1-1-2017

Reasonable Certainty & Corpus Linguistics: Judging Definiteness after Nautilus & Teva

Joseph S. Miller
Professor of Law *University of Georgia School of Law*, getmejoe@uga.edu

University of Georgia School of Law
Research Paper Series
Paper No. 2017-11

Reasonable Certainty & Corpus Linguistics: Judging Definiteness After *Nautilus* & *Teva*

Joseph Scott Miller∗

I. INTRODUCTION

> Reasonableness is not a concept of definite and unchanging content. Its meaning necessarily varies in the different fields of the law, because it is used as a convenient summary of the dominant considerations which control in the application of legal doctrines.
>
> *United States v. Trenton Potteries Co.*[1]

> An invention must be capable of accurate definition, and it must be accurately defined, to be patentable.
>
> *United Carbon Co. v. Binney & Smith Co.*[2]

The Patent Act requires that a utility patent "conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the inventor or a joint inventor regards as the invention."[3] The Patent Act has included this requirement, known as claim definiteness,[4] since 1870.[5] The requirement's rigors are now in ferment,

---

∗ Professor, University of Georgia School of Law. My thanks for helpful feedback to participants at the University of New Hampshire Law School's Sixth Annual IP Scholars' Roundtable, the Oregon Patent Law Association's Fall 2016 meeting, the UGA Law School faculty, and Greg Castanias, Suzanne Michel, Janice Mueller, Christian Turner, and Ryan Vacca.

1. 273 U.S. 392, 397 (1927).

2. 317 U.S. 228, 237 (1942). *See also* Harries v. Air King Prods. Co., 183 F.2d 158, 160 (2d Cir. 1950) (L. Hand, J.) ("A patent must be a certain guide; not a congeries of pregnant suggestions.").

3. 35 U.S.C. § 112(b) (2012). Before the 2011 enactment of the Leahy-Smith America Invents Act, Pub. L. 112-29, 125 Stat. 284 (2011), the same requirement, in nearly identical words, was found at ¶ 2 of § 112. *See* 35 U.S.C. § 112, ¶ 2 (2006) ("The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.").

4. *See* JANICE M. MUELLER, PATENT LAW 91 (5th ed. 2016) ("This statutory edict is known to patent lawyers as the claim **definiteness** requirement (although the statute does not use the word 'definite').").

having recently undergone transformative Supreme Court analysis twice in the space of nine months.[6] The last time the Supreme Court had considered claim definiteness on the merits, before its 2014 and 2015 decisions, was 1946.[7]

In *Nautilus*, rejecting the Federal Circuit's insoluble-ambiguity test as too tolerant of imprecise claim language,[8] the Supreme Court held that "a patent is invalid for indefiniteness if its claims, read in light of the specification delineating the patent, and the prosecution history, fail to inform, with reasonable certainty, those skilled in the art about the scope of the invention."[9] This is, as the Court described it, the "proper reading of the statute's *clarity and precision* demand."[10]

In *Teva*, again rejecting the Federal Circuit's approach, the Supreme Court held that resolving a dispute about the best way to construe a word or phrase in a patent claim can involve questions of fact, as well as of law.[11] As a result, when reviewing a trial court's resolution of a claim

---

5. Patent Act of 1870, ch. 230, § 26, 16 Stat. 198, 201 (1870) (requiring that "before any inventor or discoverer shall receive a patent for his invention or discovery . . . he shall particularly point out and distinctly claim the part, improvement, or combination which he claims as his invention or discovery").

6. Teva Pharm. USA, Inc. v. Sandoz, Inc., 135 S. Ct. 831 (2015); Nautilus, Inc. v. Biosig Instruments, Inc., 134 S. Ct. 2120 (2014).

