of Congress envisioned the narrow version (i.e. *carry*₂). If neither clearly expressed their meaning in the statute, it seems most consistent with the notions of fair notice and due process that this equation should balance out on the narrow side of the spectrum.

Perhaps with this in mind, Justice Ginsburg makes a plea for the application of the Rule of Lenity, stating that “[t]he sharp division in the Court on the proper reading of the measure confirms, ‘[a]t the very least, . . . that the issue is subject to some doubt. Under these circumstances, we adhere to the familiar rule that, ‘where there is ambiguity in a criminal statute, doubts are resolved in favor of the defendant.’”²⁰¹ Justice Breyer dismisses this plea for lenity, stating that “[t]he rule of lenity applies only if, after seizing everything from which aid can be derived, . . . we can make no more than a guess as to what Congress intended.”²⁰²

V. A CORPUS-BASED APPROACH

The Supreme Court’s decision in *Muscarello* unsatisfactorily addresses the meaning of *carry*, and in so doing throws into stark relief the Court’s difficulty with ordinary meaning generally. As noted above, the majority claims that *carry*₁ is the “first” or “primary” or “ordinary English meaning” of *carry* and that *carry*₂ is somehow a “different, rather special” meaning of *carry*.²⁰³ The Court

201. *Id.* at 148 (Ginsburg, J., dissenting) (citations omitted). Justice Ginsburg also responded to the text-based argument in her dissent, by, among other things, invoking the Second Amendment. *Id.* at 143 (“Surely a most familiar meaning is, as the Constitution’s Second Amendment (‘keep and bear Arms’) (emphasis added) and Black’s Law Dictionary, at 214, indicate: ‘wear, bear, or carry . . . upon the person or in the clothing or in a pocket, for the purpose . . . of being armed and ready for offensive or defensive action in a case of conflict with another person.’”). This equation of the meaning of *carry* with the meaning of *bear* in the Second Amendment would meet with what was surely an unsatisfying vindication in the Court’s opinion in *District of Columbia v. Heller*. 128 S. Ct. 2783, 2793 (2008) (“At the time of the founding, as now, to ‘bear’ meant to ‘carry.’ . . . When used with ‘arms,’ however, the term has a meaning that refers to carrying for a particular purpose—confrontation. In *Muscarello v. United States*, 524 U.S. 125 (1998), in the course of analyzing the meaning of ‘carries a firearm’ in a federal criminal statute, Justice Ginsburg wrote that ‘[s]urely a most familiar meaning is, as the Constitution’s Second Amendment’ We think that Justice Ginsburg accurately captured the natural meaning of ‘bear arms.’ Although the phrase implies that the carrying of the weapon is for the purpose of ‘offensive or defensive action,’ it in no way connotes participation in a structured military organization.”). Justice Ginsburg joined both dissents in *Heller*.

202. *United States v. Wells*, 519 U.S. 482, 499 (1997) (internal quotations omitted).

203. *Muscarello*, 524 U.S. at 127–28.

relies, among other things, upon the OED to support its claims—even though the OED takes a contrary position. The Court likewise gives in to the Sense-Ranking and Etymology fallacies discussed above. Further, the majority seems satisfied to demonstrate “more than one third” of the uses of *carry* found in a database appear to be instances of *carry*, even after entering a highly skewed query. Still, the Court insists that “[t]he *relevant linguistic facts* are that the word ‘carry’ in its ordinary sense includes carrying in a car.”²⁰⁴ But the claim that the ordinary sense of *carry* “includes” carrying in a car, is a far cry from the claim that this is the word’s first, primary, and “ordinary English” meaning. And what are the majority’s “linguistic facts,” but a hodgepodge of dictionary citations, literary references, and, as demonstrated above, fallacious inferences?

These and other deficiencies led one commentator to conclude that the *Muscarello* Court “is not entirely sincere when claiming that it actually sought the ordinary meaning.”²⁰⁵ Instead, the same commentator noted that the majority’s “interpretation falls within the range of the statute’s possible meanings”²⁰⁶ of *carry*, but not its ordinary meaning.

Though it is true that the majority’s claims are unsupported by the Court’s “linguistic facts,” I think, rather than insincerity, the Court has simply fallen victim to one of the law’s classic equivocations. When jurists speak of “ordinary meaning,” they simply are not always talking about the same thing.

Professors Hart and Sacks state that one of the purposes of the dictionary is to “answer[] hard questions of whether, in an appropriate context, a particular meaning is linguistically permissible.”²⁰⁷ It is this same concept of what is “linguistically permissible” that seems to actuate the reasoning in *Muscarello*. The Court states only that “one *can*, as a matter of ordinary English, ‘carry firearms’ in a wagon.”²⁰⁸ This “linguistically permissible” characterization of ordinary meaning also explains the majority’s satisfaction with finding *carry*, instantiated in only one-third of the sentences in its database search.²⁰⁹

204. *Id.* at 131 (emphasis added).

205. *See* Solan, *supra* note 20, at 2053.

206. *Id.*

207. HART & SACKS, *supra* note 48, at 1375–76.

208. *Muscarello*, 524 U.S. at 128 (emphasis added).

209. *Id.* at 129–130.

This approach is not without its critics. In the related case of *Smith v. United States*, Justice Scalia famously stated that “[t]he Court does not appear to grasp the distinction between how a word *can be* used and how it *ordinarily is* used.”²¹⁰ Justice Scalia has further observed that the Court’s job “is not to scavenge the world of English usage to discover whether there is any possible meaning.”²¹¹ “[O]ur job,” he says “is to determine . . . the *ordinary* meaning.”²¹²

In the end, if the *Muscarello* decision stands for anything, it is that the concept of ordinary meaning is “incompletely theorized.”²¹³ Though jurists commonly invoke the rule that statutory terms should be interpreted according to their ordinary meaning, there is no generally accepted view as to what *ordinary meaning* itself actually means, nor is there a universal justification for appealing to ordinary meaning.²¹⁴

Ordinary meaning has been characterized as “what . . . words would mean in the mouth of a normal speaker of English”²¹⁵ or how words “sound in the mind of a skilled, objectively reasonable user of words.”²¹⁶ These two characterizations are not identical. The skilled user of words and the normal speaker of English are not necessarily the same person. Still, both of these characterizations are grounded in the same principle, that ordinary meaning can be understood and analyzed in the context of ordinary usage.²¹⁷ This notion of ordinary

210. *Smith v. United States*, 508 U.S. 223, 242 (1993) (Scalia, J., dissenting) (emphasis in original).

211. *Chisom v. Roemer*, 501 U.S. 380, 410 (1991) (Scalia, J., dissenting).

212. *Id.*

213. Cass R. Sunstein, *Incompletely Theorized Agreements*, 108 HARV. L. REV. 1733, 1735 (1994-1995).

214. See Richard A. Posner, *Statutory Interpretation—In the Classroom and the Courtroom*, 50 U. CHI. L. REV. 800, 808 (1983) (“It is ironic that a principle designed to clarify should be so ambiguous.”). The same appears to be true for interpretive matters generally. See HART & SACKS, *supra* note 48, at 1169 (“The hard truth of the matter is that American courts have no intelligible, generally accepted, and consistently applied theory of statutory interpretation.”).

215. Oliver Wendell Holmes, Jr., *The Theory of Legal Interpretation*, 12 HARV. L. REV. 417, 417–18 (1899) (emphasis added); see also John F. Manning, *What Divides Textualists from Purposivists?* 106 COLUM. L. REV. 70, 92 (2006) (“Textualists give primacy to the *semantic* context—evidence about the way a reasonable person conversant with relevant social and linguistics practices *would have used the words*.” (emphasis added)).

