

9-30-2018

Corpus Linguistics in the Chevron Two-Step

Jacob Crump

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



Part of the [Applied Linguistics Commons](#), and the [Law Commons](#)

Recommended Citation

Jacob Crump, *Corpus Linguistics in the Chevron Two-Step*, 2018 BYU L. Rev. 399 (2018).

Available at: <https://digitalcommons.law.byu.edu/lawreview/vol2018/iss2/14>

This Comment 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.

Corpus Linguistics in the *Chevron* Two-Step

CONTENTS

I. INTRODUCTION	399
II. <i>CHEVRON U.S.A. v. NATURAL RESOURCES DEFENSE COUNCIL, INC.</i>	402
III. CORPUS LINGUISTICS.....	404
IV. <i>MCI TELECOMMUNICATIONS CORP. v. AT&T Co.</i>	408
V. CORPUS ANALYSIS OF <i>MODIFY</i>	410
A. Concordance Analysis.....	411
B. Collocate Analysis.....	414
VI. IMPLICATIONS	417

I. INTRODUCTION

*Chevron U.S.A. Inc. v. Natural Resources Defense Council, Inc.*¹ changed the relationship between administrative agencies and the courts that frequently review agency action.² *Chevron* established a famous two-step analysis for determining whether an agency interpretation of a statute will be given controlling weight by reviewing courts. Step One requires that a court—using “traditional tools of statutory construction”³—determine whether Congress has directly spoken to the precise point at issue or whether the statute

1. *Chevron U.S.A. Inc. v. Nat. Res. Def. Council, Inc.*, 467 U.S. 837 (1984).

2. Jack M. Beermann, *End the Failed Chevron Experiment Now: How Chevron Has Failed and Why It Can and Should Be Overruled*, 42 CONN. L. REV. 779, 781 (2010) (“In *Chevron* . . . , the Supreme Court announced a startling new two-step approach to judicial review of statutory interpretation by administrative agencies . . .”). However, not all commentators see *Chevron* as revolutionary. See, e.g., Michael Herz, *Chevron Is Dead; Long Live Chevron*, 115 COLUM. L. REV. 1867, 1867 (2015) (“*Chevron* is not a revolutionary shift of authority from the judiciary to the executive. That *Chevron* is dead.”).

3. *Chevron*, 467 U.S. at 843 n.9 (“If a court, employing traditional tools of statutory construction, ascertains that Congress had an intention on the precise question at issue, that intention is the law and must be given effect.”).

is ambiguous.⁴ If a court finds ambiguity or silence in the statute, it assumes that the ambiguity constitutes an implicit delegation by Congress to the agency to interpret the statute authoritatively.⁵ In such cases, the court proceeds to Step Two, where it asks whether the agency's interpretation is a reasonable one. If it is, courts then give controlling weight to the agency's authoritative interpretation.⁶

Step One requires courts to make difficult determinations⁷: Did Congress directly speak to the precise question at issue in this case? What does it mean for Congress to speak "directly" or "precisely?" Step Two's reasonability analysis can be even less tethered to the text of the statute, and therefore, more difficult: Is the agency's interpretation a "reasonable" one?⁸ Many of the difficulties presented to courts by the two steps are similar, and as a result, the line between Step One and Step Two can be blurry.⁹

Chevron has its fair share of critics and criticisms,¹⁰ but the criticism most relevant to this Comment is that *Chevron*—though aimed at reducing judges' ability to insert their own policy

4. *Id.* at 842–43.

5. *Id.* at 843–44.

6. *Id.* at 844 ("In such a case, a court may not substitute its own construction of a statutory provision for a reasonable interpretation made by the administrator of an agency.").

7. See GARY LAWSON, FEDERAL ADMINISTRATIVE LAW 608–57 (6th ed. 2012) (outlining difficulties in *Chevron's* Step One).

8. According to some scholars, "[t]here is no universally accepted test for determining when an interpretation is impermissible or, as modern cases tend to frame it, unreasonable." Gary Lawson & Stephen Kam, *Making Law out of Nothing at All: The Origins of the Chevron Doctrine*, 65 ADMIN. L. REV. 1, 3 n.10 (2013).

9. See, e.g., Matthew C. Stephenson & Adrian Vermeule, *Chevron Has Only One Step*, 95 VA. L. REV. 597, 597 (2009); Glob. Crossing Telecomms., Inc. v. Metrophones Telecomms., Inc., 550 U.S. 45, 55–58 (2007) (conducting a muddled *Chevron* analysis and then concluding that the agency's "determination is a reasonable one; hence it is lawful").

10. Some commentators and judges argue that the doctrine upsets the separation of powers. See, e.g., *Michigan v. EPA*, 135 S. Ct. 2699, 2712 (2015) (Thomas, J., concurring) ("*Chevron* deference raises serious separation-of-powers questions."). Cf. Richard H. Fallon, Jr., *Of Legislative Courts, Administrative Agencies, and Article III*, 101 HARV. L. REV. 915, 983 (1988) ("[A]n Article III court must exercise independent judgment concerning all questions of law that it is called upon to decide. Separation-of-powers values call for this conclusion, and sometimes emphatically so."). Others argue that *Chevron* violates the Administrative Procedure Act. See, e.g., William N. Eskridge, Jr. & Lauren E. Baer, *The Continuum of Deference: Supreme Court Treatment of Agency Statutory Interpretations from Chevron to Hamdan*, 96 GEO. L.J. 1083, 1160 (2008) (noting that *Chevron* deference is incompatible with the APA's requirements that "the reviewing court shall decide all relevant questions of law" and shall "interpret constitutional and statutory provisions").

preferences into ambiguous statutes¹¹ – still allows judges to consider those preferences.¹² Before the *Chevron* regime was introduced in 1984, questions of deference to agency interpretations of law were generally analyzed under the framework established by the U.S. Supreme Court in *Skidmore v. Swift & Co.*¹³ Under *Skidmore*, courts examine several factors to determine the persuasive weight of the agency’s interpretation.¹⁴ Because of *Skidmore*’s malleability, judges were easily able to slip policy preferences into a decision.

