

2017

Datamining the Meaning(s) of Progress

Jake Linford

Follow this and additional works at: <https://ir.law.fsu.edu/articles>



Part of the [Constitutional Law Commons](#), and the [Intellectual Property Law Commons](#)

Recommended Citation

Jake Linford, *Datamining the Meaning(s) of Progress*, 2017 *B.Y.U. L. REV.* 1531 (2017),
Available at: <https://ir.law.fsu.edu/articles/600>

This Article is brought to you for free and open access by Scholarship Repository. It has been accepted for inclusion in Scholarly Publications by an authorized administrator of Scholarship Repository. For more information, please contact bkaplan@law.fsu.edu.

Datamining the Meaning(s) of *Progress*

Jake Linford*

CONTENTS

I. CORPUS LINGUISTICS AND TRIANGULATION	1531
II. ORIGINALISM AND SEMANTIC SHIFT	1534
III. COMPETING DEFINITIONS OF <i>PROGRESS</i>	1538
A. Immersion.....	1539
B. Constitutional History	1541
C. Meaning(s) of <i>Progress</i>	1544
1. Interpretation.....	1545
2. Construction.....	1547
D. Resolving Definitional Disputes	1549
IV. USING CORPUS LINGUISTICS TO DISCOVER MODULATION.....	1551
V. LIMITATIONS OF CORPUS LINGUISTICS.....	1553

I. CORPUS LINGUISTICS AND TRIANGULATION

Lawrence Solum proposes that corpus linguistics—quantitative analysis of usage data from a broad body or corpus of contemporary texts¹—might reveal which among possible meanings of an ambiguous term is most common and, thus, the most likely “ordinary” or

* Associate Professor, Florida State University College of Law. My thanks to Franita Tolson, Jay Kesten, and participants at the 2017 BYU Law Review Symposium, “Law & Corpus Linguistics,” and a faculty workshop at Florida State University College of Law. Thanks to Alan LaCerra and Lauren Pettine for excellent research assistance. Responsibility for unliquidated ambiguities, nonce formations, and failures of clarity is mine alone to bear.

1. Stephen C. Mouritsen, *The Dictionary Is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning*, 2010 BYU L. REV. 1915, 1938.

“conventional semantic meaning” of a term.² Such evidence of contemporary usage might fruitfully contribute to a search for original meaning of constitutional language.³ Professor Solum argues that the search for original meaning should be undertaken by triangulating corpus evidence with two other methods for discovering original meaning: immersive reading of texts contemporary with the drafting of the Constitution; and careful examination of the drafting, ratification, and implementation history of the Constitution.⁴ Consilience, or agreement, between these methods would strongly indicate a dominant meaning of a given term.⁵ But proper triangulation could be a massive undertaking. As Professor Solum notes, locating the original meaning of constitutional language may well amount to a work of many lifetimes, requiring “a division of intellectual labor.”⁶ Nonetheless, digitization and computational analysis of a large number of texts created prior to, and contemporaneous with, the drafting of the Constitution might better provide information about ordinary meaning for which courts currently consult dictionaries.⁷

This response agrees in large part with Professor Solum’s prescription and focuses on how to best use corpus lexicography to confirm or refute other evidence of original public meaning.⁸ The tools to build and digitally examine corpora are relatively new, but corpus lexicography can fruitfully build on both analysis grounded in immersion and constitutional history analysis carried out in earlier scholarship.⁹ Herein I focus on how corpus lexicography might build

2. Lawrence B. Solum, *Triangulating Public Meaning: Corpus Linguistics, Immersion, and the Constitutional Record*, 2017 BYU L. REV. 1621, 1631 [hereinafter Solum, *Triangulating*].

3. *Id.* at 1623–25.

4. *Id.*

5. *Id.* at 1676–77.

6. *Id.* at 1680.

7. *See infra* notes 25–26 and accompanying text.

8. Linguistics is “[t]he scientific study of language and its structure,” *Linguistics*, OXFORD ENGLISH DICTIONARY (2017), <http://www.oed.com>, while lexicography is the art of compiling a dictionary, *Lexicography*, OXFORD ENGLISH DICTIONARY (2017), <http://www.oed.com>. Both disciplines entail a search for ordinary and unusual meanings of words as lexical units. In this response, I adopt the term corpus lexicography for the use of data analysis to uncover usual meanings, consistent with Professor Solum’s use in his essay.

9. *See infra* Sections III.III.A–III.B.

on prior scholarly work which analyzes the language of Article I, Section 8, Clause 8 of the Constitution (“the Copyright and Patent Clause”). That provision empowers Congress to secure for limited times to authors and inventors exclusive rights in copyrighted expression and patented inventions.¹⁰ Intellectual property scholars mining contemporaneous texts and constitutional history have done important work defining key terms, but they reach conflicting conclusions.¹¹ Corpus lexicography may be well suited to support or refute “traditional” linguistic analysis because it quantifies analysis that has historically been left to the interpretive faculties of the reader, scholar, or jurist.¹²

This response will also alert the reader to potential pitfalls to avoid when undertaking corpus lexicography.¹³ For instance, Professor Solum notes that corpus lexicography might often fail to recognize an attempt to use an existing word in a new way to create meaning. By definition, a use that diverges from standard meanings will be difficult to track using any measure designed to detect standard meanings. Such a modulation might not be picked up by texts contemporaneous with the drafting of the Constitution precisely because the use is new. This is problematic if the drafters intended to use an existing word in a new way. Context might reveal this new intended meaning, but Professor Solum expresses reasonable concern that corpus analysis will miss it. This response highlights one potential approach to corpus construction that might ameliorate this limitation by treating the drafting and ratification of the Constitution as an inflection point from which we might measure semantic shift—the creation of new meaning.¹⁴

10. U.S. CONST. art. I, § 8, cl. 8 (“The Congress shall have Power to . . . promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries . . .”).

11. See *infra* Sections III.III.C–III.D.

12. Cf. Beth Levin, Grace Song & B. T. S. Atkins, *Making Sense of Corpus Data: A Case Study of Verbs of Sound*, 2 INT’L. J. CORPUS LING. 23, 29–30 (1997).

13. See *infra* Part V.

14. See *infra* Part IV.

II. ORIGINALISM AND SEMANTIC SHIFT

Originalism is, with some variance between subdisciplines, a search for constitutional meaning contemporaneous with the founding era. Professor Solum in particular begins with the premise that proper originalist analysis of constitutional text should strive to define the “original public meaning” of constitutional terms. In other words, the text of the Constitution should be interpreted consistent with how the text would have been understood by the public at the time it was drafted.¹⁵ Originalist interpretation has certainly captured the imagination of scholars and judges. The originalist approach is nonetheless susceptible to the accusation that it lacks discipline. To a skeptic, an originalist inquiry might resemble Justice Scalia’s famous criticism of reliance on legislative history:¹⁶ Originalist scholars might naturally look out over an ambiguous historical record and focus on evidence consistent with their expectations.¹⁷

The inquiry into original meaning is often conducted with dictionaries. Dictionaries may have been the best available interpretive tool in earlier eras, but they are not necessarily a good tool. Many dictionaries are organized with goals distinct from or even antithetical to a search for original public meaning.¹⁸ None seem systematic in

15. See Solum, *Triangulating*, *supra* note 2, at 1629, 1637; see also Randy E. Barnett & Evan Bernick, *The Letter and the Spirit: A Unified Theory of Originalism* 3–5 (Oct. 5, 2017) (unpublished article), <https://papers.ssrn.com/abstract=3049056>.

16. Justice Scalia commonly used an analogy first announced by Judge Leventhal. See, e.g., Nicholas S. Zeppos, *Legislative History and the Interpretation of Statutes: Toward a Fact-Finding Model of Statutory Interpretation*, 76 VA. L. REV. 1295, 1310 n.58 (1990).