216. Frank Easterbrook, *The Role of Original Intent in Statutory Construction*, 11 HARV. J.L. & PUB. POL’Y 59, 62 (1988).

217. See Frankfurter, *Some Reflections*, *supra* note 42, at 529 (“If a statute is written for

meaning as common usage is consistent with what the OED refers to as the linguistic definition of *ordinary*, which says “2. d. Of language, usage, discourse, etc.: that most commonly found or attested.”²¹⁸ Almost certainly without intending to, judges who have invoked this “ordinary-meaning-as-common-usage” characterization have implicated not merely a theory of ordinary meaning, but also a method for its analysis. Unlike the more qualitative notions of ordinary meaning—like the “linguistically permissible” approach discussed above—the question of the “common usage” of a statutory term may be answered quantifiably through a linguistic methodology called corpus linguistics.

A. Brief Introduction to Corpus Linguistics

In general terms, corpus linguistics may be thought of as a linguistic methodology that analyzes language function and use by means of an electronic database called a corpus.²¹⁹ A leading corpus linguist, Douglas Biber, has identified four unifying characteristics of the corpus approach:

1. it is empirical, analyzing the actual pattern of use in natural text;
2. it utilizes a large and principled collection of natural texts, known as a ‘corpus,’ as the basis for analysis;
3. it makes extensive use of computers for analysis, using both automatic and interactive techniques; and
4. it depends on both quantitative and qualitative analytical techniques.²²⁰

The data in the corpus are considered “natural” because they were not elicited for the purpose of study. That is, generally no one

ordinary folk, it would be arbitrary not to assume that Congress intended its words to be read with the minds of ordinary men. If they are addressed to specialists, they must be read by judges with the minds of the specialists. And so we assume that Congress uses common words in their popular meaning, as used in the common speech of men.”); *see also* *United States v. Lopez*, 590 F.3d 1238, 1249 (11th Cir. 2009) (“When a statutory term is undefined, courts give it its ‘ordinary meaning’ or ‘common usage.’”).

218. 10 OXFORD ENGLISH DICTIONARY 912 (2d ed. 1989).

219. Douglas Biber, *Corpus-based and Corpus-driven Analyses of Language Variation and Use* in THE OXFORD HANDBOOK OF LINGUISTIC ANALYSIS 159 (Bernd Heine & Heiko Narrog eds., 2009).

220. *See* BIBER ET AL., *supra* note 112, at 4.

asks the speakers or writers whose words are represented in the corpus to speak or write for the purpose of subjecting their words to linguistic scrutiny. Instead the architect of the corpus assembles her collection of speech and writing samples after the fact, from newspapers, books, transcripts of conversations, or interviews, etc.²²¹

Because the speech and writing sampled in a corpus are both naturally occurring, we may think of a corpus as “essentially tell[ing] us what a language is like, and the main argument in favor of using a corpus is that it is a more reliable guide to language use than native speaker intuition.”²²² Writing in 2004, Professor Lawrence Solan observed:

Over the past decade, a great deal of work has been published in the growing field of corpus linguistics. . . . Access to computers now makes it relatively simple to see how words are used . . . in common parlance. This allows judges to easily become their own lexicographers. If they perform that task seriously, they stand to learn more about how words are ordinarily used, than by today’s method of fighting over which dictionary is the most authoritative.²²³

Since that time, and certainly before, very little research has been published examining the legal question of “ordinary meaning” using a corpus-based approach.²²⁴ What follows is a brief outline of how an electronic corpus might be put to that task. I make no claim to any generalized theory of ordinary meaning. This Section will merely demonstrate that the question of which contested sense of a term is

221. TONY MCENERY, RICHARD XIAO & YUKIO TONO, *CORPUS-BASED LANGUAGE STUDIES: AN ADVANCED RESOURCE BOOK* 126 (2006) (“Any selection of texts is a sample. Whether or not a sample is ‘representative,’ however, depends first of all on the extent to which it is selected from the range of text types in the target population.”); TONY MCENERY & ANDREW WILSON, *CORPUS LINGUISTICS: AN INTRODUCTION* 77–82 (2001) (discussing sampling methods for corpus representativeness); VINCENT V. Y. OOI, *COMPUTER CORPUS LEXICOGRAPHY* 52–56 (1998).

222. See HUNSTON, *CORPORA IN APPLIED LINGUISTICS*, *supra* note 115, at 20.

223. See Solan, *supra* note 20, at 2059–60.

224. *But see* Brief for the Project on Government Oversight et al. as Amici Curiae Supporting Petitioners, *FCC v. AT&T Inc.*, No. 09-1279 (U.S. Nov. 16, 2010) (citing data from a number of corpora supporting the proposition that the phrase *personal privacy* in the Freedom of Information Act has reference to *human privacy* not *corporate privacy*); *see also*, Clark D. Cunningham & Charles J. Fillmore, *Using Common Sense: A Linguistic Perspective on Judicial Interpretations of “Use a Firearm,”* 73 Wash. U. L. Q. 1159, 1207–09 (1995) (using the British National Corpus to find examples of the competing senses of the verb *to use*).

ordinary in a given context is an empirical question and the answer may be quantified by an appeal to a corpus.

1. Selecting a corpus. As noted above, the chief concern in both corpus design and corpus selection is that of representativeness.²²⁵ In order for a corpus to be of any utility in the examination of ordinary usage, it must contain a representative sample. This sample must contain speech and text from the linguistic community which it purports to represent.

One way to think about selecting a speech community is to assume that the law ought to reflect the common usage of those it attempts to regulate. As noted above, we might “assume that Congress uses common words in their popular meaning, as used in the common speech of men”²²⁶ and according to the understanding of “the citizens subject to it.”²²⁷ Applying this assumption to a federal statute would suggest that we are looking for a broad-based corpus of contemporary American language use.

One such corpus is the Corpus of Contemporary American English (“COCA”), “the largest freely-available corpus of English, and the only large and balanced corpus of American English.”²²⁸

The corpus contains more than 410 million words of text and is equally divided among spoken [English], fiction, popular magazines, newspapers, and academic texts. It includes 20 million words each year from 1990–2010 and the corpus is also updated every six to nine months.²²⁹

Each of the 410 million words in the COCA is encoded for grammatical content or tagged.²³⁰ Nouns are electronically labeled as nouns, and verbs as verbs. In the COCA, this tagging process is

225. See *supra* note 81. The term *representativeness*, as used here, must not be confused with the decision-making heuristic of the same name from behavioral economics. See RICHARD H. THALER & CASS R. SUNSTEIN, *NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS* 27–31 (2008).

226. See Frankfurter, *Some Reflections*, *supra* note 42, at 529.

227. *Green v. Bock Laundry Mach. Co.*, 490 U.S. 504, 528 (1989) (Scalia, J. concurring).

228. See CORPUS OF CONTEMPORARY AMERICAN ENGLISH (COCA), <http://www.americancorpus.org/> (last visited Jan. 12, 2011).

229. *Id.*

230. PAUL BAKER, ANDREW HARDIE, & TONY MCENERY, *A GLOSSARY OF CORPUS LINGUISTICS* 154 (2006) (defining tagging as an “informal term for the act of applying additional levels of annotation to corpus data. A tag usually consists of a code, which can be attached to a phoneme, morpheme, word, phrase or longer stretch of text.”).

automated (though other corpora employ human editors to review this automated electronic tagging process).