Chevron was supposed to help solve this free agent problem, but it may not have. Step One allows for enough discretion that courts can produce results all over the map, and Step Two’s inquiry into reasonableness or permissibility is even more nebulous. What seems *reasonable* to one judge can appear absurd to another, because our individual experiences easily influence our perception of reasonableness.¹⁵

Especially problematic *Chevron* cases are those where the question of agency deference hinges on a single term or phrase. In these cases, the temptation is for judges to reflexively turn to dictionaries to marshal support for their own intuitions about linguistic ambiguity and the reasonableness of various interpretations. But the problem is, this type of reasoning allows judges to look out over the crowd of dictionary definitions and pick out their friends.¹⁶

So what can be done? One possible solution is to focus on more data-driven, transparent approaches to discerning language meaning. This Comment discusses one possible tool in a judge’s interpretive toolkit—corpus linguistics—and suggests it can be

11. See *supra* note 6.

12. See Beermann, *supra* note 2, at 783 (“*Chevron* is so pliable that courts applying it can still reach any desired result”); see also Brett M. Kavanaugh, *Fixing Statutory Interpretation*, 129 HARV. L. REV. 2118, 2154 (2016) (reviewing ROBERT A. KATZMANN, *JUDGING STATUTES* (2014)) (“*Chevron* . . . is . . . indeterminate—and thus can be antithetical to the neutral, impartial rule of law”).

13. *Skidmore v. Swift & Co.*, 323 U.S. 134 (1944).

14. *Id.* at 140.

15. Cf. Jeffrey Usman, *Capital Punishment, Cultural Competency, and Litigating Intellectual Disability*, 42 U. MEM. L. REV. 855, 899–900 (2012) (“For judges, cultural attitudes, values, and beliefs will shape thinking and decision making in a case.”) (internal alterations and quotations omitted).

16. See ANTONIN SCALIA, *A MATTER OF INTERPRETATION: FEDERAL COURTS AND THE LAW* 36 (1997) (criticizing the “variety and specificity” of results facilitated by the use of legislative history in statutory interpretation).

used to achieve more objective, data-driven results in *Chevron* cases. Corpus linguistics provides access to patterns and trends of actual language usage and allows for empirical determinations regarding linguistic meaning. It thus has significant implications for administrative law by helping courts arrive at more demonstrably accurate semantic conclusions in *Chevron* cases.

To demonstrate how this can be done, this Comment examines the Supreme Court's 1994 decision in *MCI Telecommunications Corp. v. AT&T Co.*¹⁷ To begin, Part II of this Comment briefly reviews the *Chevron* decision. Part III then introduces corpus linguistics and identifies advantages this approach has over the use of traditional dictionaries in many contexts. Part IV describes the Supreme Court's reasoning and decision in *MCI*. Part V then applies a corpus linguistic analysis to the Court's decision in *MCI* and concludes that the Court's opinion, which held that the term *modify* could not reasonably be interpreted to include fundamental change, was probably incorrect as a matter of ordinary language use. Data compiled from the Corpus of Historical American English (COHA),¹⁸ the Corpus of Contemporary American English (COCA),¹⁹ and the News on the Web (NOW) Corpus²⁰ demonstrate that *modify* can reasonably be used to signify substantial changes, in addition to minor ones. Part VI then briefly suggests ways in which corpus linguistics can impact the world of judicial deference to agency interpretation of law under *Chevron*.

II. *CHEVRON U.S.A. v. NATURAL RESOURCES DEFENSE COUNCIL, INC.*

In 1977, Congress amended the Clean Air Act.²¹ In implementing the amendments, the EPA promulgated regulations adopting a

17. *MCI Telecomms. Corp. v. AT&T Co.*, 512 U.S. 218 (1994).

18. Mark Davies, CORPUS HIST. AM. ENG. (COHA), <http://corpus.byu.edu/coha/> (last visited June 22, 2018). COHA contains more than 400 million words from the period 1810 to 2010. *Id.*

19. Mark Davies, CORPUS CONTEMP. AM. ENG. (COCA), <http://corpus.byu.edu/coca/> (last visited June 26, 2018). COCA contains more than 560 million words of American English from the period 1990 to 2017. *Id.*

20. Mark Davies, NEWS ON WEB CORPUS (NOW), <http://corpus.byu.edu/now/> (last visited June 26, 2018). NOW is updated daily and currently contains more than 6 billion words from 2010 to the present time. *Id.*

21. Pub. L. No. 95-95, 91 Stat. 685 (1977).

plant-wide definition of the term “stationary source,” allowing certain types of polluting entities to make changes to their facilities so long as the total polluting emissions from the plant did not increase.²²

The Natural Resources Defense Council challenged the EPA’s regulation. The D.C. Circuit set aside the new regulations, reasoning that in light of ambiguity found in the amendments and in the amendments’ legislative history, “the purposes of the non-attainment program should guide [the court’s] decision.”²³ The court found the EPA’s regulation to be at odds with those purposes and therefore concluded that the regulation was contrary to law.²⁴