7. Halliburton Oil Well Cementing Co. v. Walker, 329 U.S. 1, 9 (1946) (invalidating a claim on a machine for measuring the distance from a well top to the oil surface below on the ground that "[t]he language of the claim . . . describes th[e] most crucial element in the 'new' combination in terms of what it will do rather than in terms of its own physical characteristics or its arrangement in the new combination apparatus. We have held that a claim with such a description of a product is invalid as a violation of Rev. Stat. s 4888."). A subsequent case, *Faulkner v. Gibbs*, had appeared as if it might occasion further exploration of the analysis from *Halliburton Oil*, but after briefing and argument the Court thought otherwise. 338 U.S. 267, 267 (1949) (per curiam) ("The record, briefs and arguments of counsel lead us to the view that *Halliburton* . . . is inapposite.").

8. *Nautilus*, 134 S. Ct. at 2130 ("To tolerate imprecision just short of that rendering a claim 'insolubly ambiguous' would diminish the definiteness requirement's public-notice function and foster the innovation-discouraging 'zone of uncertainty' against which this Court has warned.") (quoting United Carbon Co. v. Binney & Smith Co., 317 U.S. 228, 236 (1942)). The Federal Circuit's "insolubly ambiguous" test arose no later than 2001, with the decision in Exxon Research & Eng'g Co. v. United States, 265 F.3d 1371, 1375 (Fed. Cir. 2001) ("If a claim is insolubly ambiguous, and no narrowing construction can properly be adopted, we have held the claim indefinite. If the meaning of the claim is discernible, even though the task may be formidable and the conclusion may be one over which reasonable persons will disagree, we have held the claim sufficiently clear to avoid invalidity on indefiniteness grounds.").

9. *Nautilus*, 134 S. Ct. at 2124.

10. *Id.* (emphasis added). On remand, the Federal Circuit concluded, as it had before, that the asserted claims were not invalid for indefiniteness. Biosig Instruments, Inc. v. Nautilus, Inc., 783 F.3d 1374, 1376 (Fed. Cir.) ("On remand, we maintain our reversal of the district court's determination that Biosig's patent claims are indefinite."), *cert. denied*, 136 S. Ct. 569 (2015).

11. *Teva*, 135 S. Ct. at 838 ("While we held in *Markman* that the ultimate issue of the proper construction of a claim should be treated as a question of law, we also recognized that in patent construction, subsidiary factfinding is sometimes necessary. Indeed, we referred to claim

construction dispute, an appeals court must apply the standard of review for factual findings, setting them aside only if they are "clearly erroneous."[12] The problem with the claim term at issue in *Teva*—"molecular weight"—was, importantly, purported indefiniteness: accused infringer Sandoz "argued . . . that, in the context of th[e asserted] patent claim, the term 'molecular weight' might mean any one of three different things."[13] Depending on which meaning was established, infringement may or may not occur, presenting just the type of claim scope uncertainty condemned in *Nautilus*. And the definiteness question turned on disputed facts: "[T]his case provides a perfect example of the factfinding that sometimes underlies claim construction."[14] The trial court received testimony from conflicting experts, credited the patentee's expert over the accused infringer's, and made a factual finding about how the person of ordinary skill in the art would have understood the term "molecular weight"; that finding, in turn, informed the trial court's construction of the term in the context of the asserted patent.[15] The appeals court rejected the trial court's finding without analyzing whether

---

construction as a practice with 'evidentiary underpinnings,' a practice that 'falls somewhere between a pristine legal standard and a simple historical fact.' We added that sometimes courts may have to make 'credibility judgments' about witnesses. In other words, we recognized that courts may have to resolve subsidiary factual disputes.") (quoting Markman v. Westview Instruments, Inc., 517 U.S. 370, 378, 388, 390, 389 (1996)).