If the linguistic community we are seeking to study is that governed by a federal statute, then the COCA presents itself as an obvious alternative for the analysis.

2. Query design. Like any other computational system, the results of a corpus query are only as good as the query itself. Because electronic corpora have almost never before been employed in the search for ordinary meaning in the legal sense, there are no guiding principles for how to do so. A few basic considerations emerge.

a. Syntax. The query ought to sweep in all uses of the term within a given grammatical class. In the COCA, this is done by surrounding the operative legal word (in this case *carry*) in brackets and then appending the expression $.[v^*]$, as in (1) below:

(1) $[carry].[v^*]$ ²³¹

This search will reveal all the *verbal* uses of the term *carry* (e.g., *carry*, *carries*, *carrying*, *carried*, and one archaic use of *carried*), while excluding other grammatical uses.²³²

b. Context. Generally, statutes are only concerned with the use of a term in a given context. As Justice Ginsburg correctly notes in her dissent, at issue in *Muscarello* “is not ‘carries’ at large but ‘carries a firearm.’”²³³ Thus, while examining how the operative term *carry* is used generally may be an important step in framing the analysis, the ultimate question in the case is not the ordinary meaning of *carry* at large, but the meaning of that term when it co-occurs with *firearm(s)* and its synonyms.

c. Dice Loading. Though context is an essential part of the question of ordinary meaning, some contextual information must be excluded. To ascertain ordinary meaning of *carry a firearm*, we are not simply asking the database to tell us how often *carry* is used to mean *convey*. Instead we are concerned with relative frequency, that is, how often *carry*₁ is used with reference to firearms versus how often *carry*₂ is used with reference to firearms. This was the fatal flaw in the *Muscarello* Court’s database analysis. If we include all three

231. I will use throughout the linguistic convention of labeling sample sentences and phrases with numbers in parenthesis.

232. Like most corpora, the COCA will exclude proper names from this search, like the proper name *Carry*.

233. See *Muscarello v. United States*, 524 U.S. 125 (1998) (Ginsburg, J., dissenting).

operative terms, *carry*, *firearm*, and *vehicle*, we would expect the database to show us only sentences instantiating *carry*. If, however, *vehicle* is excluded, we will be able to parse out the uses of *carry*₁ and *carry*₂ and determine which is more common.

3. Concordance lines and Key Words in Context (KWIC). Entering the search [carry].[v*] into the COCA reveals 82,687 uses of the verb *to carry*. These returns are listed in concordance lines, which display the word with its surrounding context.²³⁴ In these concordance lines, “each occurrence of the chosen word is presented on a single line, with the word in the middle and the context on each side.”²³⁵ This type of display is referred to as a Key Word in Context or KWIC display.²³⁶

In order to render the review of the ordinary meaning of *carry* more manageable than an examination of all 82,687 concordance lines, the COCA allows for a randomized sample to be displayed. Thus, by selecting the “500” icon, next to the heading “SAMPLE,” the corpus displays five hundred randomized instances of *carry* in KWIC format.

4. Results from the KWIC search. The first five hundred randomized concordance lines reveal several different uses of the verb *carry*. What follows is not an exhaustive account of the range of possible meanings of *carry*. As discussed above, when phrasal verbs like *carry on* are included, the OED features some fifty-six separate definitions for *carry*.²³⁷ As Justice Breyer rightly noted, only two meanings of *carry* are implicated in *Muscarello*.²³⁸ Before proceeding to a discussion of those two senses, I will briefly outline the other senses of *carry* found in the corpus, but only in the broadest terms.

The most common uses of *carry* found in the corpus were the non-physical senses of *carry*. These include metaphorical uses, as in (2), a broadcasting sense, where *carry* is used as a synonym for *feature*, as in (3), and an intangible sense in which what is carried is not a physical object, as in (4).

(2) The term *gender* while often used synonymously for *women* did not carry the same political weight.

234. See BIBER ET AL., *supra* note 112, at 26.

235. *Id.*

236. See Appendix III, Figure 8.

237. See *supra* notes 102–108.

238. *Muscarello*, 524 U.S. at 126.

(3) The Golf Channel now will carry early rounds and some tournaments.

(4) Our legacy copper wiring just can't carry the data to support HD-video streaming.

These non-physical senses of *carry* account for some twenty-nine percent of the uses found in the sample. None of these senses of *carry* is implicated in the statute.

Another very common use of *carry* is the phrasal-verb use of the term like *carry on* in (5) and *carried away* in (6).

(5) Stop complaining. You'd think I was killing you the way you carry on.

(6) If anything, he wants to win so badly he sometimes gets carried away and tries too hard.

These senses of *carry* account for approximately twenty-one percent of the uses of *carry* in the sample.

A common use of *carry*, is instrumental—an agent (typically a person) carries an object (like a gun) by means of some sort of instrument (like a car, a horse, or a stretcher, etc.). Thus, *carry*₁ sentences typically have similar structure to (7):

(7) <Agent> carries <object> *by, with, or in* an <instrument>.

By contrast, *carry*₂ sentences typically have a simpler structure, as in (8):

(8) <Agent> carries <object>.

A related use of *carry* occurs where the agent or subject of the sentence is itself an instrument of conveyance as in (9):

(9) Planes with 30 or fewer seats don't have to carry transponders that signal their location to air traffic control.

Grammatically, (9) is similar to the *carry*₂ paradigm in (8) because there is an agent (*planes*) and an object (*transponders*), but no instrument. However, in an abstract sense, the planes at issue (9) are being used as instruments for unstated agents. (People design the planes to carry or not carry transponders for some human purpose). I have classified all such uses of *carry* together under the heading of “inanimate” uses because the agents in such sentences are inanimate. These sentences account for some seven percent of the uses of *carry* in the sample.

Additionally, about four percent of the uses of *carry* were indeterminable from context, even using the corpus' expanded context feature. Another two percent of the uses in the sample were mislabeled as verbs (one of the concessions a corpus architect must make to automated grammatical tagging). Thus, there were eleven nominal uses of *carry*, as in (10), and one adjectival use, as in (11).

(10) Phillips, who came into the game averaging just 2.6 yards a carry, was impressive running inside the tackles against the Falcons.

(11) Cost is \$2,995, including suitcase and carrying bag.

Finally, just under twenty-nine percent of the sentences in the sample include the *carry*₂ sense. That these uses of *carry* are properly classified as *carry*₂ is sometimes obvious from the sentence itself, as in (12), and sometimes must be inferred from context, as in (13).

(12) On their heads or in their arms they carried their last worldly possessions.

(13) "Every child must carry the stone and throw it at the occupier," admonishes a 1989 leaflet.²³⁹

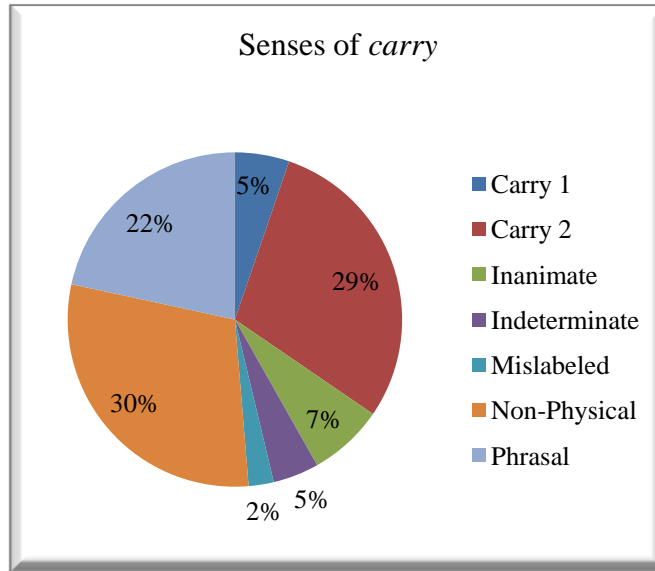
In contrast, only some five percent of the sentences in the sample included the *carry*₁ sense as in (14).