17. William N. Eskridge, Jr., *Should the Supreme Court Read The Federalist but Not Statutory Legislative History?*, 66 GEO. WASH. L. REV. 1301, 1303–07, 1312–14 (1998); Bradley C. Karkkainen, *“Plain Meaning”: Justice Scalia’s Jurisprudence of Strict Statutory Construction*, 17 HARV. J.L. & PUB. POL’Y 401, 457 (1994) (“The main danger in judicial interpretation of the Constitution—or, for that matter, in judicial interpretation of any law—is that the judges will mistake their own predilections for the law.” (quoting Antonin Scalia, *Originalism: The Lesser Evil*, 57 U. CIN. L. REV. 849, 856 (1989))); Vasan Kesavan & Michael Stokes Paulsen, *The Interpretive Force of the Constitution’s Secret Drafting History*, 91 GEO. L.J. 1113, 1138 (2003) (noting that a search for “the actual understanding of the Ratifiers” does not avoid “all the problems of use of legislative history and ascertaining collective intention”); Cass R. Sunstein, *Originalism* 3–4 (Oct. 17, 2017) (online working paper), <https://papers.ssrn.com/abstract=3055093>.

18. See Clark D. Cunningham et al., *Plain Meaning and Hard Cases*, 103 YALE L.J. 1561, 1614–16 (1994); Mouritsen, *supra* note 1, at 1935–37, 1945–46 (arguing that at best, using traditional techniques, a dictionary can provide “the most frequently encountered meaning,” not “the most frequently occurring meaning”); Malla Pollack, *What is Congress*

offering a most common meaning. Indeed, some dictionaries suggest meanings that are archaic or effectively abandoned. One cannot safely assume lexicographers compiled a given dictionary with the goal of revealing original public meaning, and dictionaries may thus be a tool that is not well-suited to the originalism often attempted with their assistance.¹⁹ Corpus lexicography thus holds out the promise that data of usage derived from the right corpora might “make originalism’s methodology more rigorous.”²⁰

To understand the importance of the originalism movement in constitutional interpretation, one must understand how the meaning of a word can change over time. As Professor Solum notes, “[w]ords and phrases have conventional semantic meanings, determined by patterns of usage.”²¹ Not every attempted linguistic change catches on. Indeed, one can trace obscure blips or “nonce formations” that are proposed and never adopted.²² Other usages can be adopted by local communities in certain contexts.²³ Some of these occasional meanings become “usual” or ordinary, and the word thus adopts that

Supposed to Promote?: Defining “Progress” in Article I, Section 8, Clause 8 of the United States Constitution, or Introducing the Progress Clause, 80 NEB. L. REV. 754, 789 (2001) [hereinafter Pollack, *What is Congress Supposed to Promote?*]; Solum, *Triangulating*, *supra* note 2, at 1631 (“Words and phrases have conventional semantic meanings, determined by patterns of usage. Lexicographers report these conventional semantic meanings in dictionary definitions, but such definitions are secondary evidence of the patterns of usage.”).

19. Solum, *Triangulating*, *supra* note 2, at 1641. On occasion, the Court clearly recognizes that a statutory provision has incorporated a meaning that is not the most common meaning, but is revealed through context. *See, e.g.*, *Gen. Dynamics Land Sys., Inc. v. Cline*, 540 U.S. 581, 594–98 (2004) (rejecting the argument that “age” as used in the ADEA must have the same meaning in every instance, and using context to determine which meaning was probably intended).

20. Lee J. Strang, *The Original Meaning of ‘Religion’ in the First Amendment: A Test Case for Originalism’s Utilization of Corpus Linguistics* 1, 1 (Sept. 1, 2017) (unpublished paper); accord Caleb Nelson, *Originalism and Interpretive Conventions*, 70 U. CHI. L. REV. 519, 535 (2003) (“Madison believed in the inevitability of linguistic change, [but] did not think that this change should affect the meaning of laws or constitutions.”).

21. Solum, *Triangulating*, *supra* note 2, at 1631.

22. *See* Jake Linford, *A Linguistic Justification for Protecting “Generic” Trademarks*, 17 YALE J.L. & TECH. 110, 131 (2015) [hereinafter Linford, *“Generic” Trademarks*]. One recent example of an attempted modulation on film was featured in the movie *Mean Girls*. The effort of one character to use “fetch” as a synonym for “cool” or “awesome” is blocked by a friend with a frosty “Stop trying to make ‘fetch’ happen. It’s not going to happen.” *See* Jake Linford, *The False Dichotomy Between Suggestive and Descriptive Trademarks*, 76 OHIO ST. L.J. 1367, 1411 (2015) [hereinafter Linford, *False Dichotomy*] (quoting *MEAN GIRLS* (Paramount Pictures 2004)).

23. *See supra* note 22.

meaning irrespective of context.²⁴ When you ask someone what a given word means, the answer is likely one of a handful of publicly accessible, relatively well-established and thus acontextual meanings.

As English speakers, it is easy to assume that we use words much the same way those words were used centuries ago, but as many high school students learn when they first encounter Shakespeare, that assumption is often incorrect. Likewise, an appeal to modern meaning might lead to fundamental error in constitutional interpretation.²⁵ A contemporary reader of English is likely to presume key terms in the Constitution have meanings consistent with one or more common contemporary meanings. In reality, words often experience semantic shift over time.²⁶ Meanings broaden or narrow; offensive meanings are ameliorated and inoffensive meanings become pejorative; and words add new meanings either by adding close meanings, or by shifting away from traditional meanings when similar foreign words conflict.²⁷ Words used in the Constitution often undergo semantic shift between the founding era and the modern era. Failure to account for the shift can lead to interpretive error, but corpus lexicography can reveal those differences.

At its core, corpus lexicography is interpretation through data analysis. The term *corpus* refers to a dataset of words, culled from numerous texts. There are two standard types of analysis. First, statistical analysis can show the key word in context (KWIC). Each use of the word will appear in a sentence or other string of characters, and a reader can consider each use as it appears in the original text. If the number of occurrences is unwieldy, a randomly determined sample of those instances could be reviewed instead.

Corpus analysis also can provide evidence of collocation. Collocation is “the habitual juxtaposition or association” of a given word with other words at a frequency higher than chance.²⁸ A search

24. For a more detailed discussion of modulation and semantic shift, see Linford, “Generic” Trademarks, *supra* note 22, at 130–45; Linford, *False Dichotomy*, *supra* note 22, at 1391–98.

25. See, e.g., Solum, *Triangulating*, *supra* note 2, at 1645–46.

26. *Id.* at 1639–41.

27. See Linford, “Generic” Trademarks, *supra* note 22, at 132–33.

28. *Collocation*, OXFORD ENGLISH DICTIONARY, <https://en.oxforddictionaries.com/definition/collocation> (last visited Jan. 13, 2018).

for collocates will display how often a given word appears within a certain distance of the target word.²⁹ Collocation thus provides information about words that are often fellow travelers. Evidence of collocation can highlight the context of usage. Understanding which words frequently appear near the target word can thus indicate ordinary usage of the target word.³⁰ Collocation evidence may also allow a researcher to track semantic shift over time.³¹

Professor Solum provides an example of semantic shift in the Copyright and Patent Clause. The clause empowers Congress “[t]o promote the Progress of Science and useful Arts.”³² Our modern understanding of the term *science* would lead the reader to conclude that Congress is empowered to promote progress in “hard sciences”—chemistry, physics, and the like—through the grant of copyright and patent protection.³³ This presumption is incorrect, and corpus lexicography can reveal the difference between founding era and modern usage. Professor Solum’s article summarizes evidence from one corpus, the Corpus of Historical American English (COHA), which includes texts as old as 1810. Professor Solum’s search reveals that most of the collocates for *science of* are not references to STEM sciences but to a broader collection of knowledge bases, including “*government, politics, art, law, religion, and theology*.”³⁴ Indeed, consistent with established precedent, copyright protection does not extend to scientific invention but instead protects creative expression—the authors’ “writings.”³⁵ Our modern understanding of *science* does not fit how the term is understood as part of the Constitution’s grant of power to Congress. This change between

29. Mouritsen, *supra* note 1, at 1958, 1963.

30. Collocation evidence can be adjusted to control for the frequency of each word’s appearance in the corpus generally, to avoid overreporting the frequency of collocation with common words or underreporting collocation with rare ones.