On appeal, the Supreme Court reversed. The Court held that

[w]hen a court reviews an agency’s construction of the statute which it administers, it is confronted with two questions. First, always, is the question whether Congress has directly spoken to the precise question at issue. If the intent of Congress is clear, that is the end of the matter; for the court, as well as the agency, must give effect to the unambiguously expressed intent of Congress. If, however, the court determines Congress has not directly addressed the precise question at issue, the court does not simply impose its own construction on the statute, as would be necessary in the absence of an administrative interpretation. Rather, if the statute is silent or ambiguous with respect to the specific issue, the question for the court is whether the agency’s answer is based on a permissible construction of the statute.²⁵

The Court noted that this deference applied to both explicit and implicit delegations of authority by Congress to agencies.²⁶ In either case, the Court concluded that “a court may not substitute its own construction of a statutory provision for a reasonable interpretation made by the administrator of an agency.”²⁷

Thus was born one of the most transformative doctrines in modern Supreme Court history – the *Chevron* Two-Step – from what quickly became one of the most important administrative law

22. *Chevron U.S.A., Inc. v. Nat. Res. Def. Council, Inc.*, 467 U.S. 837, 840 (1984).

23. *Nat. Res. Def. Council, Inc. v. Gorsuch*, 685 F.2d 718, 726 n.39 (D.C. Cir. 1982), *rev’d*, 467 U.S. 837, 840 (1984).

24. *Id.* at 726.

25. *Chevron*, 467 U.S. at 842–43.

26. *Id.* at 843–44.

27. *Id.* at 844.

cases ever decided.²⁸ But despite its importance (or perhaps because of it), *Chevron* often proves difficult for courts to apply consistently. A *Chevron* analysis combines the difficulty of statutory interpretation with judgments about reasonability and, therefore, presents courts with questions to which there are frequently no easy answers. Fortunately, a branch of linguistics offers ways to bring more objectivity, empirical accuracy, and transparency into the semantic inquiry *Chevron* demands.

III. CORPUS LINGUISTICS

Corpus linguistics is an approach to the study of language use and function that employs large databases of language, essentially a “big-data” approach to discovering linguistic meaning. A corpus linguistics approach typically boils down to “detailed searches for words and phrases in multiple contexts across large amounts of text.”²⁹ A corpus can be composed of any language material,³⁰ and there are reasons for using specialized corpora or constructing corpora in certain ways.³¹ At bottom, however, essentially any person who examines large amounts of naturally occurring language to discern patterns of use and meaning is using a corpus linguistic approach. In this way, corpus linguistics analysis is something that most of us do in our heads every day: We take instances of language use – from what we read, what we hear, what we speak, etc. – and analyze them to piece together meaning.

Today, corpora are often enormous in size, with millions or billions of words, and are analyzed using powerful computer

28. See Antonin Scalia, *Judicial Deference to Administrative Interpretations of Law*, 1989 DUKE L.J. 511, 512 (1989) (“*Chevron* has proven a highly important decision—perhaps the most important in the field of administrative law since *Vermont Yankee* . . .”).

29. Michael McCarthy & Anne O’Keeffe, *Historical Perspective: What Are Corpora and How Have They Evolved?*, in THE ROUTLEDGE HANDBOOK OF CORPUS LINGUISTICS 3 (Anne O’Keeffe & Michael McCarthy eds., 2012).

30. TONY MCENERY, RICHARD XIAO & YUKIO TONO, CORPUS-BASED LANGUAGE STUDIES: AN ADVANCED RESOURCE BOOK 126 (2006) (“Any selection of texts is a sample. Whether 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 . . .”).

31. See *id.*; see also Randi Reppen, *Building a Corpus: What Are the Key Considerations?*, in THE ROUTLEDGE HANDBOOK OF CORPUS LINGUISTICS 31–32 (Anne O’Keeffe & Michael McCarthy eds., 2012) (discussing considerations of size, representativeness, collection methods, and annotation of corpora).

software.³² Computer analysis of corpora can provide “a new perspective on the familiar” by allowing analysis on a scale too vast to be done by hand.³³ It can also provide a more scientific, objective approach to determining linguistic meaning.³⁴ Because of this, corpus linguistic data can have significant advantages over traditional approaches to deciding language meaning in some contexts.

For example, corpus linguistics analysis can have decided advantages over simple appeals to dictionaries when it comes to deciding the “ordinary meaning” of a term.³⁵ The classic case goes like this: A judge, in resolving a case that requires her to determine the “ordinary meaning” of a word, will reason that she has found the ordinary meaning of term X because dictionary Y lists it as the first definition. But the problem with using dictionaries to make claims about the ordinary meaning of words is that “dictionaries do not tell us how words are commonly or ordinarily used.”³⁶ Most general dictionaries do not claim to rank the definitions they contain according to relevancy or frequency of use.³⁷ Therefore, appeals to dictionaries claiming that the first definition listed should have some sort of special weight are generally misguided, and these types of arguments have been aptly labeled the “sense-ranking fallacy.”³⁸ When a court’s holding rests on this type of reasoning, the conclusion can lack foundation.³⁹

32. Jane Evison, *What Are the Basics of Analysing a Corpus?*, in THE ROUTLEDGE HANDBOOK OF CORPUS LINGUISTICS 122 (Anne O’Keeffe & Michael McCarthy eds., 2012).

33. SUSAN HUNSTON, CORPORA IN APPLIED LINGUISTICS 3 (2002).

34. See Thomas R. Lee & Stephen C. Mouritsen, *Judging Ordinary Meaning*, 127 YALE L.J. 788, 828–29 (2018).

35. See generally 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).

36. J.M.W. v. T.I.Z. (*In re Adoption of Baby E.Z.*), 2011 UT 38, ¶ 98, 266 P.3d 702, 727 (Utah 2011) (Lee, J., concurring).