(14) Donkeys are disappearing in some parts of this country. . . . They don't use them as much to carry things anymore. They use tractors and big cars because they hold more.

The distribution of these senses of *carry* is illustrated in Figure 1 below.

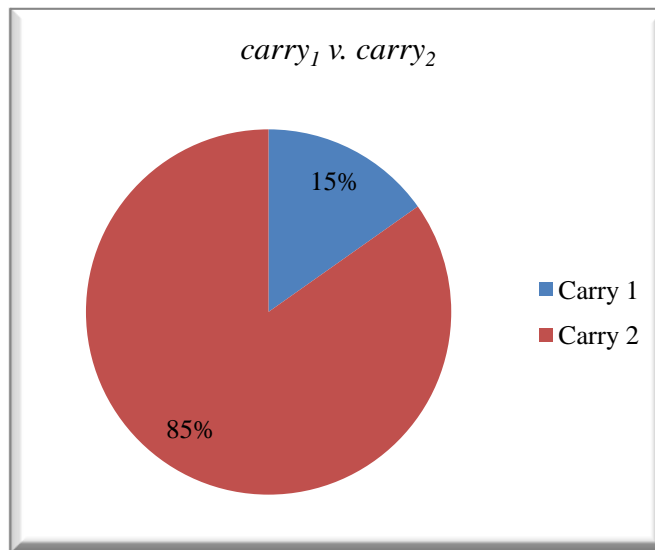
239. There is no reference to hands, arms, or even pockets in this sentence, but the context suggests that these stones were physically hefted and not conveyed in a vehicle.

Fig. 1



The ratio of *carry*₁ to *carry*₂ is illustrated in Figure 2.

Fig. 2



It is important to pause here and observe what these data do and do not tell us. My contention is not that because *carry*₂ is far more common than *carry*₁, § 924(c) ought to be interpreted with the *carry*₂ meaning. Such a reading would be arbitrary. There are undoubtedly circumstances in which Congress employs the less frequent of two senses of a word. However, these data do serve to undermine the Court's contention that *carry*₁ is the word's primary or ordinary meaning, while *carry*₂ is a "different, rather special way" to use *carry*. Instead, the circumstances are reversed. When it comes to the support or movement of a physical object, *carry*₂ is the primary or common use of *carry*, while *carry*₁ is the specialized sense of the term.

5. Collocation. As noted above, at issue in *Muscarello* "is not 'carries' at large but 'carries a firearm.'"²⁴⁰ While the data above undermine the majority's contention that *carry*₁ is the primary or ordinary meaning of *carry*, they do not give us a clear picture of how *carry* is used in the context of the statute—that is, when *carry* co-occurs with *firearm(s)*. One examination of *Muscarello* sought to expand the context in which *carry*'s use with *firearm* is analyzed by searching a database for a number of firearms, listed by name.²⁴¹ However, because both the parties and the Court framed the analysis in terms of how *carry* is ordinarily used, I think the better approach is to determine those synonyms of *firearm* with which *carry* ordinarily co-occurs. The corpus makes this possible through what is called a collocation display.

"Collocation is the tendency of words to be biased in the way they co-occur,"²⁴² that is, the tendency of certain words to be used in the same semantic environment as other words. A collocation program calculates collocation rates based on a *node word*, in our case, *carry*. The program proceeds by "count[ing] the instances of all words occurring within a particular span, for example, four words to the left of the node word and four words to the right."²⁴³

240. *Muscarello*, 524 U.S. at 143.

241. Note, *The Supreme Court 1997 Term, Leading Cases*, 112 HARV. L. REV. 355, 363 n.69 (1998) ("Search of Westlaw, USNEWS Database (Aug. 21, 1998) (search for records containing "(gun 'side arm' weapon ammunition ammo firearm shotgun rifle pistol revolver Uzi AK-47 handgun semiautomatic semi- automatic MAC-10 M-16 .38 .40 .44 .45 .357 9-mm 9-millimeter Beretta (Smith /3 Wesson) (Hechler /3 Koch) Magnum Glock Colt) /s (trunk) & DA(1993)).").

242. See HUNSTON, *supra* note 115, at 68.

243. *Id.* at 69.

Collocation statistics “can be helpful to the corpus user in summarizing some of the information to be found in concordance lines, thereby allowing more instances of a word to be considered.”²⁴⁴ Unlike the KWIC approach performed above, in which only 500 out of 82,687 uses of *carry* were examined, the collocation program will review every occurrence of the word in question and give a read-out of its collocation statistics. Put simply, the collocation data will show the words that are most commonly used with the word *carry*. Collocation is therefore “something of a short-cut to the information that could be obtained from concordance lines,”²⁴⁵ and may be used to confirm the data already extracted from the corpus.

In order to display the collocation information in the COCA, enter [carry].[v*] under the heading “Search String,” next to the search field labeled “Word(s),” then click on “Collocation.”²⁴⁶ Then click “Search.” This will display the first one hundred collocates of *carry* in order of their statistical frequency. The read-out will likewise display the total number of sentences in which each collocate appears.²⁴⁷ I have reprinted the first one hundred collocates, together with their frequency ranking and number of occurrences in Appendix IV.²⁴⁸

There are numerous purposes to which collocation displays can be put to use, each of them related to determining the semantic range of a word or words.²⁴⁹ For present purposes, the collocation display shows that *firearm* and four of its synonyms (together with their plurals) are listed among *carry*’s one hundred most common collocates.²⁵⁰ They are *gun*, *handgun*, *rifle*, and *pistol*. Consequently, in order to determine the ordinary use of *carry* when used with *firearm*, I have examined *carry* in the context of these related terms.

6. Carries a firearm. Having determined that the word *carry* commonly collocates with the words *firearm*, *gun*, *handgun*, *rifle*,

244. *Id.* at 76.

245. *Id.* at 77.

246. See Appendix III, Figure 9.

247. *Id.*

248. See Appendix IV.

249. See, e.g., BIBER ET AL., *supra* note 112, at 43–51.

250. I have excluded *weapon(s)* here even though the term appears among the top-hundred collocates of *carry*. Though all firearms are weapons, the reverse is not true and a preliminary examination of *weapon(s)* in the corpus suggested that many if not most of the weapons referenced were not firearms.

and *pistol* we can return to the KWIC display to search for sentences in which *carry* and *firearm* or its synonyms co-occur. This is done by entering [carry].[v*] into the “WORD(S)” search field and one of (15) through (19) in the “COLLOCATE” field.

- (15) [firearm].[n*]
- (16) [gun].[n*]
- (17) [handgun].[n*]
- (18) [rifle].[n*]
- (19) pistol²⁵¹

These searches will reveal every co-occurrence of any verb form of *carry* and any nominal form (singular or plural) of *firearm* and its synonyms. I have examined fifty concordance lines for each nominal search term, twenty-five for the singular form and twenty-five for the plural.