31. See *infra* notes 109–13 and accompanying text.

32. U.S. CONST. art. I, § 8, cl. 8.

33. Solum, *Triangulating*, *supra* note 2, at 1640 (“The contemporary meaning of the word is usually limited to the so-called hard sciences such as physics, chemistry, and biology . . .”).

34. *Id.* at 1646.

35. Jake Linford, *The Institutional Progress Clause*, 16 VAND. J. ENT. & TECH. L. 533, 559 (2014) (citing *Golan v. Holder*, 132 S. Ct. 873, 887–88 (2012)).

founding era and contemporary definitions of science is an example of semantic shift.³⁶

Other terms in the Copyright and Patent Clause present similar interpretive challenges. For example, several scholars have engaged in a close reading of texts contemporaneous with the Constitution in an attempt to define a meaning for *progress* as it is used in the Copyright and Patent Clause.³⁷ Other articles examine the limited history of the ratification debates over the clause to discover similar information.³⁸ The next part discusses this evidence and highlights a few conflicts of interpretation and construction that this research has revealed.

III. COMPETING DEFINITIONS OF *PROGRESS*

Prior attempts to pin down the meaning of key terms in the Copyright and Patent Clause have frequently relied on immersion or an examination of constitutional history. But scholars pursuing these methods of inquiry have reached conflicting conclusions about the proper interpretation of key terms and the proper legal construction of those terms. For instance, scholars reach different conclusions about the proper meaning of *progress* as used in the Copyright and Patent Clause.³⁹

These conflicting interpretations might be resolved if corpus linguistic evidence of public meaning favors one meaning. Thus, assuming that original public meaning should set the bounds of judicial construction, properly defining *progress* is therefore crucially important.⁴⁰ One fundamental debate concerns whether the stated goal of “promot[ing] the Progress of Science and useful Arts”⁴¹ grants broad power to Congress to implement copyright and patent protection as it sees fit, or limits Congress to enacting grants

36. Other scholars refer to these phenomena as semantic change, semantic progression, or semantic drift. See, e.g., ELIZABETH CLOSS TRAUOGOTT & RICHARD B. DASHER, REGULARITY IN SEMANTIC CHANGE 1 (2002) (defining semantic change as a shift “from one linguistically coded meaning to another”).

37. See *infra* Section III.III.A.

38. See *infra* Section III.III.B.

39. See *infra* Section III.C.

40. Margaret Chon, *Postmodern “Progress”: Reconsidering the Copyright and Patent Power*, 43 DEPAUL L. REV. 97, 145 (1993) (“[W]e cannot decide how to ‘promote the Progress of Science and useful Arts’ before we have debated the terms of progress.”).

41. U.S. CONST. art. I, § 8, cl. 8.

consistent with a narrow definition of *progress*.⁴² One cannot answer that question without defining *progress*. But once we understand the nature of the progress that Congress is empowered to promote, we can properly evaluate claims that the copyright and patent regimes are overbroad or underprotective in whole or in part.⁴³

A. Immersion

Many scholars turn to historical accounts to trace common meanings of *progress*. The majority of these historical accounts indicate that *progress* was a relatively modern idea in the “seventeenth and eighteenth centuries,”⁴⁴ although other scholars suggest that a fairly contemporary concept of *progress* was captured or expressed by classical authors.⁴⁵ Some legal scholars investigate the work of one or two authors, and others examine multiple volumes.⁴⁶ This may be

42. Compare *Graham v. John Deere Co. of Kan. City*, 383 U.S. 1, 5 (1966) (suggesting that Congress’s patent power is limited to promoting progress in the useful arts), with *Golan v. Holder*, 565 U.S. 302, 326 (2012) (reading broadly Congress’s power to promote progress by extending copyright protection to works that have lost copyright protection for deficiencies like failing to observe certain formalities).

43. See, e.g., *In re Bilski*, 545 F.3d 943, 1001–02 (Fed. Cir. 2008) (Mayer, J., dissenting), *aff’d but criticized sub nom.* *Bilski v. Kappos*, 561 U.S. 593 (2010) (“[P]atentable processes must ‘be in the technological arts so as to be in consonance with the Constitutional purpose to promote the progress of ‘useful arts.’ Business method patents do not promote the ‘useful arts’ because they are not directed to any technological or scientific innovation.” (emphasis omitted) (citation omitted)).

44. Michael D. Birnhack, *The Idea of Progress in Copyright Law*, 1 BUFF. INTELL. PROP. L.J. 3, 8 (2001) (first citing JOHN BAGNELL BURY, *THE IDEA OF PROGRESS: AN INQUIRY INTO ITS ORIGIN AND GROWTH* 7 (1932); and then citing DANIEL J. BOORSTIN, *THE SEEKERS: THE STORY OF MAN’S CONTINUING QUEST TO UNDERSTAND HIS WORLD* 184 (1998)); see also Chon, *supra* note 40, at 116 (first citing HENRY F. MAY, *THE ENLIGHTENMENT IN AMERICA* (1976); and then citing RAYMOND WILLIAMS, *KEYWORDS: A VOCABULARY OF CULTURE AND SOCIETY* 205–06 (1976)).

45. Birnhack, *supra* note 44, at 7 (citing ROBERT NISBET, *HISTORY OF THE IDEA OF PROGRESS*, at xii (1994)).

46. Pollack, *What is Congress Supposed to Promote?*, *supra* note 18, at 803–09 (first citing A. R. J. Turgot, *Turgot, A Philosophical Review of the Successive Advances of the Human Mind*, in *TURGOT ON PROGRESS, SOCIOLOGY AND ECONOMICS* 41, 41–44 (Ronald L. Meek ed. & trans., 1973); then citing ANTOINE-NICOLAS DE CONDORCET, *SKETCH FOR A HISTORICAL PICTURE OF THE PROGRESS OF THE HUMAN MIND* 33, 38, 42, 73–76, 92–93, 99–106, 117–20, 136–40, 164, 171, 186–88 (June Barraclough trans., Noonday Press 1955) (1795); then citing THE MARQUIS OF CONDORCET, *THE FRENCH ACADEMY OF SCIENCES, THE LIFE OF M. TURGOT, COMPTROLLER GENERAL OF THE FINANCES OF FRANCE IN THE YEARS 1774, 1775, and 1776* (1787); then citing Michael Kiernan, *Preface to FRANCIS BACON, THE ADVANCEMENT OF KNOWLEDGE*, at vii, vii (Michael Kiernan ed., Clarendon Press 2000); then citing FRANCIS

second-order immersion, relying on historians to define *progress*. Legal scholars can then use the interpretation provided to motivate a proper construction. Of course, a historical exegesis of the usage of *progress* may or may not precisely dovetail with the common meaning of the term.

Two scholarly attempts at immersion look somewhat like proto-corpus lexicography. Malla Pollack analyzed every usage of the word *progress* in the *Pennsylvania Gazette*,⁴⁷ a newspaper often touted as the paper of record for the founding era.⁴⁸ Professor Pollack identified five definitions of *progress* as used in the *Pennsylvania Gazette*, as well as quotations of the Copyright and Patent Clause, ordered in the footnote below from most to least frequently used.⁴⁹ She concluded