37. See, e.g., WEBSTER’S THIRD NEW INTERNATIONAL DICTIONARY 17a (2002) (“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.*”) (emphasis added); 1 THE OXFORD ENGLISH DICTIONARY xxxi (1st ed. 1933) (“[T]hat sense is placed first which was actually the earliest in the language: the others follow in the order in which they appear to have arisen.”).

38. Mouritsen, *supra* note 35, at 1926–38.

39. Yet the United States Supreme Court has employed this reasoning. For example, when determining the “ordinary” meaning of the phrase “carries a firearm” in *Muscarello v. United States*, 524 U.S. 125 (1998), Justice Breyer’s majority opinion used the following as primary support for its conclusion: “Consider first the word’s primary meaning. The Oxford

Further, language is in constant change, so dictionaries cannot claim to be either completely exhaustive or completely up to date in their coverage. It is potentially problematic, then, when a court's reasoning is "X cannot possibly mean Y because we can find no dictionary to support that definition."⁴⁰ Dictionaries are not written to make such determinations. But judges often use these types of dictionary arguments to bolster their conclusions.

These problems are just as apparent in cases where a court's job is not to decide the *correct* interpretation of a statute but to decide whether the interpretation advanced by one of the parties is a *permissible* or *reasonable* one – the inquiries required by *Chevron*. The problem of determining what *permissible* or *reasonable* means in the context of statutory interpretation is bad enough. This difficulty is compounded by the question of where judges should turn to make these determinations of reasonableness or permissibility. Judges often end up leaning even more heavily on their own linguistic intuition about whether a term can have a given meaning and then use dictionaries to support their personal intuition. This is problematic because people's own linguistic intuitions are often far from accurate when compared with actual data of language use.⁴¹ Claims about *reasonable* and *permissible* interpretations of statutory terms in some *Chevron* cases⁴² are on shaky ground when the court's

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')." *Muscarello*, 524 U.S. at 128 (original typefaces retained).

40. But Supreme Court Justices and Federal Circuit Court judges have done this too. See, e.g., *Marx v. Gen. Revenue Corp.*, 133 S. Ct. 1166, 1180 (2013) (Sotomayor, J., dissenting) ("[T]he majority does not cite even a single dictionary definition in support of that reading, despite the oft-cited principle that a definition widely reflected in dictionaries generally governs over other possible meanings."); *Gibson v. Parish*, No. 08-7103, 2010 WL 104553, at *974, *980 (10th Cir. Jan. 13, 2010) ("[Plaintiff] does not cite to any dictionary to support her argument, and we have found none adopting that definition.").

41. See Reppen, *supra* note 31, at 31 ("Unlike straightforward grammaticality judgments, when [people] are asked to reflect on language use, [their] recall and intuitions about language often are not accurate.").

42. This Comment does not argue for a general application of corpus linguistics in all types of *Chevron* cases. Rather, the type of case in which corpus linguistics has the most to offer is a case like *MCI*, where the question of deference hinges on the reasonable or permissible meaning(s) of a single term or phrase. While a discussion of other types of

main authority for its conclusion is a judge's linguistic intuition buttressed by a dictionary.

Common corpus linguistic methodologies can help avoid some of these problems. One such method is the use of concordance lines. Concordance analysis involves a computer-aided search of a database where the user searches for a single term or phrase. The search results are displayed as lines of text from the corpus that include the search term, usually in the center of a line or section of text.⁴³ Concordance analyses are useful in examining the "sense,"⁴⁴ or connotation of a word, in a given context.⁴⁵ The target word's semantic context, as evidenced by the concordance line, can help determine the appropriate sense in which the word is being used. While a concordance analysis, like a dictionary, cannot claim to be exhaustive or detail every possible sense of any given term (because it would be impossible to ever build a database that includes every instance of a word's use in the history of a language), it does allow for a much closer analysis of whether a given term can bear a certain meaning.

A related corpus method is the study of collocates. Collocates are words that co-occur in significant ways.⁴⁶ In essence, collocate analysis adapts the old adage that a person can know a *word* by looking at the company it keeps.⁴⁷ Many corpora allow examination

judicial deference to agency views is beyond the scope of this Comment, the problems discussed here are also apparent in cases applying what is known as *Seminole Rock* or *Auer* deference to agency interpretations of their own regulations. In such cases, a court is required to give controlling weight to agency interpretations of their own regulations unless the agency's interpretation is "plainly erroneous or inconsistent with the regulation." *Auer v. Robbins*, 519 U.S. 452, 461 (1997) (quoting *Robertson v. Methow Valley Citizens Council*, 490 U.S. 332, 359 (1989)).

43. See JOHN SINCLAIR, *CORPUS, CONCORDANCE, COLLOCATION* 32 (1991) ("A concordance is a collection of the occurrences of a word-form, each in its own textual environment. In its simplest form, it is an index. Each word-form is indexed, and a reference is given to the place of each occurrence in a text.")

44. "In linguistics a *word sense* is the meaning ascribed to a word in a given context." Alice Ruggeri & Luigi Di Caro, *Linguistic Affordances: Making Sense of Word Senses*, in *PROCEEDINGS AIC 2013: 1ST INTERNATIONAL WORKSHOP ON ARTIFICIAL INTELLIGENCE AND COGNITION* 136 (2013).

45. See Rosamund Moon, *What Can a Corpus Tell Us About Lexis?*, in *THE ROUTLEDGE HANDBOOK OF CORPUS LINGUISTICS* 204–05 (Anne O'Keeffe & Michael McCarthy eds., 2012).