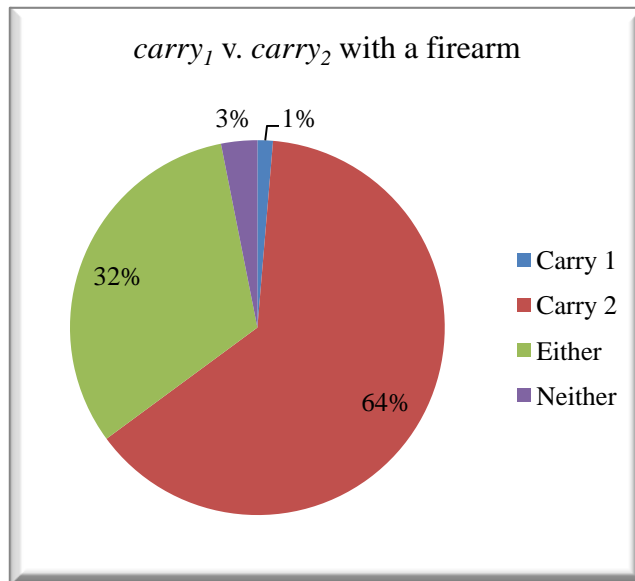
Virtually all of the concordance lines returned in the searches outlined above featured uses of *carry* related to the physical carrying of an object, though it could not always be determined whether *carry*₁ or *carry*₂ was intended. Consequently, I have classified the data returned from these searches under four broad headings: (1) *carry*₁, (2) *carry*₂, (3) *either carry*₁ or *carry*₂ or (4) *neither carry*₁ nor *carry*₂. These results are listed in Figure 3 below and represented graphically in Figure 4.

Fig. 3

	<i>carry</i> ₁	<i>carry</i> ₂	either	neither
<i>firearm(s)</i>	0	28	20	2
<i>gun(s)</i>	0	35	13	2
<i>handgun(s)</i>	1	24	24	1
<i>rifle(s)</i>	0	41	8	1
<i>pistol</i>	2	15	7	1
total	3	143	72	7

251. I have excluded the plural form *pistols* because only the singular form was among the top 100 collocates of *carry*.

Fig. 4



Like the statute at issue in *Muscarello*, a number of sentences in the sample above do not clearly implicate *carry*₁ or *carry*₂. These are instances in which *either carry*₁ or *carry*₂ is meant, as in (20) and (21).

(20) I would like you to carry your pistol on this job.

(21) It's safe to assume that the handguns they carry are duly registered.

In both cases, the firearms referenced could be carried either on one's person or in a car. However, even resolving all such instances in favor of *carry*₁ (which we have no reason to do), the data still demonstrate that in the overwhelming majority of cases, when *carry* is used in the context of a firearm, *carrying on one's person* is the intended meaning.

Therefore, if what is meant by ordinary meaning is that sense of a word that is most commonly employed, then the ordinary meaning of *carry* is undoubtedly *carry*₂. This conclusion does not necessarily resolve the questions at issue in *Muscarello*, nor does it tell us how long Frank Muscarello ought to spend in jail. As discussed above, Justice Breyer's real argument—his “purely legal” argument—is that

Congress was concerned about the violence that ensues when criminals bring guns to drug deals. The same Congress that wanted to stop criminals from bringing guns to drug deals on their person, would have likewise wanted to stop criminals from schlepping guns to drug deals in their cars. This is an entirely reasonable conclusion and is consistent with Justice Breyer's purposivist jurisprudence outlined above. An equally plausible argument made by Justice Ginsburg is that if Congress wishes to apply a five-year sentence enhancement to criminal defendant who drove to a drug deal with a gun locked in his glove box, Congress must say so explicitly. Still, as noted, the data at the very least undermine the majority's contention that *carry*₁ is the primary or ordinary meaning of *carry*. If the ultimate conclusion reached by the Court is based even in part on the erroneous belief that *carry*₁ is that term's ordinary meaning, that conclusion may likewise be called into question.

B. Some Common Criticisms of the Corpus Method

The corpus method is not without its critics and the application of the corpus approach to questions of ordinary meaning implicates several familiar criticisms. The first of these is that ultimately, no matter how large the data set or how well designed the corpus, the corpus will fail to present a full picture of the linguistic community analyzed. "The reason that no corpus will ever be big enough to embrace all the knowledge that a native speaker can access through introspection is that . . . the actual set of English sentences is an infinitely extendable one."²⁵² Thus, "even with the assistance of data from very large corpora, a native speaker can still have accurate intuitions that permit her to say, 'I know that such-and-such is a part, or a sentence, or an expression in English, even though it's not in the corpus.'"²⁵³

252. Charles Fillmore, Panel Discussion, Northwestern University/Washington University Law School, Law and Linguistics Seminar Conference, Evanston, Illinois, 73 WASH. U. L. Q. 791, 794-95 (1995) [hereinafter Panel Discussion]. This criticism of the corpus approach originated with Noam Chomsky, who noted: "Any natural corpus will be skewed. Some sentences won't occur because they are obvious, others because they are false, still others because they are impolite. The corpus, if natural, will be so wildly skewed that the description would be no more than a mere list." See TONY MCENERY & ANDREW WILSON, CORPUS LINGUISTICS: AN INTRODUCTION 10 (2001) (quoting Noam Chomsky, paper presented at the University of Texas, 3d Texas Conference on Problems of Linguistic Analysis in English published in STUDIES IN AMERICAN ENGLISH, at 159 (1962)).

253. See Panel Discussion *supra* note 252.

This criticism is less troublesome in the present circumstances. Although there may be uses of *carry* not captured by the data presented in the COCA, the two senses at issue in this case—*carry*₁ and *carry*₂—are well represented. Likewise, while cases may arise where Congress has regulated certain conduct using highly specialized language, such is not the case in *Muscarello*.

A second criticism of corpus data—particularly data derived from a contemporary corpus—is anachronism. After all, the earliest texts in the COCA date from 1990, but the statute at issue in *Muscarello* was passed in 1968. What if the use of *carry* has changed in the ensuing decades? Writing in 2003, Sydney Landau observed that “[m]ost existing lexicographic corpora are also limited in time, extending back perhaps to the 1970s, rarely much earlier. . . . There are historical corpora, but as of now they only scratch the surface of past use.”²⁵⁴ It has been suggested that “[u]ntil this deficit is supplied, the [corpus] approach . . . is not practicable for old statutes.”²⁵⁵

Of course, one obvious response to this criticism is to observe that the twenty-two year span between the Act and the earliest texts in the COCA is at the very least shorter than the distance between the passage of the Act and the Court’s citations to Melville, Defoe, or the translators of the King James Bible, not to mention to centuries old citations found in the OED. However, since this Comment was initially accepted for publication, the architect of the COCA has released the Corpus of Historical American English (or “COHA”). The COHA is now “the largest structured corpus of historical English (or of any language, for that matter).”²⁵⁶ It allows users to “search more than 400 million words of text of American English from 1810 to 2009” by decade.²⁵⁷ Because the Act was passed in 1968, a relevant inquiry into the use of *carry* and *firearm(s)* contemporaneous to the passage of the Act should focus on uses that date from the 1960s.

254. See LANDAU, *supra* note 51, at 321–22.

255. Rickie Sonpal, Note, *Old Dictionaries and New Textualists*, 71 FORDHAM L. REV. 2177, 2219 (2003).

256. Mark Davies, THE CORPUS OF HISTORICAL AMERICAN ENGLISH (COHA), available at <http://corpus.byu.edu/coha/> (last visited Jan. 27, 201). The COHA was developed by Professor Mark Davies with a grant from the National Endowment for the Humanities.