BACON, THE ADVANCEMENT OF LEARNING 9 (Oxford Univ. Press ed., 2000) (1605); then citing MERIC CASAUBON, A TREATISE CONCERNING ENTHUSIASM (1655); then citing Lawrence E. Klein, *Introduction to SHAFTESBURY, CHARACTERISTICS OF MEN, MANNERS, OPINIONS, TIMES*, at vii, passim (Lawrence E. Klein ed., Cambridge Univ. Press 1999); then citing 2 BERNARD MANDEVILLE, THE FABLE OF THE BEES: OR, PRIVATE VICES, PUBLIC BENEFITS 43 (Clarendon Press 1924) (1714); then citing F. B. Kaye, *Introduction to BERNARD MANDEVILLE, THE FABLE OF THE BEES: OR PRIVATE VICES, PUBLIC BENEFITS*, at xvii, xxxix, lx–lxi (Liberty Fund 1988) (1732); then citing BERNARD MANDEVILLE, A LETTER TO DION 40 (Bonamy Dobree ed., Univ. Press of Liverpool 1954) (1732); then citing David Berman, *Introduction to GEORGE BERKELEY, ALCIPHON OR THE MINUTE PHILOSOPHER IN FOCUS 1*, 1–2 (David Berman ed., 1993) (1732); then citing GEORGE BERKELEY, ALCIPHON, OR THE MINUTE PHILOSOPHER IN FOCUS 6, 12, 24, 29, 52, 158 (David Berman ed., 1993) (1732); then citing FRANCIS HUTCHINSON, REFLECTIONS UPON LAUGHTER AND REMARKS UPON THE FABLE OF THE BEES (Garland Publ’g 1971) (1750); then citing FRANCIS HUTCHINSON, AN INQUIRY INTO THE ORIGINAL OF OUR IDEAS OF BEAUTY AND VIRTUE (Garland Publ’g 1971) (2d ed. 1726); then citing FRANCIS HUTCHESON, ON HUMAN NATURE (Thomas Mautner ed., Cambridge Univ. Press 1993) (containing both “Reflections on Our Common Systems of Morality,” and “On the Social Nature of Man”); then citing ADAM SMITH, THE THEORY OF MORAL SENTIMENTS 236 (Liberty Classics 1982) (1759); and then citing ADAM SMITH, THE WEALTH OF NATIONS (Modern Library 2000) (1776).

47. Pollack, *What is Congress Supposed to Promote?*, *supra* note 18, at 798 (“I consider the word usage of the *Pennsylvania Gazette* the best currently available evidence of what 1789 American residents would have understood from the word ‘progress’ in the Progress Clause.”); *cf.* Randy E. Barnett, *New Evidence of the Original Meaning of the Commerce Clause*, 55 ARK. L. REV. 847, 856–66 (2003) (using the *Pennsylvania Gazette* to explicate meaning of “commerce” and “to regulate”).

48. Charles E. Clark & Charles Wetherell, *The Measure of Maturity: The Pennsylvania Gazette, 1728-1765*, 46 WM. & MARY Q. 279, 280 (1989) (arguing that the *Pennsylvania Gazette* “[i]n its essential character, although not in its unusual longevity, . . . was representative of the great majority of the newspapers of the provincial period”).

49. Pollack, *What is Congress Supposed to Promote?*, *supra* note 18, at 798:

in light of this evidence that *progress* was most commonly used to indicate some type of physical movement or spread of the thing progressing, as a fire might progress.⁵⁰

Orrin Hatch and Thomas Lee conducted a similar computer-assisted study of the use of *progress* in the Federalist Papers.⁵¹ That analysis also supports the conclusion that the dominant common meaning of *progress* was to spread.⁵² A search of a broader corpus of texts contemporaneous with the drafting and ratification of the Constitution, like the forthcoming Corpus of the Founding Era American English (COFEA), could support or challenge that conclusion.⁵³

B. Constitutional History

Some provisions of the Constitution were subject to exhaustive debate during the Constitutional Convention. Others, like the Copyright and Patent Clause, were agreed upon with very little controversy.⁵⁴ More problematically, for the purpose of discovering original meaning, there was little effort in the early days post-

Definition	Occurrences	Percentage
physical movement without implication of qualitative improvement, e. g. progress of a fire or a traveler	213	0.370
change or action towards a pre-set goal, e. g. progress towards finishing a book	125	0.217
qualitative improvement	124	0.216
movement through time, i.e. a chronologically arranged account without implication of qualitative improvement	80	0.139
numerical increase without implication of qualitative improvement	21	0.037
a quote of the Constitutional phrase	6	0.010

50. *Id.* at 809.

51. Orrin G. Hatch & Thomas R. Lee, “To Promote the Progress of Science”: The Copyright Clause and Congress’s Power to Extend Copyrights, 16 HARV. J.L. & TECH. 1 (2002).

52. *Id.* at 8–10, 10 n.42.

53. LAW & CORPUS LINGUISTICS CONFERENCE, <http://lawcorpus.byu.edu/> (last visited Jan. 15, 2018).

54. Edward C. Walterscheid, *The Preambular Argument: The Dubious Premise of Eldred v. Ashcroft*, 44 IDEA—J.L. & TECH. 331, 351 n.105 (2004) [hereinafter Walterscheid, *Preambular Argument*] (citing 2 RECORDS OF THE FEDERAL CONVENTION 508–10 (Max Farrand ed., 1911)) (stating that James Madison’s notes from the Constitutional Convention “indicate only that the Clause was approved without debate or dissent”).

ratification to define key terms or set the bounds of congressional authority.⁵⁵ Indeed, the Supreme Court has never defined *progress*.⁵⁶

Scholars attempting to flesh out the meaning of *progress* have turned to drafts of proposed language that were not accepted to provide some insight. For example, Dotan Oliar examined draft language proposed by both James Madison and Charles Pinckney, and suggests that differences between those proposed clauses and the Clause as adopted assist in interpreting *progress*. For example, the language of draft clauses proposed by both Madison and Pinckney would have extended plenary power to Congress to grant patent and copyright protection. Neither proposal included language like the Progress phrase that was eventually adopted.⁵⁷ The decision to reject these broader grants of power and adopt instead the Clause as enacted tends, as Professor Oliar reads it, “to prove that the Progress [language] was added as a limitation.”⁵⁸

Due to the paucity of evidence from the drafting and ratification debates, scholars often rely on post-ratification activity by Congress and the courts to echolocate meanings of constitutional terms.⁵⁹ To wit, scholars have found meaningful congressional inactivity in defining the contours of *progress*. For instance, Congress chose not to act in response to President Washington’s call to found a national university,⁶⁰ and likewise failed to pass legislation that would have provided direct subsidies for scientific research and exploration.⁶¹

55. Edward C. Walterscheid, *Divergent Evolution of the Patent Power and the Copyright Power*, 9 MARQ. INTELL. PROP. L. REV. 307, 317 (2005).

56. Christina Mulligan et al., *Founding-Era Translations of the U.S. Constitution*, 31 CONST. COMMENT. 1, 27 (2016).

57. Dotan Oliar, *Making Sense of the Intellectual Property Clause: Promotion of Progress as a Limitation on Congress’s Intellectual Property Power*, 94 GEO. L.J. 1771, 1776 (2006).

58. *Id.* at 1777.

59. See, e.g., Joshua D. Sarnoff, *Patent-Eligible Inventions After Bilski: History & Theory*, 63 HASTINGS L.J. 53, 65–66 (2011) (arguing that early limitations on patented inventions imposed by a Patent Board created by Congress in 1790 reflected an intent to restrict patent protection to inventions that promoted the progress of science and useful arts).

60. Walterscheid, *Preambular Argument*, *supra* note 54, at 348 (citing 13 THE DOCUMENTARY HISTORY OF THE FIRST FEDERAL CONGRESS OF THE UNITED STATES 1221 (Helen E. Veit et. al eds., 1994)).

61. *Id.* at 348–49 (citing American State Papers, Miscellaneous, Doc. No. 74, 4th Cong., 1st Sess. (1796); Annals of Congress, Feb. 3, 1796, at 288; A. HUNTER DUPREE, SCIENCE IN THE FEDERAL GOVERNMENT: A HISTORY OF POLICIES AND ACTIVITIES TO 1940, at 14 (1957)).

These instances of inaction suggest Congress interpreted narrowly its authority to encourage progress through means other than the grant of copyright or patent protection.⁶²

Professor Solum suggests that evidence of congressional inactivity may underdetermine original meaning. Interpreting the Constitution in light of early congressional action (or inaction) or judicial decisions might too easily reflect the political agendas or subconscious biases of legislators or judges, leading to interpretations inconsistent with the meaning the public might make from the language chosen.⁶³ But reliance on post-ratification (in)activity may well meet the expectations of at least some founders. For example, James Madison seems to have expected that some ambiguity would be resolved, liquidated, or made plain post-ratification in precisely this manner.⁶⁴ Evidence of post-ratification (in)activity would be particularly important for provisions for which there is little or no record of debate or discussion.