46. See HUNSTON, *supra* note 33, at 68 ("Collocation is the tendency of words to be biased in the way they co-occur.")

47. See Lee & Mouritsen, *supra* note 34, at 832.

of the words that frequently appear in close proximity to a given search term. A collocate search will examine every use of the “node word” and give a statistical summary of words with which it frequently appears—i.e., the word’s collocates. Collocation analysis is extremely useful in identifying the semantic properties of a given word or phrase.⁴⁸

This Comment uses both methodologies discussed above—concordance and collocate analysis—to test Justice Scalia’s claims about the semantic range of *modify*. Both analyses suggest that Justice Scalia’s argument—that *modify* could not connote substantial or fundamental change—was incorrect from a semantic point of view.⁴⁹

IV. MCI TELECOMMUNICATIONS CORP. v. AT&T CO.

This brings us to the Supreme Court’s decision in *MCI Telecommunications Corp. v. AT&T Co.*⁵⁰ At first blush, the case seems relatively routine—a fight over whether the Federal Communication Commission’s construction of the statutory provision at issue qualified for deference under *Chevron*. But the case stands out because the majority relied heavily on dictionary arguments to reach its conclusions about the meaning of *modify*.

AT&T has a long history of government regulation, primarily due to its long-held monopoly over many types of telecommunications services in the United States.⁵¹ One of the statutes passed to regulate the industry was the Telecommunications Act of 1934.⁵² As amended, § 203(a) of the Act requires communications common carriers to file tariffs with the Federal Communications

48. See DOUGLAS BIBER, SUSAN CONRAD & RANDI REPPEN, CORPUS LINGUISTICS: INVESTIGATING LANGUAGE STRUCTURE AND USE 43–51 (1998).

49. It is probably a good time to note that sometimes even an extended corpus linguistic analysis can fail to demonstrate conclusively which sense of the word is being used in a given example. Some language is truly ambiguous, and even a corpus linguistic approach cannot solve the problem of *true* ambiguity. But often, ambiguous language has probative value to the issue being analyzed. This is clearly the case in *Chevron*’s Step One, where ambiguity is the threshold determination courts are required to make.

50. *MCI Telecomms. Corp. v. AT&T Co.*, 512 U.S. 218 (1994).

51. *Id.* at 220.

52. *Id.*

Commission (FCC).⁵³ The FCC, as the agency charged with implementing the Act, is authorized under § 203(b) to “modify any requirement” of the tariff filing provision.⁵⁴

The Telecommunications Act aimed, at least in part, to make sure that the rates being charged by AT&T for their long-distance services were reasonable and did not discriminate.⁵⁵ Eventually, as the market became more competitive, the FCC decided to relax the tariff filing requirement in some ways for non-dominant carriers.⁵⁶ MCI Telecommunications was one of those non-dominant carriers.⁵⁷ Eventually, the FCC made the tariff filing completely optional for nondominant carriers, and AT&T challenged the regulation.⁵⁸ The FCC claimed authority to make these changes under its power to “modify,” as described in § 203(b).⁵⁹ The D.C. Circuit disagreed,⁶⁰ and MCI petitioned for certiorari.

For the Supreme Court, the case was not a difficult one. The dispute, according to the majority, “turn[ed] on the meaning of the phrase ‘modify any requirement’ in § 203(b)(2).”⁶¹ The Court disagreed with the Commission’s contention that § 203(b) gave it “authority to make even basic and fundamental changes” to the tariff filing requirement.⁶² To support its conclusion, the Court first examined the etymology of the word *modify*.⁶³ It then cited multiple dictionary definitions of *modify*, all with the aim of demonstrating that the term signified only moderate change rather than major or fundamental change.⁶⁴ The Court noted that only a single dictionary—*Webster’s Third New International Dictionary*—allowed

53. 47 U.S.C. § 203(a) (2012) (“Every common carrier, except connecting carriers, shall, within such reasonable time as the Commission shall designate, file with the Commission and print and keep open for public inspection schedules showing all charges for itself and its connecting carriers for interstate and foreign wire or radio communication . . .”).

54. *Id.* § 203(b)(2) (“The Commission may, in its discretion and for good cause shown, modify any requirement made by or under the authority of this section . . .”).

55. *MCI*, 512 U.S. at 220.

56. *Id.* at 221.

57. *Id.*

58. *Id.* at 221–22.

59. *Id.* at 221.

60. *AT&T v. FCC*, 978 F.2d 727 (D.C. Cir. 1992).

61. *MCI*, 512 U.S. at 225.

62. *Id.*

63. *Id.*

64. *Id.* at 225–26.

modify to mean significant or major change.⁶⁵ For the majority, interpreting *modify* to connote both major change and minor change was ridiculous.

When the word “modify” has come to mean *both* “to change in some respects” *and* “to change fundamentally” it will in fact mean *neither* of those things. It will simply mean “to change,” and some adverb will have to be called into service to indicate the great or small degree of the change.⁶⁶

The Court continued by stating that “an elimination of the crucial provision of the statute for 40% of a major sector of the industry is much too extensive to be considered a ‘modification.’ What we have here . . . is a fundamental revision of the statute”⁶⁷ Because it found that the Commission’s interpretation went “beyond the meaning that the statute [could] bear,” the Court concluded that no deference was due to the agency’s interpretation of *modify* under *Chevron*.⁶⁸

V. CORPUS ANALYSIS OF *MODIFY*

To examine the Court’s claims about the reasonability and permissibility of the FCC’s construction of the verb *modify*, I examined the term using COHA, COCA, and NOW. To determine whether the FCC’s construction of the term was “a permissible construction of the statute,” I examined the term using both a concordance analysis and a collocation analysis.⁶⁹ The results show that the Court was correct in asserting that the most common meaning of *modify* is “to change moderately or in minor fashion.”⁷⁰ But the data also show that *modify* is sometimes used to connote serious or even fundamental change. Hence, under the standard established by *Chevron*, the Court likely should have accepted the FCC’s regulation as a “permissible” or “reasonable” construction of

65. *Id.* at 227.