Teva thus put to rest an issue that had divided the Federal Circuit for almost twenty years. In *Markman v. Westview Instruments, Inc.*, the Supreme Court rejected the patentee's contention that it had a Seventh Amendment right to a jury verdict on the claim construction dispute in the case. 517 U.S. 370, 372 (1996) ("We hold that the construction of a patent, including terms of art within its claim, is exclusively within the province of the court."). In the wake of that decision, the Federal Circuit reviewed claim construction decisions *de novo* on appeal. *See* MUELLER, *supra* note 4, at 645–59 (discussing this line of cases). But the Federal Circuit remained internally divided on the proper standard of review for claim construction, leading to two split *en banc* cases on the practice in the years between *Markman* and *Teva*. *See* Lighting Ballast Control LLC v. Philips Elecs. N. Am. Corp., 744 F.3d 1272 (Fed. Cir. 2014) (en banc) (reaffirming *de novo* review of claim construction by a 6-4 vote); Cybor Corp. v. FAS Techs., Inc., 138 F.3d 1448 (Fed. Cir. 1998) (en banc) (reaffirming *de novo* review of claim construction by 9-3 vote). *Teva*'s holding on the factual underpinnings of claim construction is still rippling through the patent system. That process is a slow one. *See* Jason Rantanen, *Teva, Nautilus, and Change Without Change*, 18 STAN. TECH. L. REV. 430, 432 (2015) ("It seems to be business as usual at the Federal Circuit.").

12. *Teva*, 135 S. Ct. at 836; *see also* FED. R. CIV. P. 52(a)(6) ("Findings of fact, whether based on oral or other evidence, must not be set aside unless clearly erroneous, and the reviewing court must give due regard to the trial court's opportunity to judge the witnesses' credibility.").

13. *Teva*, 135 S. Ct. at 836. The trial court had concluded that the claim was not fatally indefinite. *Id.* The Federal Circuit disagreed, striking down the claim as indefinite. *Id.* The procedural facet of the case the Supreme Court squarely rejected was that "the Federal Circuit reviewed *de novo* all aspects of the District Court's claim construction, including the District Court's determination of subsidiary facts." *Id.*

14. *Id.* at 840.

15. *Id.* at 842–43.

it was clearly erroneous, and "in failing to do so, the Federal Circuit was wrong."[16]

The Federal Circuit has also recently tackled a claim-definiteness question in a different part of the patent system, reviewing the way the Patent Office implements the clarity requirement during the initial examination of a patent application.[17] One key difference between patent examination and patent litigation is that, in the former process, an applicant can fix a legal defect in imprecise claim language simply by amending that claim language.[18] By contrast, in an enforcement case such as *Nautilus* and *Teva*, the terms in an issued patent claim remain fixed: However badly bungled the claim language may be, a court cannot rewrite the claim to save it.[19] In *In re Packard*, a case of first impression handed down about a month before the Supreme Court's *Nautilus* decision, the Federal Circuit affirmed that,

> when the USPTO has initially issued a well-grounded rejection that

---

16. *Id.* at 843. On remand, the Federal Circuit concluded, as it had before, that the asserted claims were invalid for indefiniteness. Teva Pharm. USA, Inc. v. Sandoz, Inc., 789 F.3d 1335, 1338 (Fed. Cir. 2015).

17. *See* 35 U.S.C. § 131 (2012) ("The Director shall cause an examination to be made of the application and the alleged new invention; and if on such examination it appears that the applicant is entitled to a patent under the law, the Director shall issue a patent therefor.").

18. *See* Halliburton Energy Servs., Inc. v. M–I LLC, 514 F.3d 1244, 1255 (Fed. Cir. 2008) ("We note that the patent drafter is in the best position to resolve the ambiguity in the patent claims, and it is highly desirable that patent examiners demand that applicants do so in appropriate circumstances so that the patent can be amended during prosecution rather than attempting to resolve the ambiguity in litigation.").

19. *See* Chef Am., Inc. v. Lamb-Weston, Inc., 358 F.3d 1371, 1374 (Fed. Cir. 2004) ("This court . . . repeatedly and consistently has recognized that courts may not redraft claims, whether to make them operable or to sustain their validity."); Rhine v. Casio, Inc., 183 F.3d 1342, 1345 (Fed. Cir. 1999) ("We . . . have consistently employed the caveat, 'if possible,' to our instruction that claims should be construed to sustain their validity. We have also admonished against judicial rewriting of claims to preserve validity.") (citations omitted). As the Supreme Court cautioned 130 years ago,

> > [s]ome persons seem to suppose that a claim in a patent is like a nose of wax, which may be turned and twisted in any direction, by merely referring to the specification, so as to make it include something more than, or something different from, what its words express. The context may, undoubtedly, be resorted to, and often is resorted to, for the purpose of better understanding the meaning of the claim; but not for the purpose of changing it, and making it different from what it is. The claim is a statutory requirement, prescribed for the very purpose of making the patentee define precisely what his invention is; and it is unjust to the public, as well as an evasion of the law, to construe it in a manner different from the plain import of its terms.