Indeed, such evidence has been persuasive in two recent Supreme Court cases considering the proper scope of congressional authority to extend copyright protection. In both *Eldred v. Ashcroft* and *Golan v. Holder*, the Supreme Court concluded that Congress had frequently exercised broad constitutional power to expand copyright protection,

62. Walterscheid, *Preambular Argument*, *supra* note 54, at 348.

63. Solum, *Triangulating*, *supra* note 2, at 1659–60; *see also* Craig W. Dallon, *Original Intent and the Copyright Clause: Eldred v. Ashcroft Gets It Right*, 50 ST. LOUIS L.J. 307, 357 (2006) (“Longevity of a practice alone, particularly where a constitutional challenge had never before been considered, should not define the reach of the Constitution. Interpreting the meaning of the Constitution based on the conduct of Congress undermines judicial review and diminishes the force of the Constitution.” (footnote omitted)); L. Ray Patterson, *What’s Wrong with Eldred? An Essay on Copyright Jurisprudence*, 10 J. INTELL. PROP. L. 345, 349 (2003) (criticizing the Court’s use of congressional practice to determine the original meaning instead of using the Constitution to evaluate congressional practice). Consider, however, circumstances in which Justice Scalia, an ardent originalist, turns to early post-ratification practice to locate original meaning. *See, e.g.*, Antonin Scalia, *Response*, in A MATTER OF INTERPRETATION 129, 145–46 (Amy Gutmann ed., 1998) (“[I]t is entirely clear that capital punishment, which was widely in use in 1791, does not violate the abstract moral principle of the Eighth Amendment.”); Thomas B. Colby & Peter J. Smith, *Living Originalism*, 59 DUKE L.J. 239, 296 (2009).

64. *See, e.g.*, Caleb Nelson, *Stare Decisis and Demonstrably Erroneous Precedents*, 87 VA. L. REV. 1, 13 (2001); Paul G. Ream, Note, *Liquidation of Constitutional Meaning Through Use*, 66 DUKE L.J. 1645 (2017); Will Baude, *Some Cold Water on the Liquidation Debate*, WASH. POST: THE VOLOKH CONSPIRACY (Apr. 14, 2017), https://www.washingtonpost.com/news/volokh-conspiracy/wp/2017/04/14/some-cold-water-on-the-liquidation-debate/?utm_term=.85fb61af0b47.

and construed the Copyright and Patent Clause accordingly.⁶⁵ Originalists must consider the possibility that in some cases early congressional (in)activity and judicial interpretation provide evidence of meaning that may well trump other indicators,⁶⁶ even if the interpretation driven by that evidence is not in consilience with corpus lexicography, immersion, or other evidence of constitutional history.⁶⁷

C. *Meaning(s) of Progress*

Prior scholarship has sought to interpret *progress* using immersion and historical methods, but to date, those analyses have not decided the question. Ambiguities remain unresolved. In areas like these, corpus lexicography might serve as a tie-breaker, or at least a thumb on the scale in favor of one meaning or another.

As Professor Solum notes, some originalists distinguish between interpretation and construction of constitutional language.⁶⁸ Interpretation “discovers the communicative content of the constitutional text.”⁶⁹ Construction “determines the legal effect of the constitutional text, including the decision of constitutional cases and the legal content of constitutional doctrines.”⁷⁰ Professor Solum’s article in this volume focuses on interpretation instead of

65. *Golan v. Holder*, 565 U.S. 302, 320 (2012) (holding that Congress could constitutionally re-extend copyright protection to works in the public domain, and finding such a grant consistent with the Copyright Act of 1790, enacted by the First Congress); *Eldred v. Ashcroft*, 537 U.S. 186, 194 (2003) (granting a twenty-year increase in the term of copyright protection to both existing and future works was consistent with the decision in the Copyright Act of 1790 to grant a fourteen-year renewal term to both existing and future works). *But see, e.g., id.* at 228 (Stevens, J., dissenting) (“Because the content of that first legislation, the debate that accompanied it, and the differences between the initial versions and the bills that ultimately passed provide strong evidence of early Congresses’ understanding of the constitutional limits of the Copyright/Patent Clause, I examine both the initial copyright and patent statutes.”); *id.* at 230, 237.

66. *Cf.* Oliver Wendell Holmes, *The Path of the Law*, 10 HARV. L. REV. 457, 458 (1897) (“[A] legal duty so called is nothing but a prediction that if a man does or omits certain things he will be made to suffer in this or that way by judgment of the court;—and so of a legal right.”).

67. On the dangers of appeals to the historical record in legal scholarship more generally, see Justin Hughes, *Copyright and Incomplete Historiographies: Of Piracy, Propertization, and Thomas Jefferson*, 79 S. CAL. L. REV. 993, 996–97 (2006) (critiquing the use of “incomplete historical claims” to argue for a recent “propertization” of copyright that in truth is hard to distinguish from eighteenth-century conceptions of copyright).

68. Solum, *Triangulating*, *supra* note 2, at 1628.

69. *Id.*

70. *Id.*

construction. But interpretation and construction are intimately intertwined. Indeed, scholars who have sought original meaning through immersion or constitutional history reach competing interpretations of *progress*, and those interpretations drive competing constructions of the proper scope of legislative authority to provide copyright and patent protection and judicial review of those enactments.⁷¹

1. Interpretation

As discussed above, scholars like Professor Pollack, then-Professor Lee, and Senator Hatch conclude, using datamining techniques from a limited corpus, that the most common meaning of *progress* is likely spread, dissemination, or distribution.⁷² Despite the evidence amassed, the “spread” definition has not been embraced by the majority of scholars.⁷³ Instead, most scholars have embraced a series of meanings that coalesce around the notion that *progress* means advancement in knowledge,⁷⁴ using phrases like “the encouragement of learning”⁷⁵ to refine the concept. Advancement can then be measured either in

71. Pollack, *What is Congress Supposed to Promote?*, *supra* note 18, at 804 (“The Idea of Progress, of course, allows the word ‘progress’ to accumulate [] disparate meanings”); *see also* Malla Pollack, *The Democratic Public Domain: Reconnecting the Modern First Amendment and the Original Progress Clause (a.k.a. Copyright and Patent Clause)*, 45 JURIMETRICS 23, 25, 26, 28 (2004); Malla Pollack, *Dampening the Illegitimacy of the United States’ Government: Reframing the Constitution from Contract to Promise*, 42 IDAHO L. REV. 123, 206 (2005).

72. *See supra* notes 47–52 and accompanying text.

73. *See, e.g.*, Jeanne C. Fromer, *The Intellectual Property Clause’s External Limitations*, 61 DUKE L.J. 1329, 1373–74 (2012) (critiquing and ultimately rejecting Pollack’s analysis).

74. Ned Snow, *Discrimination in the Copyright Clause*, 67 ALA. L. REV. 583, 594 (2016) (summarizing examples of “advances in knowledge” interpretations).

75. EDWARD C. WALTERSCHEID, *THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE: A STUDY IN HISTORICAL PERSPECTIVE* 150–51 (2002); L. Ray Patterson & Craig Joyce, *Copyright in 1791: An Essay Concerning the Founders’ View of the Copyright Power Granted to Congress in Article I, Section 8, Clause 8 of the U.S. Constitution*, 52 EMORY L.J. 909, 938 (2003).

quantitative increase⁷⁶ or qualitative improvement.⁷⁷ This advancement definition relies in part on founding era assumptions about the inevitability of progress.⁷⁸

Support for both the advancement and spread meanings can be unearthed in founding era dictionaries. Several dictionaries published close in time to the ratification of the Constitution include an advancement definition. For example, the 1785 edition of Samuel Johnson's *Dictionary of the English Language* offers, among its definitions, the following meaning of *progress*: "Intellectual improvement; advancement in knowledge; proficiency."⁷⁹ Noah Webster's 1828 edition of his *American Dictionary of the English Language* provides a definition that is essentially identical.⁸⁰ But these same dictionaries also offer definitions consistent with the spread meaning. For example, the first listed definition of *progress* in Webster's 1828 dictionary indicates that *progress* can be "a moving or going forward; a proceeding onward."⁸¹

Another group of scholars critique the presumptions undergirding a classic interpretation of *progress*. Advances in scientific knowledge or the technology we use to disseminate information do not guarantee a "state of 'betterment' for others."⁸² These "postmodern" analyses instead conclude *progress* is best understood to indicate advancement

76. See, e.g., Paul J. Heald & Suzanna Sherry, *Implied Limits on the Legislative Power: The Intellectual Property Clause as an Absolute Constraint on Congress*, 2000 U. ILL. L. REV. 1119, 1173; L. Ray Patterson, *Eldred v. Reno: An Example of the Law of Unintended Consequences*, 8 J. INTEL. PROP. L. 223, 234 (2001) ("[T]he creation of a new work is the unalterable condition for copyright, a condition that the Framers . . . made a part of the Copyright Clause" by authorizing Congress to "promote the Progress of Science.").