66. *Id.*

67. *Id.* at 231.

68. *Id.* at 229.

69. Although it is likely corpus linguistics could have also resolved the case at *Chevron* Step One, this Comment challenges Justice Scalia’s analysis where it seems to have focused longest—at Step Two.

70. *MCI*, 512 U.S. at 225.

the statute, even though the Court did not believe it was the best possible construction.⁷¹

A. Concordance Analysis

I first used a concordance analysis to approach the question of *modify*'s meaning, examining data from both COHA and COCA. I analyzed two sets of concordance lines, each composed of 100 randomized samples returned by my search. The first data set was composed of concordance lines between the years 1930 and 1939. (As Justice Scalia points out in his opinion, "the most relevant time for determining [the] statutory term's meaning" is 1934, the year the Communications Act became law.⁷²) To determine how much, if any, the connotation of *modify* has changed since the 1930s, I also examined 100 randomized concordance lines from COCA from 1990 to 1994 (the years immediately surrounding the *MCI* decision). In both searches, I used simple syntax, searching for "modify.[v*]," which instructs the corpus to return all inflections of the verb *to modify*.

To determine which sense of *modify* was being used in each concordance, I examined the greater context of the concordance. Based on indicators like modifiers, synonyms, and other structural features, I then assigned each use of *modify* to one of three categories: (1) uses of *modify* that allow for only moderate change (*i.e.*, the definition of the word favored by the *MCI* majority opinion); (2) ambiguous uses (*i.e.*, uses that would allow for both a moderate only sense of *modify* and a substantial-change sense); and (3) uses where *modify* is employed to mean substantial, fundamental, or major change.

The data from both time periods were consistent and seemed to indicate that the Court was incorrect in asserting both (1) that *modify* was unambiguous and (2) that construing it to mean major or fundamental change was an impermissible reading of the statute. In the results from the 1930s, nearly one-quarter of the uses

71. See *supra* Part II. An important caveat: This Comment does not attempt to establish a definitive line marking what counts as "reasonable." That is a debate for another paper (or ten). Rather, this Comment assumes that a use that appears roughly one quarter of the time in general language-use data is likely a "reasonable" interpretation under the strong deference established under *Chevron*.

72. *MCI*, 512 U.S. at 228.

of *modify* either strongly suggested that it indicated substantial change or at least would allow such a definition. In the sample, 77% of the tokens examined indicated that *modify* suggested only moderate change. This constituted a strong majority of the uses of *modify*, and this result was not surprising. But 23% of the uses either allowed for a more expansive sense of *modify* or required it. Of that 23%, 18% of the uses were ambiguous, and would allow for either a moderate-scope connotation or a substantial-change connotation. The remaining 5% indicated strongly that *modify* connoted substantial, significant change. Examples of each type of use are shown below in *Figure 1*.

Figure 1. COHA sample concordance lines.

Sense	KWIC ⁷³
Moderate-only	“ . . . it is difficult <i>to change or even to <u>modify</u>.</i> ”
Ambiguous	“ . . . passive immunity has provided medical men with the means to <u>modify</u> <i>to a greater or a lesser extent such bacterial diseases as . . .</i> ”
Fundamental	“ . . . The House Ways & Means Committee managed to <u>modify</u> <i>the principle beyond all recognition . . .</i> ”

As can be seen, the context of each example provides clues to the intended meaning of *modify*. In the moderate-only sense, *modify* is set in contrast to the broader verb *to change*. As if to drive the point home, *modify* is preceded by *even*, which is used to stress the difference in scope between *change* and *modify*. The second example of *modify* is ambiguous, and the verb in this context could bear either sense. The phrase immediately following *modify* is most helpful in determining its possible scope; “to a greater or lesser extent” plainly suggests that *modify* could mean to change substantially *or* slightly. The last example strongly suggests that *modify* means substantial change—even change beyond recognition. If something can be *modified* to the extent that it is

73. “Key word in context” (KWIC).

unrecognizable, then clearly a substantial, fundamental change of some sort has taken place.

The data from COCA (1990–94) were similar. Twenty-eight percent of the uses of *modify* either strongly suggested that the term meant substantial change or were ambiguous and would have allowed for that meaning. Twenty percent of the tokens were ambiguous, 8% suggested that *modify* meant serious change, and the remaining 72% of the samples all suggested a moderate-only sense for *modify*. So although the percentage of the ambiguous/substantial uses of *modify* increased slightly from the 1930s, the overall proportion remained similar – just over a quarter of the uses of *modify* defy Justice Scalia’s claim about its permissible meaning. Examples from COCA are shown below in *Figure 2*.

Figure 2. COCA sample concordance lines.

Sense	KWIC
Moderate-only	“... replaying the tapes for the majors to <u>modify</u> and clarify any aspect of the interview . . .”
Ambiguous	“... students must first share their answers with each other – and possibly <u>modify</u> and/or discuss them – before their group is finished.”
Fundamental	“... and the State Board of Health to <u>modify</u> existing school posture examinations, <i>transforming them</i> . . .”