White v. Dunbar, 119 U.S. 47, 51–52 (1886); *see also* Cimiotti Unhairing Co. v. Am. Fur Ref. Co., 198 U.S. 399, 410 (1905) ("In making his claim the inventor is at liberty to choose his own form of expression, and while the courts may construe the same in view of the specifications and the state of the art, they may not add to or detract from the claim."). More concisely: "The claim is the measure of the grant." Universal Oil Prods. Co. v. Globe Oil & Ref. Co., 322 U.S. 471, 484 (1944).

> identifies ways in which language in a claim is ambiguous, vague, incoherent, opaque, or otherwise unclear in describing and defining the claimed invention, and thereafter the applicant fails to provide a satisfactory response, the USPTO can properly reject the claim as failing to meet the statutory [definiteness] requirements of § 112(b).[20]

A "satisfactory response" from the would-be patentee can be "a modification of the language identified as unclear, a separate definition of the unclear language, or, in an appropriate case, a persuasive explanation for the record of why the language at issue is not actually unclear."[21] In explaining its approach, the Federal Circuit couched the definiteness requirement in terms remarkably similar to those the Supreme Court would later use in *Nautilus*: "[T]his requirement is not a demand for unreasonable precision. . . . Rather, how much clarity is required necessarily invokes some *standard of reasonable precision* in the use of language in the context of the circumstances."[22]

Nautilus rejects the lax definiteness test the Federal Circuit had used since at least 2001, instead requiring that a claim provide reasonably certain notice.[23] *Teva* rejects the claim construction plenary-review paradigm the Federal Circuit had used since 1996, instead treating claim construction as grounded as much in technological fact as in documentary law.[24] *Packard* fully endorses a definiteness inquiry at the Patent Office that demands more of applicants who proffer seemingly unclear claims.[25]

As the Patent Office, the courts, patent applicants, and litigants adjust to this new reality, a critical question looms: What constitutes the requisite "*reasonable* certainty" that *Nautilus* demands? Put differently, *how*, exactly, is one to establish—in the Patent Office, or in the courts—whether a disputed claim term provides *reasonably* certain notice? After all, the term "'[r]easonable,' the most ubiquitous legal adjective, is not self-defining."[26] The Federal Circuit, in its remand decision in *Nautilus*,

---

20. 751 F.3d 1307, 1311 (Fed. Cir. 2014), *cert. denied*, 135 S. Ct. 2310 (2015).

21. *Id.*

22. *Id.* at 1313 (emphasis added). The *Packard* court's reference to context is also reminiscent of the court's conclusion, in an earlier era, that definiteness is a function of the technological art to which the invention pertains. *See* Orthokinetics, Inc. v. Safety Travel Chairs, Inc., 806 F.2d 1565, 1568, 1576 (Fed. Cir. 1986) (upholding the definiteness of "[t]he phrase 'so dimensioned'" in a claim reciting a pediatric wheelchair designed to be put in a car, on the ground that the phrase "is as accurate as the subject matter permits, automobiles being of various sizes").

23. *See* Nautilus, Inc. v. Biosig Instruments, Inc., 134 S. Ct. 2120, 2129–30 (2014).

24. *See* Teva Pharm. USA, Inc. v. Sandoz, Inc., 135 S. Ct. 831, 835 (2015).

25. *See Packard*, 751 F.3d at 1313–14.