77. Fromer, *supra* note 73, at 1373 ("To promote the Progress of Science and useful Arts' generally refers to the goal of encouraging the advancement of systematic knowledge, cultural knowledge, and technology. . . . [This] understanding can further mean improvement in a knowledge base's quality or quantity.").

78. Birnhack, *supra* note 44, at 19–20; Chon, *supra* note 40, at 117.

79. 2 SAMUEL JOHNSON, A DICTIONARY OF THE ENGLISH LANGUAGE (6th ed. 1785).

80. 2 NOAH WEBSTER, AN AMERICAN DICTIONARY OF THE ENGLISH LANGUAGE (Johnson Reprint Corp. 1970) (1828) (*progress* defined as "advance in knowledge; intellectual or moral improvement; proficiency").

81. *Id.*

82. Jessica Silbey, *Promoting Progress: A Qualitative Analysis of Creative and Innovative Production*, in THE SAGE HANDBOOK OF INTELLECTUAL PROPERTY 515, 519 (Matthew David & Debora Halbert eds. 2014).

of civilization collectively or the improved condition of humankind,⁸³ through responsible stewardship and distributional fairness.⁸⁴

2. Construction

These competing interpretations of *progress* can lead to differing constructions of congressional power and the proper scope of judicial review. Some scholars conclude that the progress language is primarily an open conveyance of power. For instance, Craig Dallon concludes that the Copyright and Patent Clause should be “understood broadly to empower Congress to make changes to copyright law that it rationally believes improve the overall copyright scheme.”⁸⁵ Robert Merges similarly concludes that “[g]iven a constitutional provision rooted in a blind faith in ‘progress,’ we cannot read in historically contingent limitations on patentable subject matter.”⁸⁶ Others conclude that the progress language conveys only limited power to Congress, and thus limits Congress’s authority to extend protection to copyrighted expression or patented inventions. For example, Professor Oliar concludes after his intensive study of drafts from the framing era that “[t]he Framers intended the progress language in the Clause—‘to promote the progress of science and useful art’—to limit Congress’s power to grant IP rights.”⁸⁷

Those scholars who see the Progress phrase as a limit also differ as to *how* the phrase cabins congressional authority. For example, Paul

83. Similar language is found in several state copyright statutes from the Confederation period. *See, e.g.*, Dallon, *supra* note 63, at 325 n.131 (quoting COPYRIGHT ENACTMENTS: LAWS PASSED IN THE UNITED STATES SINCE 1783 RELATING TO COPYRIGHT 4, 8–9 (Copyright Office ed., Bulletin No. 3 rev. 1963)) (“The Massachusetts, New Hampshire, and Rhode Island copyright acts made reference to the need to encourage ‘the progress of civilization.’”). The Constitution embeds two notions of progress: intellectual progress, or progress regarding knowledge, and social and political organization (a better frame for copyright law), and material progress, for example, progress in technology and natural sciences (a better frame for patent law). *See* Birnhack, *supra* note 44, at 16–17.

84. *See, e.g.*, Chon, *supra* note 40, at 101–02.

85. Dallon, *supra* note 63, at 357; *see also* 1 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 1.03[B] (Matthew Bender rev. ed. 1963), Lexis (arguing that the introductory phrase has expanded, rather than limited, Congress’s authority).

86. Robert P. Merges, *As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform*, 14 BERKELEY TECH. L.J. 577, 587 (1999).

87. Dotan Oliar, *The (Constitutional) Convention on IP: A New Reading*, 57 UCLA L. REV. 421, 423 (2009).

Heald and Suzanna Sherry argue this limitation must be operationalized by requiring an express quid pro quo from authors and inventors before Congress can grant an exclusive right.⁸⁸ A parallel argument is marshaled by several scholars in support of subject matter limitations on patentable inventions.⁸⁹ Senator Hatch and Professor Lee instead conclude that a grant is constitutionally permissible, following the spread meaning, if it does no more than encourage the dissemination of already created works.⁹⁰ Interestingly, (and perhaps troublingly for Professor Solum's anticipated consilience), Professor Pollack likewise concludes *spread* is the most common meaning of *progress* but nevertheless construes the Copyright and Patent Clause to bar new grants that do not tie protection to the creation of new works.⁹¹

Similarly, some scholars have argued that *progress* requires an increase in the quality of copyrighted expression and patented inventions. Professor Oliar suggests this limit should require courts to consider whether new legislative grants of rights to copyright or patent owners provide benefits greater than the costs imposed.⁹² But Barton

88. Heald & Sherry, *supra* note 76, at 1162 (“[A]n author or inventor may not be given something for nothing; the author or the inventor must give the public something it did not have before to earn a grant of exclusive rights from Congress.”).

89. See, e.g., Brief of Eleven Law Professors and AARP as Amici Curiae in Support of Respondent at 33, *Bilski v. Kappos*, 561 U.S. 593 (2010) (No. 08-964) (citing Suzanne Scotchmer, *Standing on the Shoulders of Giants: Cumulative Research and the Patent Law*, 5 J. ECON. PERSP. 29 (1991)) (discussing the effects of patents on sequential innovation).

90. Hatch & Lee, *supra* note 51, at 16–20.

91. Pollack, *What is Congress Supposed to Promote?*, *supra* note 18, at 761 (“If a court thoughtfully considers ‘progress’ (under any definition), the CTEA should be held unconstitutional in all its applications. The *Eldred* court merely invoked the alleged upside of the change without considering the downside—an improper way to do any type of cost/benefit analysis.”); see also *id.* at 763 (“[T]he CTEA only claims to promote ‘progress,’ if ‘progress’ means ‘economic value.’ . . . [But giving] copyright holders more of the financial value of the work . . . [or] supporting a strong export industry . . . [does not] conceivably promote ‘progress’ if that word means either ‘quality improvement’ or ‘spread.’”).

92. Oliar, *Making Sense of the Intellectual Property Clause*, *supra* note 57, at 1789–1807, 1840–41 (arguing that *progress* is consistent with “advancement” or promotion, and that courts should declare unconstitutional those statutory enactments that impose costs greater than their marginal benefits); Hatch & Lee, *supra* note 51, at 3 (“[P]rogress . . . encompass[es] . . . an increase in quantity or quality of works . . .”); cf. Matthew Sag & Kurt Rohde, *Patent Reform and Differential Impact*, 8 MINN. J.L. SCI. & TECH. 1, 2 (2007) (“Technology heavyweights . . . fear that, rather than encouraging the ‘progress of science and the useful arts’ as required by the United States Constitution, declining patent quality and overly broad patent rights are reducing incentives to invest in manufacturing, research and development.” (citations

Beebe cautions that efforts to use copyright protection to increase the aesthetic quality of copyrighted expression is not only inconsistent with the intent of the Framers,⁹³ but also likely futile, ignoring differences between disciplines of knowledge and art.⁹⁴ Scholars also disagree over whether *progress* is naturally promoted by incentivizing and rewarding creation,⁹⁵ or whether *progress* instead requires favoring privileges for the use and reuse of copyrighted works and patented inventions.⁹⁶

D. Resolving Definitional Disputes

Professor Solum anticipates that in some cases, the tools of immersion, examination of the constitutional record, and corpus linguistic analysis will lead to consilience. For example, in his essay, Professor Solum provides a first cut analysis of collocates of *science* from 1810 forward, using COHA,⁹⁷ which suggest that *science* as used in the Copyright and Patent Clause is a category that encompasses more branches of knowledge than STEM disciplines. There is widespread agreement on this point. The Supreme Court has so

omitted)). *But see* Pollack, *What is Congress Supposed to Promote?*, *supra* note 18, at 803 (“[L]ate eighteenth century speakers of English more commonly used ‘improvement,’ ‘perfection,’ or ‘advancement’ (as opposed to ‘progress’) when referring to the betterment of mankind’s knowledge base.”).