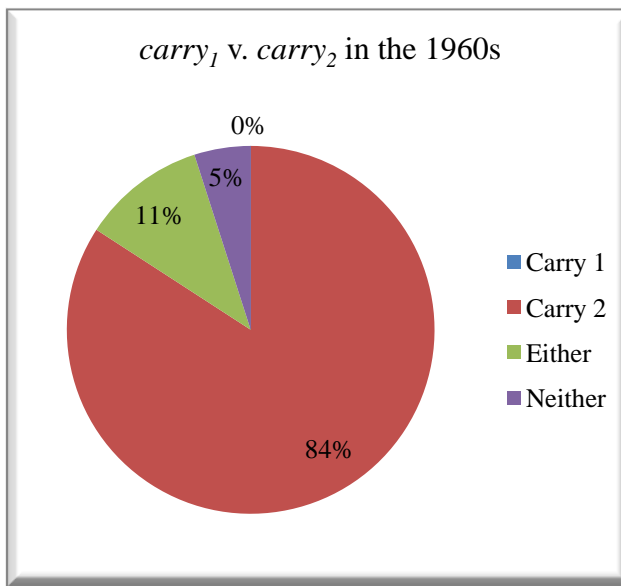
257. *Id.*

Although the COHA is approximately the same size as the COCA, the COHA is made up of texts spanning two centuries. This means that the sample sentences available for a given word in a given decade can be limited. Consequently, instead of limiting the analysis to fifty sentences for each term, I have examined every sentence in which *carry* collocates with any of the synonyms of *firearm(s)* listed above during the relevant timeframe. I have used the same classification system for these collocations of *carry* and *firearm(s)* with the addition of listing the total number of sentences in which the terms co-occur.

Fig. 5

	collocations	<i>carry</i> ₁	<i>carry</i> ₂	either	neither
<i>firearm(s)</i>	3	0	2	1	0
<i>gun(s)</i>	59	0	47	7	5
<i>rifle(s)</i>	27	0	26	1	0
<i>pistol</i>	12	0	10	2	0
total	101	0	85	11	5

Fig. 6



The results of COHA searches are more striking than those from contemporary usage. Among the sample sentences returned, not a single occurrence of *carry*₁ was found. Instead, nearly all of the references to *carry* in relation to *firearm(s)* and its synonyms were of the *carry*₂ variety. This result suggests that at the time of the Act's passage, *carry*₂ was by far the more common sense of the word.

A final criticism has to do with the notion of representativeness. Professor Solan has stated the problem as follows:

When the legal system decides to rely on the ordinary meaning of a word, it must also determine which interpretive community's understanding it wishes to adopt. This choice is made tacitly in legal analysis, but becomes overt when the analysis involves linguistic corpora because the software displays the issue on a screen in front of the researcher.²⁵⁸

This is both a strength and a weakness of the corpus-based approach to ordinary meaning. The corpus can only definitively say how a term is ordinarily used *within the corpus*. Given the infinite permutability of human language, the corpus can never capture every possible human utterance, even in a narrowly-defined speech community. The corpus architect must therefore justify her conclusion that the corpus is representative based on certain premises—none of which can be verified by an examination of the complete language use of the community as a whole. On the other hand, unlike the judge who often “tacitly” relies upon intuition or introspection to determine what uses of a word are ordinary, the premises of the corpus architect are “overt” as are the data used to determine ordinary usage.

Consider a scenario in which two judges on a panel are called upon to interpret a word in a statute. Judge A intuitively feels that the ordinary meaning of the word is X, and finds supporting definitions in several dictionaries. Judge B intuitively feels that the ordinary meaning is Y and likewise finds dictionary-based support. Assuming that there is no additional evidence as to which sense of the word should control, this dispute is intractable. After all, the judges' disagreement is about their differing intuitions.

Now suppose that Judge A bases her conclusion about the meaning of the word on a corpus-based analysis. If Judge B

258. Lawrence M. Solan, *The New Textualist's New Text*, 38 LOY. L.A. L. REV. 2027, 2059 (2005).

disagrees, she will have numerous “overt” bases to challenge Judge A’s conclusion. First, Judge B may attack the relevance of the corpus, noting that the data is not relevant to the speech community at issue. A general corpus like the COCA may not show the technical use of a term at issue in, say, a patent dispute. Second, Judge B may attack the construction of the corpus, noting that the methods employed by the corpus architect are not likely to represent the usage of the speech community the corpus claims to represent. Finally, Judge B may repeat the corpus-based research conducted by Judge A, using the same corpus data, and conclude that Judge A has reached an erroneous result.

The difference between these two approaches is one of the chief benefits of the corpus method. The first scenario is a metaphysical debate; the second is an empirical one.

VI. CONCLUSION

Justice Felix Frankfurter stated that “[a] problem in statutory construction can seriously bother courts only when there is a contest between probabilities of meaning.”²⁵⁹ If that is the case, then the process of interpretation may benefit from some method to quantify these comparative “probabilities.”

The corpus method is not a panacea. The use of corpus data will not do away with disagreements as to the meaning of statutory terms. Instead, the corpus method removes the determination of ordinary meaning from the black box of the judge’s mental impression and renders the discussion of ordinary meaning one of tangible and quantifiable reality.

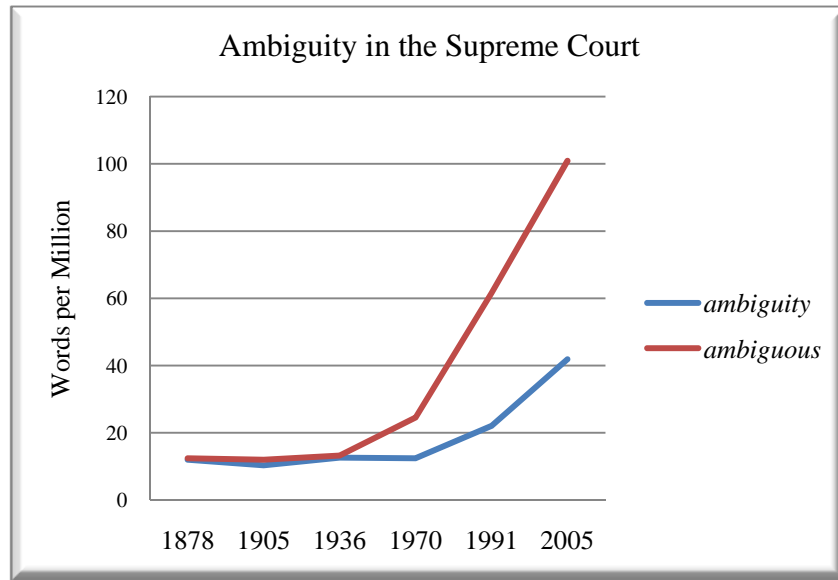
*Stephen C. Mouritsen**

259. Felix Frankfurter, *Some Reflections on the Reading of Statutes*, 47 COLUM. L. REV. 527, 527–28 (1947).

* Stephen C. Mouritsen received his JD from BYU in the spring of 2010 and an MA in linguistics with a corpus-linguistics emphasis from BYU in 2007. He currently serves as a law clerk for the Honorable Justice Thomas Rex Lee on the Utah Supreme Court. A version of this paper received the John S. Welch Award (First Prize) for Outstanding Legal Writing in 2010.