The first example suggests a moderate-only sense for *modify* by pairing it with the word *clarify*. Both words suggest that the “majors” are merely reviewing recordings of an interview to suggest any changes that should be made. The phrase “modify and clarify” essentially excludes the possibility that the “majors” will be able to fundamentally change the interviews they are reviewing. The second example is ambiguous. The students in the sentence will be given the chance to “modify and/or discuss” their answers to a question before they are finished working in a group. A student in this situation could “modify” his or her answer either by slightly

changing it or completely changing it. In this context *modify* could bear either meaning. The third example suggests that *modify* means some sort of substantial change, as the modifications made were sufficient to “transform” the examinations. *Transform* obviously suggests something much more substantial than slight change.

The data collected from the corpus and represented by the examples in *Figure 1* and *Figure 2* suggest that the Court’s claims about possible connotations of *modify* were incorrect. Although the clear majority of occurrences of *modify* did in fact suggest only minor change, in both time periods nearly one-quarter of the uses of *modify* permit (and sometimes *require*) connotations of major, fundamental change. This is certainly at odds with Justice Scalia’s statement that “*modify does not contemplate fundamental changes.*”⁷⁴ Under the highly deferential standard established by *Chevron*, a connotation established by nearly 25% of all uses of a word should qualify as a “permissible” construction of the word.

B. Collocate Analysis

I also analyzed the collocates of *modify*, in an analysis suggested in *MCI* by Justice Scalia himself. He suggested that if *modify* could mean both fundamental and minor change, then it had come to be essentially a synonym for the verb *to change*. We would know this, he hypothesized, if adverbs are pressed into service to describe the degree of change indicated by *modify*.⁷⁵ To examine this hypothesis, I looked at collocate data from three corpora: COHA, COCA, and NOW. The data from COHA were used to examine *modify* during the time when the Telecommunications Act was originally written, while the COCA and NOW data were taken from the period in which the Court decided *MCI*.

The search was again simple: I performed a search for adverbial collocates of [modify].[v*], set the minimum frequency at five, and sorted the results by Mutual Information (MI) score.⁷⁶ I also

74. *MCI*, 512 U.S. at 228 (emphasis added) (quotations omitted).

75. *Id.* at 227.

76. A Mutual Information (MI) score “expresses the extent to which observed frequency of co-occurrence differs from what we would expect (statistically speaking). In statistically pure terms this is a measure of the strength of association between words x and y.” HarperCollins Publishers, *A Guide to Statistics: T-Score and Mutual Information*, COLLINS

specified that I wanted only adverbs that occurred within two words on either side of the node word, *modify*. The point was to see whether the data showed adverbs of scope or degree as significant collocates of *modify*. If no adverbs of scope were employed among the top adverbial collocates of *modify*, it would suggest that *modify* connoted only one thing. But if adverbs of scope collocated frequently vis-à-vis other non-scope adverbs, as Justice Scalia theorized, it would suggest that *modify* has some sort of inherent ambiguity and adverbs are used to bring clarity. I performed the same search for each of the three corpora and examined the twenty most frequent collocates.⁷⁷

The data suggest that *modify* has undergone the very process suggested by the Court – to some degree, it has become a synonym of *to change*. Several adverbs of scope or degree collocated frequently with *modify*, including adverbs both expanding and contracting its scope. For example, in COHA, *radically*, *materially*, *profoundly*, *greatly*, and *seriously* are all in the top six collocates. These adverbs expand the scope of *modify* beyond what would be allowed if the MCI Court had been right – *i.e.*, if *modify* could only connote small, incremental, or measured change. The data suggest that even during the decade in which the Telecommunications Act was written (1930–39), *modify* was also used to suggest serious change. But what is most important is the use of adverbs at all. If, as the Court claimed, *modify* always and only connoted incremental change, then adverbs suggesting a limited scope would be redundant. And using adverbs that broadened *modify*'s meaning would be contradictory and nonsensical.

Looking at data from the time of the decision in *MCI*, the COCA data show a similar pattern, but with an important distinction: The collocates from COCA include words like *substantially* and *significantly*, which broaden the scope of *modify* and are similar to the collocates from COHA. But the COCA collocates also contain words like *slightly*, which do the exact opposite. *Slightly* narrows the scope or degree of *modify*. As mentioned above, if *modify* could

WORDBANK ONLINE (2008), https://wordbanks.harpercollins.co.uk/other_doc/statistics.html. Essentially, the higher the MI score, the higher a particular collocate's relevancy is to the node word's semantic properties.

77. I included the top seventeen collocates from COHA rather than the top twenty because only seventeen collocates occurred with a frequency of five or greater.

only ever connote moderate, incremental change, *slightly* would be redundant. Again, the most important takeaway is that adverbs of scope are being used with *modify* at all. That use suggests exactly what Justice Scalia claimed was impossible—that *modify* actually means something like *to change*, and people use adverbs to show how much or how little.⁷⁸

Finally, the data from NOW show a pattern similar to the data from COCA. Interestingly, *slightly* ranked even higher in NOW than in COCA.

To investigate how similar the meaning of *modify* is to the meaning of *to change*, I also analyzed the top twenty adverbial collocates of the verb *to change* during the same time periods. My search methodology was the same as the searches I performed for *modify*. The results showed patterning similar to the collocates of *modify*: adverbs of scope or degree such as *radically*, *subtly*, and *fundamentally* all appeared as top collocates for *change*. This held true across data from all three time periods and corpora.