93. Barton Beebe, Bleistein, *The Problem of Aesthetic Progress, and the Making of American Copyright Law*, 117 COLUM. L. REV. 319, 338 (2017) (concluding from the drafting evidence that the decision of the Framers to exclude “fine arts from the language of the Progress Clause appears to have been a deliberate act”); *see also id.* at 340.

94. *Id.* at 341–42.

95. Steven R. Englund, Note, *Idea, Process, or Protected Expression?: Determining the Scope of Copyright Protection of the Structure of Computer Programs*, 88 MICH. L. REV. 866, 891 (1990) (citing *Sony Corp. of Am. v. Universal City Studios*, 464 U.S. 417, 429 (1984)) (“Copyright promotes progress by acting as an incentive to authors to create new works and ultimately placing those works in the hands of the public.”); *see also* Roberta Rosenthal Kwall, *Governmental Use of Copyrighted Property: The Sovereign’s Prerogative*, 67 TEX. L. REV. 685, 730 (1989) (“[T]hose rewards are essential to effectuating copyright law’s major objective of enhancing societal progress, because an absence of monetary protections might well result in diminished creativity.”).

96. Jessica Litman, *Readers’ Copyright*, 58 J. COPYRIGHT SOC’Y U.S.A. 325, 350 (2010) (“[C]opyright’s purpose is (and should be) to encourage reading, listening, and viewing” both by “creative remixers” and “passive consumers.”); *see also id.* at 339 (“But a moment’s reflection should reveal that a copyright system with no readers, listeners or viewers to enjoy the copyrighted works that the system produces has no plausible mechanism for promoting the progress of anything.”).

97. Solum, *Triangulating*, *supra* note 2, at 1646.

stated,⁹⁸ and scholars who immerse themselves in key texts have reached a similar conclusion.⁹⁹ There is little reason to suspect that analysis using the forthcoming Corpus of the Founding Era American English (COFEA), a corpus containing texts contemporaneous with the drafting and ratification era, will provide evidence of use that leads to a different interpretation of *science*.

Where immersion and the constitutional record cut in different directions or are susceptible to multiple readings, as discussed in the examples above, corpus linguistic evidence of usage may allow a scholar to triangulate the original public meaning with a higher level of confidence. To date, computer-assisted linguistic analysis of the meaning of *progress* using a large corpus of texts contemporaneous with the framing of the Constitution has not been conducted.¹⁰⁰ Linguistic analysis using a corpus like COFEA may well aid courts in resolving disputes about its proper construction. Recent Supreme Court precedent indicates the continued importance of proper construction. While the Supreme Court has heretofore resisted calls to apply heightened scrutiny to congressional copyright statutes grounded either in the First Amendment or the Progress language, the Court signaled that a decision by Congress to tamper with a “traditional contour”¹⁰¹ of copyright law like the idea/expression divide or fair use might trigger heightened scrutiny.¹⁰² Determining whether congressional activity falls outside the scope of its authority in those areas may well turn in part on the proper definition of *progress*. For example, if the most common public meaning of *progress* is spread

98. *Golan v. Holder*, 565 U.S. 302, 324 (2012) (“Perhaps counterintuitively for the contemporary reader, Congress’ copyright authority is tied to the progress of science; its patent authority, to the progress of the useful arts.”).

99. Beebe, *supra* note 93, at 323 (concluding that science, as distinguished from useful arts, “refer[s] to systematic theoretical and empirical knowledge,” but “neither category encompassed the fine arts”).

100. *But see supra* notes 44–50 and accompanying text.

101. These traditional contours work like buffers between the First Amendment and copyright protection. *See, e.g.*, Christopher Sprigman, *Reform(aliz)ing Copyright*, 57 STAN. L. REV. 485, 527 (2004).

102. *Eldred v. Ashcroft*, 537 U.S. 186, 221 (2003) (“[W]hen, as in this case, Congress has not altered the traditional contours of copyright protection, further First Amendment scrutiny is unnecessary.”). *But see id.* at 264 (Breyer, J., dissenting) (“We cannot avoid the need to examine the statute carefully by saying that ‘Congress has not altered the traditional contours of copyright protection,’ for the sentence points to the question, rather than the answer. Nor should we avoid that examination here.” (internal citations omitted)).

or distribution, courts should construe fair use and the idea-expression line consistent with that interpretation,¹⁰³ which will generally mean a broader fair use exception to copyright protection, and more elements of a work falling on the unprotected idea side of the line.

IV. USING CORPUS LINGUISTICS TO DISCOVER MODULATION

In his article, Professor Solum cautions that modulation—an attempt to use old words in new ways—might go unrecognized by corpus lexicography.¹⁰⁴ Corpus analysis should reveal commonly used meanings, but the meanings revealed will necessarily work in many contexts. To that extent, those meanings are acontextual. On the other hand, by definition, modulation deviates from established meaning. Corpus lexicography reveals the ways a word is typically used, but it will not disclose whether a given usage is an attempted modulation, whereby the user requires an existing term to bear new meaning in a specific context. An approach focused on uncovering ordinary public meaning will thus normally fail to catch a modulation in meaning.

Context can help reveal modulation in some cases. For example, Article II, Section 2, Clause 3 of the Constitution gives the President the power to “fill up all Vacancies” in appointed offices “during the Recess of the Senate.”¹⁰⁵ Professor Solum notes that the term *recess* read acontextually and in accordance with its ordinary public meaning could indicate any break in activity.¹⁰⁶ Thus, the Recess Appointments Clause could be read to allow the President to fill a vacancy and appoint officers of the United States during any break of Senate

103. See, e.g., Michael D. Murray, *Reconstructing the Contours of the Copyright Originality and Idea-Expression Doctrines Regarding the Right to Deny Access to Works*, 1 TEX. A&M L. REV. 921, 930 (2014) (“The design of the contours of copyright relies on the policy of access to information in the arts and sciences. If access is denied or cut-off, or made practically inaccessible through onerous barriers, then the primary purpose of the copyright scheme is unfulfilled.”).

104. Solum, *Triangulating*, *supra* note 2, at 1648; see also Lawrence B. Solum, *Congress’s Power to Promote the Progress of Science: Eldred v. Ashcroft*, 36 LOY. L.A. L. REV. 1, 46 (2002) [hereinafter Solum, *Congress’s Power*] (“Evidence that the primary or most frequent usage of ‘progress’ in the founding era was spatial or geographic does not answer the question as to whether that was the use made by those who framed or ratified the constitution.”).

105. U.S. CONST. art. II, § 2, cl. 3 (“The President shall have Power to fill up all Vacancies that may happen during the Recess of the Senate, by granting Commissions which shall expire at the End of their next Session.”).

106. Solum, *Triangulating*, *supra* note 2, at 1637.

business, even a lunch break.¹⁰⁷ However, in context, *recess* most likely refers to a sustained break in the regular business of the Senate, which would necessitate the President filling vacancies without the advice and consent of the Senate.¹⁰⁸ This modulated use of *recess* is easy to detect. But one might reasonably worry that other key constitutional terms are modulated uses that could be more difficult to detect. Evidence of regular usage that predates or is contemporaneous could reasonably fail to reflect the novel use.