APPENDIX I: DATABASE OF DECISIONS FROM THE
UNITED STATES REPORTS

Fig. 1



To analyze the frequency of use of the terms ambiguity and ambiguous, I have constructed a database of every U.S. Supreme Court decision from the inception of the Court until 2005. In Figure 1 above, each of the six dates (x axis) represent the publication of the 100th volume of the United States Reports (“USR”) (i.e. volume 100 was published in 1878, volume 200 in 1905). Because these volumes will have a different number of pages and therefore a different number of words, I have regularized the data returned by calculating a words per million (“WPM”) score for each variable. The WPM is on the y axis. So, in the period between 1878 and 1905 (represented by 100 U.S. 1 through 199 U.S. 617) the word ambiguity occurred about 10.33 WPM, while ambiguous occurred some 11.97 WPM.

Fig. 2

<i>Vol. #s</i>	<i>Dates</i>	<i>Total Word Count</i>
1 to 99	1754 ¹ –1878	22,966,526
100 to 199	1879–1905	22,550,126
200 to 299	1906–1936	16,099,370
300 to 399	1937–1970	20,470,173
400 to 499	1971–1991	27,517,650
500 to 536	1991–2005	9,006,844

Figure 2 above represents the number of total words in each of six text files, and the USR volumes represented in each. The USR represents a corpus of 118,610,689 words. Because the volumes are arranged in chronological order, a comparison of the relative frequency of a given word in all six files should reveal significant information about whether the use of the terms examined has increased or decreased in Supreme Court opinions.

Fig. 3

	1878	1905	1936	1970	1991	2005
<i>ambiguity</i>	11.97	10.33	12.61	12.41	22.02	41.86
<i>ambiguous</i>	12.41	11.97	13.23	24.57	61.52	100.92

Figure 3 shows the average WPM for the terms ambiguity and ambiguous for the period represented on the x axis.

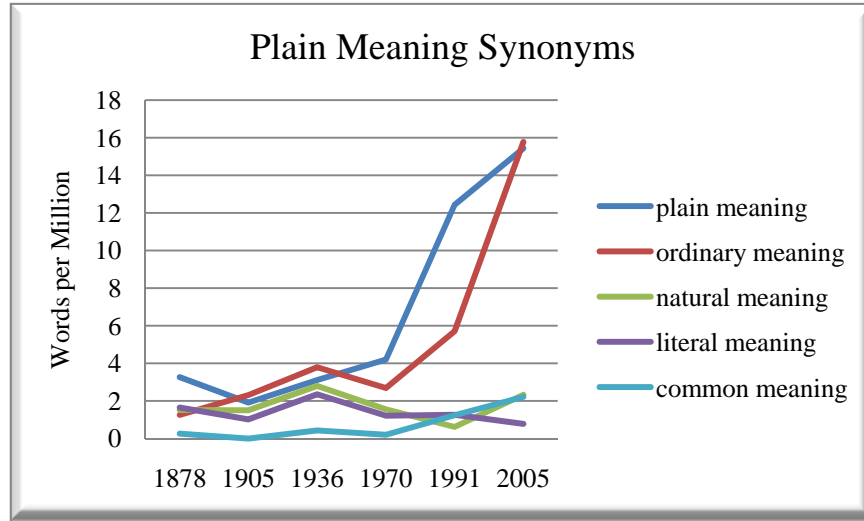
1. The Dallas reporter represents volumes one through four of the USR, and contains numerous Pennsylvania Supreme Court cases. As such, the earliest case reported in volume one of the USR predates the inception of the United States Supreme Court. *Anonymous*, 1 U.S. (1 Dall.) 1 (1754).

1915

The Dictionary Is Not a Fortress

APPENDIX II: PLAIN MEANING IN THE UNITED STATES REPORTS

Fig. 4



To analyze the frequency of use of the phrase plain meaning and its synonyms, I have used the same electronic database described in Appendix I.

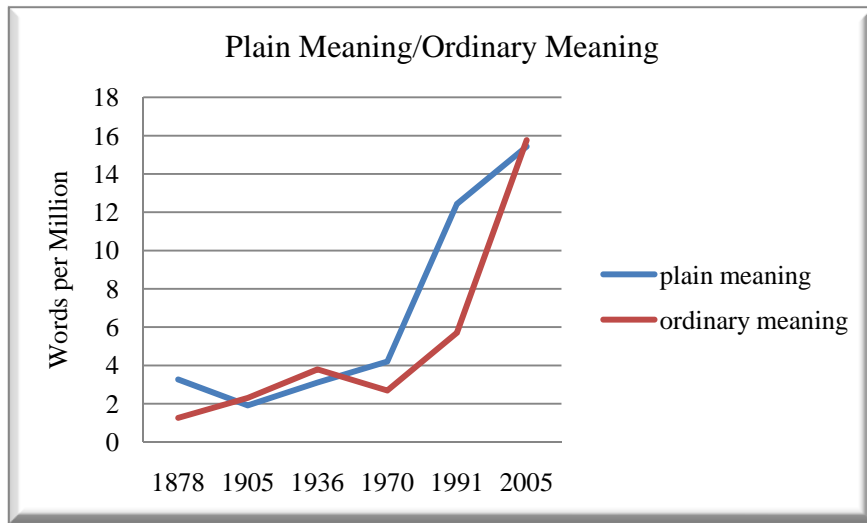
Fig. 5

	1878	1905	1936	1970	1991	2005
<i>plain meaning</i>	3.27	1.91	3.11	4.2	12.43	15.43
<i>ordinary meaning</i>	1.26	2.31	3.79	2.69	5.71	15.77
<i>natural meaning</i>	1.52	1.51	2.8	1.56	0.62	2.33
<i>literal meaning</i>	1.65	1.02	2.36	1.22	1.27	0.78
<i>common meaning</i>	0.26	0	0.43	0.2	1.24	2.22

Figure 5 shows the WPM for the term plain meaning and several of its synonyms for the periods represented on the x axis.

1973

Fig. 6



In Figure 6, I have isolated the frequency of plain meaning and ordinary meaning over the period represented on the x axis.

Fig. 7

	1878	1905	1936	1970	1991	2005
<i>plain meaning</i>	3.27	1.91	3.11	4.2	12.43	15.43
<i>ordinary meaning</i>	1.26	2.31	3.79	2.69	5.71	15.77

Figure 7 compares the average WPM for the terms plain meaning and ordinary meaning for the period represented on the x axis.