The data seem to show that, to a degree, *modify* is a synonym for *to change* and is used frequently in combination with the same types of adverbs of scope and degree. This suggests, like results of the concordance analysis, that *modify* can indeed connote fundamental, major change. This in turn also suggests that the FCC's construction of the phrase "modify any requirement" was likely a reasonable construction of the statute. Under *Chevron*, that is enough, and the Court probably should have deferred to the agency's interpretation.

The corpus data suggest not only that *modify* is a near synonym for *change*, but also that *modify* can connote a change of substantial,

78. These observations about *modify* track similar observations made recently by the Michigan Supreme Court in *People v. Harris*, 885 N.W.2d 832 (Mich. 2016). In *Harris*, the Michigan Supreme Court considered the question of whether the statutory term *information* could include false statements made by police officers during an internal investigation. The court concluded—relying at least in part on data obtained from a corpus—that *information* includes both true and false statements. *Id.* at 842. The court noted that the term was frequently used in conjunction with adjectives that suggest the information may be either true or false. *Id.* at 839. "This," the court concluded, "strongly suggests that the unmodified word 'information' can describe either true or false statements. . . . Quite simply, 'information' in common parlance describes perceptions conveyed about the world around us, which may be true or false." *Id.* The patterning for *modify* here is similar and warrants a similar inference.

even fundamental scope. These conclusions could not have been reached by a simple appeal to dictionaries – as *MCI* clearly shows. Neither could these conclusions have been reliably reached by individual linguistic introspection. As we see, corpus linguistics can help judges reach these sorts of objective conclusions about language meaning in cases where the semantic properties of words can be outcome determinative.⁷⁹

VI. IMPLICATIONS

This Comment suggests that corpus linguistics can help judges determine the ordinary meaning of a term at issue in a *Chevron* case. Corpus linguistic analysis discourages arbitrary decision-making about linguistic meaning, adds transparency, and has the potential to bring greater uniformity to areas of law – like *Chevron* cases – where it is severely lacking.

79. Of course, none of this is to say that using corpus linguistics in legal interpretation has gone uncriticized. In their recent article, Lee and Mouritsen outline several major criticisms aimed at corpus linguistics in the law. See Lee & Mouritsen, *supra* note 34, at 865–76. First is the “proficiency critique,” which says “judges and lawyers should leave the linguistic analysis to professional linguists – meaning, in practice, to expert witness reports or testimony.” *Id.* at 865. While admitting that this criticism “has some bite” because “corpus linguistics is not ‘plug and play’ analysis,” Lee and Mouritsen note that this critique is more a word of caution to judges to acknowledge the limits of the corpus linguistics tool than it is a judgment that the tool shouldn’t be used in the first place. *Id.* at 866. The second critique Lee and Mouritsen note is the “propriety critique”: Judges should avoid conducting an independent investigation of facts that goes beyond the issues briefed by any of the parties. *Id.* at 868–69. But, as Lee and Mouritsen note, this criticism is misaimed: The relevant ethics rules aim at independent investigation of *adjudicative facts*, not independent investigation of linguistic meaning. *Id.* 869–70. A rule to the contrary would often outlaw many accepted interpretive practices, including some uses of dictionaries. *Id.* at 870. A third concern has been raised that corpus data may show a specific sense to be a more common iteration of a term in the real world, but the most factually common iteration of a term may not necessarily reflect the sense that lawmakers, or even the governing public, had in mind. *Id.* at 873–74. Lee and Mouritsen note that this does not deprive corpus data of their probative value. The “top-of-mind sense” of a term will often help determine the “ordinary public meaning” by indicating the sense with which the average member of the public would likely be most familiar. *Id.* And in cases where the corpus data indicate a *lack* of a dominant sense, these data will be probative as well, perhaps indicating that it’s time for judges to employ a canon of construction or some other rule to function as a tiebreaker. *Id.* at 875–76. Finally, Lee and Mouritsen note that corpus linguistics could be attacked as only applying to their chosen theories of interpretation: originalism and textualism. *Id.* at 876. However, they assert that even avowed “anti-textualists” have an interest in demonstrating that the text is indeterminate and therefore inconclusive. *Id.* at 877. This is not a comprehensive list of critiques that have been or will yet be leveled against the use of corpus linguistics, but it does present some of the more salient criticisms and some brief responses.

First, corpus linguistics can discourage arbitrariness in determining a term or phrase's meaning. Corpus data allow judges to make decisions on actual language use in the real world rather than on internal semantic intuition.

Next, corpus linguistics can bring added transparency and consistency to decisions about word meaning. If a corpus analysis is done correctly, it is replicable and can be examined by anyone willing to invest the time in an analysis. Judges who disagree about the semantic properties of a term or phrase will be able to debate the issue in a way that encourages objectivity. Some disagreements may remain, but corpus linguistics moves the debate from the inside of judges' minds into the open. This, as a simple rule-of-law value, is preferable to the alternative: decisions about the meaning of words that have no support but a dictionary definition and a judge's intuition.

Finally, corpus linguistics can help bring uniformity to the question of what "ambiguous" means in the *Chevron* context. The data for *modify* showed a rough 75/25 split: 75% of the time, *modify* connoted moderate change, while 25% of the time it allowed or required a more expansive sense. In the author's opinion, 25% is enough to suggest that a term is ambiguous, but that is beside the point. More importantly, corpus linguistics yields actual data. Such data can provide a settled background against which judges can more accurately draw a line defining ambiguity and, thereby, create and apply a more consistent standard.

Corpus linguistics is not a magic bullet and cannot solve every problem judges face when deciding *Chevron* cases. It can, however, provide a clear, objective way for judges to analyze questions of meaning in a way that reflects actual language usage.

*Jacob Crump**

* J.D., April 2018, J. Reuben Clark Law School, Brigham Young University.