However, corpus lexicography might also allow one to see clearly how the usage of key terms change after an inflection event like the passage of the Constitution. Consider some recent modulations in American English that have become ordinary, acontextual meanings. The term *sandy* has modulated to add a new meaning, referring to the hurricane that hit the Eastern seaboard of the United States in October 2012.¹⁰⁹ Computational linguists posit that this sort of semantic shift can be detected in one of three ways. First, a spike in the *frequency* of word usage could signal meaningful change. For example, an increase in references to *Sandy* during coverage of and discussion about Hurricane Sandy as tracked by Google Trends presaged a new ordinary meaning.¹¹⁰ Second, semantic shift often occurs through a change in *syntax* when an existing word begins to be used as a new part of speech. For instance, *apple* is a noun describing a fruit, but it shifted around 1984 to add a proper noun sense as the public became more aware of and more interested in Apple computers.¹¹¹ Finally, collocation data can indicate a shift in the meaning of a term as it begins to appear in new contexts. For example, in 1900, the term *gay* frequently appeared with words like *dapper* and *courteous*, but by 1990, the term frequently occupied the same semantic space as *homosexual* and *transgender*.¹¹²

It is also possible that semantic shift could closely follow the ratification of the Constitution. The drafting and ratification of the

107. *Id.*

108. U.S. CONST. art. II, § 2, cl. 3; *see also* NLRB v. Noel Canning, 134 S. Ct. 2550, 2559 (2014) (interpreting the Recess Appointments Clause and “put[ting] significant weight upon historical practice”).

109. *See* Vivek Kulkarni et al., *Statistically Significant Detection of Linguistic Change*, in WWW '15 PROCEEDINGS OF THE 24TH INT'L CONF. ON WORLD WIDE WEB 625, 626 (2015).

110. *Id.*

111. *Id.* at 627.

112. *Id.* at 625.

Constitution was a critical inflection point in American history. Such an event might create linguistic ripples. One might therefore detect post-ratification changes in ordinary meanings for key terms used in the Constitution. Founding era corpora could be analyzed for similar changes in frequency, syntax, and context to identify whether there is a post-ratification shift in meaning.

For example, corpus linguistic data might be mined to track changes in the use of *progress*, especially as the use collocates with terms like *science* and *art*. Such a shift might reasonably be attributed to post-ratification usage. Post-ratification usage may not only reflect increased collocation between those terms but also modulation in meaning over time, proceeding from ratification. Indeed, in her analysis of the usage of *progress* in the *Pennsylvania Gazette*, Professor Pollack notes that some uses of *progress* that post-date ratification embodying the idea of *progress* as qualitative improvement might “be unreflective echoes of the constitutional phrase.”¹¹³ Those might also reflect a modulation—meaning acquired precisely because of the ratification event. Further examination of corpora that can be mined chronologically may reveal additional evidence that supports or refutes such a shift.

V. LIMITATIONS OF CORPUS LINGUISTICS

Corpus lexicography done well should provide better evidence of original meaning than an appeal to a dictionary alone.¹¹⁴ But as Professor Solum concludes, corpus lexicography cannot replace study of the constitutional record or immersion. As the discussion of modulation reveals,¹¹⁵ corpus lexicography is not well suited to catch attempted unconventional uses, although it may be used to show the effect of a modulation that has become an ordinary meaning. More generally, corpus lexicography, like textual immersion, might sometimes be maladapted to uncover meaning that matters. As Professor Solum cautions, textual immersion might likewise misfire if the corpus of texts selected are inadequate to reflect usage across a

113. Pollack, *What is Congress Supposed to Promote*, *supra* note 18, at 801.

114. Solum, *Triangulating*, *supra* note 2, at 1668, 1670.

115. *See id.* at 1650–51.

broad swath of the public.¹¹⁶ He thus concludes corpus lexicography may provide better evidence of ordinary meaning, in part because one can sample from a much larger body of literature using computing technology, including selecting texts from the appropriate subculture.¹¹⁷ A search for original public meaning permits a level of agnosticism regarding which texts may or must be included in a corpus. If every text is equally likely to reveal public meaning through usage by ordinary speakers (in the case of transcripts) and writers, then any text can be included in the corpus. But a poorly constructed corpus can be both voluminous and unrepresentative. For example, corpus linguists note the importance of a variety of sources if a corpus is to provide generalizable findings.¹¹⁸

The use of corpus lexicography to conduct an originalist inquiry also may mistakenly presume the answer to an important question that Professor Solum posed in earlier work anticipating the Supreme Court's decision in *Eldred v. Ashcroft*¹¹⁹ some fifteen years ago. Leaving aside attempted modulation, can we assume the Framers intended to adopt the most frequently used meaning, or that the Ratifiers were ratifying that meaning?¹²⁰ A search for ordinary public meaning embraces that simplifying assumption. To the extent that assumption is misguided, corpus lexicography data might not only be misapplied but actually impede the correct interpretation and construction of constitutional language.¹²¹

116. *Id.* at 1667.

117. *Id.*

118. Douglas Biber, *Representativeness in Corpus Design*, 8 LITERARY & LINGUISTIC COMPUTING 243, 243 (1993) (“[I]f a corpus does not represent the range of text types in a population, it will not represent the range of linguistic distributions”); Geoffrey Leech, *Corpora and Theories of Linguistic Performance*, in DIRECTIONS IN CORPUS LINGUISTICS: PROCEEDINGS OF NOBEL SYMPOSIUM 82, at 105, 112–13 (Jan Svartvik ed., 1992).

119. *Eldred v. Ashcroft*, 537 U.S. 186 (2003).

120. Solum, *Congress's Power*, *supra* note 104, at 46 (“Evidence that the primary or most frequent usage of ‘progress’ in the founding era was spatial or geographic does not answer the question as to whether that was the use made by those who framed or ratified the constitution.”).

121. Using inapplicable but easily accessible data can contribute to behavior based on bad information, triggering a “streetlight” effect, named after the famous (and likely apocryphal) story about a drunk looking on 3rd Street for keys he lost on 1st Street because the light is better on 3rd. See Karen M. Gebbia, *Circuit Splits and Empiricism in the Supreme Court*, 36 PACE L. REV. 477, 556 (2016); David H. Freedman, *Why Scientific Studies Are So Often Wrong: The Streetlight Effect*, DISCOVER, July–Aug. 2010, <http://discovermagazine.com/2010/jul-aug/29-why-scientific-studies-often-wrong-streetlight-effect>.

In the end, if corpus lexicography is to improve originalist analysis, its use must effectively constrain judicial discretion. But analysis even from a balanced, representative corpus can still yield biased results if the analyst consciously or unconsciously biases the search.¹²² It would be erroneous to make too much of this concern; originalism itself is an imperfect constraint.¹²³ As with originalism, imperfections in corpus data mining do not invalidate its use.¹²⁴

Despite these limitations, corpus lexicography has much to offer for deciphering original public meaning. Evidence of usage may confirm or challenge interpretation based on textual immersion and constitutional history. Consilience between those methods may liquidate ambiguity and thus harmonize disputed interpretations.¹²⁵ In addition, corpus analysis may prove useful in identifying modulations that become standard meanings.

122. Dominic Stewart, *Safeguarding the Lexicogrammatical Environment: Translating Semantic Prosody*, in *CORPUS USE AND TRANSLATING: CORPUS USE FOR LEARNING TO TRANSLATE AND LEARNING CORPUS USE TO TRANSLATE* 29, 41–42 (Allison Beeby, Patricia Rodríguez Inés & Pilar Sánchez-Gijón eds., 2009) (arguing that it can be difficult to cabin the analyst’s personal insights and prevent bias in the results).

123. William Baude, *Originalism as a Constraint on Judges*, 84 U. CHI. L. REV. (forthcoming 2018) (positing that originalism cannot perfectly constrain judges, but may serve as a helpful tool for judges who seek to constrain themselves).

124. Sue Atkins et al., *Corpus Design Criteria*, 7 LITERARY & LINGUISTIC COMPUTING 1, 6 (1992) (“It would be short-sighted indeed to wait until one can scientifically balance a corpus before starting to use one, and hasty to dismiss the results of corpus analysis as ‘unreliable’ or ‘irrelevant’ because the corpus used cannot be proved to be ‘balanced.’”).

125. Cf. THE FEDERALIST NO. 37 (James Madison) (explaining the Constitution would be “obscure and equivocal, until [its] meaning be liquidated and ascertained by a series of particular discussions and adjudications”).