APPENDIX III: SCREEN SHOTS FROM CORPUS OF CONTEMPORARY AMERICAN ENGLISH

Fig. 8

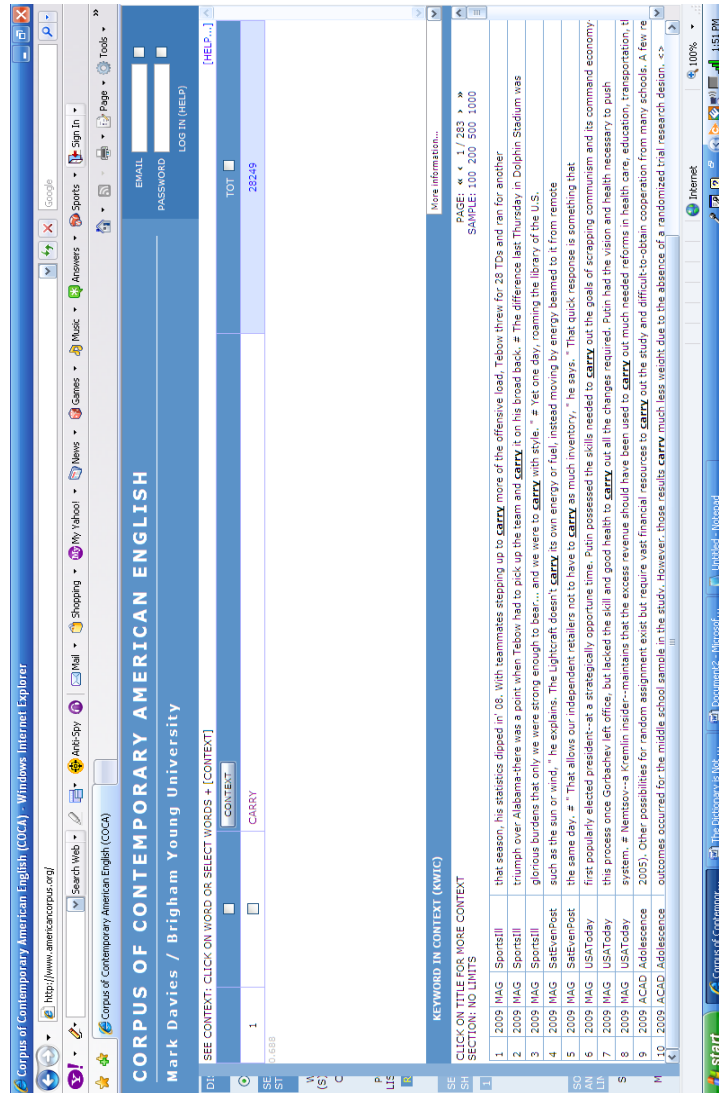
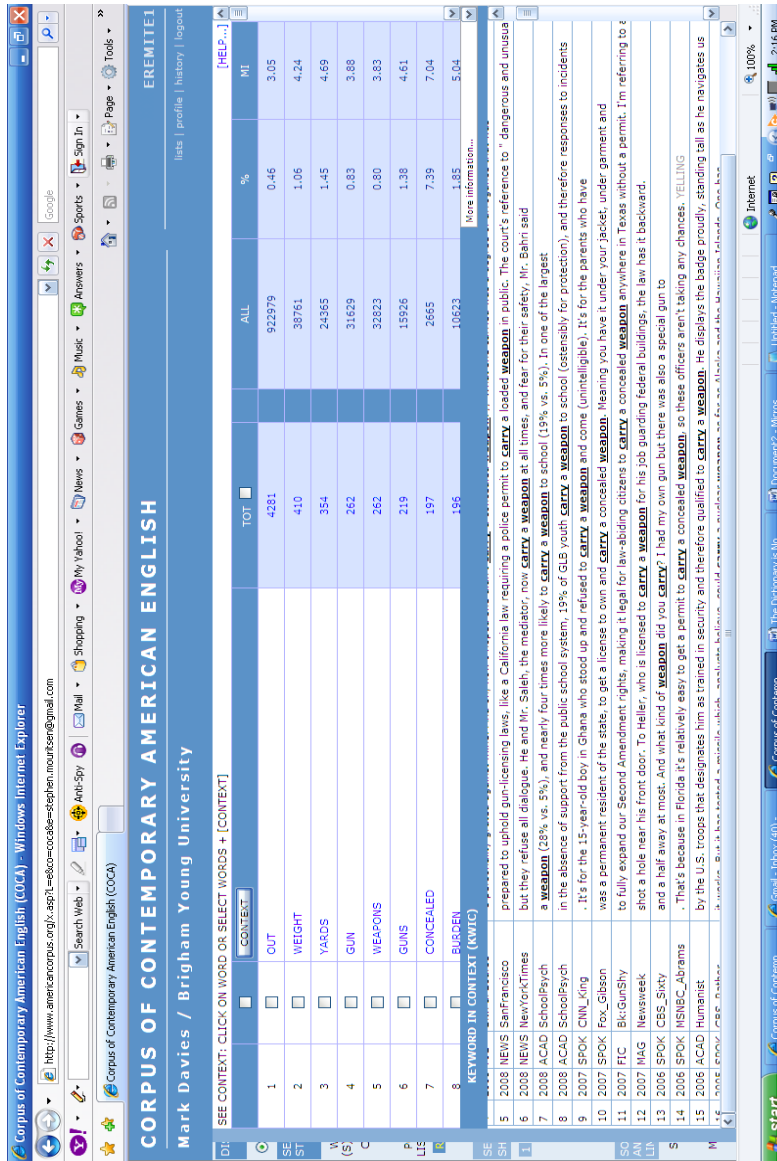


Fig. 9



APPENDIX IV: COLLOCATES OF *CARRY*

Fig. 10

Rank	Collocate	Co-occurrences
1	OUT	12674
2	YARDS	1303
3	WEIGHT	895
4	BAG	778
5	GUN	648
6	WEAPONS	565
7	HEAVY	538
8	BAGS	462
9	GUNS	443
10	LOAD	405
11	PASSENGERS	402
12	BURDEN	381
13	CONCEALED	296
14	WEAPON	283
15	TRAY	244
16	BAGGAGE	190
17	TORCH	187
18	BRIEFCASE	185
19	CARGO	184
20	SHIPS	180
21	GENE	173
22	VIRUS	169
23	SUITCASE	168
24	SENTENCE	168
25	GENES	159
26	LOADS	158
27	TASKS	155
28	MESSAGES	152
29	BACKPACK	148
30	RIFLE	148
31	TRUCKS	148
32	GEAR	148
33	MOMENTUM	139
34	FLAG	134
35	OXYGEN	131

Rank	Collocate	Co-occurrences
36	PURSE	127
37	BASKET	126
38	DUTIES	119
39	LUGGAGE	117
40	RESPONSIBILITIES	113
41	BANNER	111
42	GROCERIES	108
43	SACK	99
44	RIFLES	98
45	MISSILES	97
46	BREEZE	95
47	PISTOL	87
48	BUCKET	81
49	SCENT	80
50	STRETCHER	78
51	FREIGHT	74
52	BASKETS	74
53	TAG	72
54	KNIVES	68
55	TOUCHDOWNS	68
56	BANNERS	67
57	COFFIN	67
58	WALLET	67
59	NOTEBOOK	66
60	BUCKETS	65
61	STIGMA	65
62	UMBRELLA	65
63	BUNDLE	64
64	FLAGS	64
65	BURDENS	63
66	HANDGUNS	62
67	CURRENTS	62
68	PACKS	61
69	EXPLOSIVES	59
70	WARHEADS	58
71	SUITCASES	58
72	TRAYS	58
73	SCARS	56
74	SACKS	56

1915

The Dictionary Is Not a Fortress

Rank	Collocate	Co-occurrences
75	PENALTIES	55
76	ALOFT	53
77	MPH	53
78	SPACECRAFT	50
79	6-IRON	49
80	FIREARMS	49
81	CARDBOARD	49
82	PAYLOAD	48
83	CONNOTATIONS	48
84	FIREARM	48
85	PASSPORT	48
86	BOMBINGS	48
87	HANDGUN	47
88	AMBULANCE	47
89	BUNDLES	46
90	CASKET	46
91	DUFFEL	45
92	BACKPACKS	44
93	AMMUNITION	43
94	AVERAGING	42
95	POUCH	41
96	MOSQUITOES	41
97	CONVOY	40
98	BELONGINGS	40
99	CLOUT	40
100	EXECUTIONS	38

1979