26. Robert C. Ellickson, *Cities and Homeowners Associations*, 130 U. PA. L. REV. 1519, 1530 (1982). As Professor David Zaring recently observed, in analyzing the best way to conceptualize

recognized this breadth, observing that "[r]easonableness is the core of much of the common law, and 'reasonable certainty' has been defined in broad spectra of the law."[27]

Reasonableness is a function of context. The context here is marking claim scope, using well-chosen, well-arranged words and phrases to establish a boundary that people skilled in the relevant art can identify and thus respect (or fairly be held to account if they trespass). *Nautilus* mediates between dysfunctional extremes—a demand for perfect clarity that language cannot achieve, and a disregard for public notice that the statute forbids—using *reasonable* certainty.[28] This mediating role is a conventional one for legal reasonableness.[29] In addition, "patent law is

---

judicial review of administrative-agency action,

> [r]easonableness works all over the legal system. It undergirds the law of negligence, which depends upon the "reasonable person" standard; the Fourth Amendment, which prohibits "unreasonable searches and seizures"; antitrust law, which turns ever more increasingly on a "rule of reason" used to evaluate restraints upon trade; securities law; contract law; and the list could go on.

David Zaring, *Rule by Reasonableness*, 63 ADMIN. L. REV. 525, 527 (2011); *see also* George P. Fletcher, *The Right and the Reasonable*, 98 HARV. L. REV. 949, 949 (1985) ("W[e] lawyers should listen to the way we talk. . . . One of the most striking particularities of our discourse is its pervasive reliance on the term 'reasonable.' We routinely refer to reasonable time, reasonable delay, reasonable reliance, and reasonable care. In criminal law, we talk incessantly of reasonable provocation, reasonable mistake, reasonable force, and reasonable risk. Within these idioms pulse the sensibilities of the reasonable person. For all the supposed concreteness of the common law, we can hardly function without this hypothetical figure at the center of legal debate. We cannot even begin to argue about most issues of responsibility and liability without first asking what a hypothetical reasonable person would do under the circumstances.") (citation omitted).

27. Biosig Instruments, Inc. v. Nautilus, Inc., 783 F.3d 1374, 1379 (Fed. Cir.), *cert. denied*, 136 S. Ct. 569 (2015).

28. *See Nautilus*, 134 S. Ct. at 2128–29 (discussing these improper poles); *see also id.* at 2129 ("To determine the proper office of the definiteness command, therefore, we must reconcile concerns that tug in opposite directions.").

29. "Reasonableness can thus be conceived as [a] quest for a *practical equilibrium*, in an attempt to bring into balance different normative possibilities, measures, and arguments in relation to different circumstances." REASONABLENESS AND LAW xi (Giorgio Bongiovanni et al. eds., 2009); *see also* Neil MacCormick, *Reasonableness and Objectivity*, 74 NOTRE DAME L. REV. 1575, 1579 (1999) ("The reasonable person has the virtue of prudentia and uses this in action. It is a virtue that is incompatible with fanaticism or apathy, but holds a mean between these, as it does between excessive caution and excessive indifference to risk."); Benjamin C. Zipursky, *Reasonableness in and out of Negligence Law*, 163 U. PA. L. REV. 2131, 2138 (2015) ("Frequently in the law, reasonableness is utilized to qualify prices, rates, and costs. It is often used to qualify 'time,' as in 'within a reasonable time.' In all of these cases, 'reasonable' is roughly the opposite of 'excessive.'") (footnote omitted). Relatedly, we can use "reasonable" to highlight that a characteristic exists along a continuum, rather than only at one of two extremes:

> Not only do the[] qualifiers ["reasonable" and "reasonably"] ensure that it is a moderate level of the quality being designated, they also ensure that the one applying the law (be it legal actor or judge) is being guided in a manner that requires the exercise of judgment, not simply the identification of a clear-cut attribute.

Id. at 2146.

'a field where so much depends upon familiarity with specific scientific problems and principles not usually contained in the general storehouse of knowledge and experience.'"[30] As a result, to determine whether a patentee has marked out a reasonably certain boundary, one must have a firm grasp of the full range of language usage options that were available to a claim drafter in the relevant technological art at the time the patent in question was filed. For unless one establishes the usage options from among which such artisans demonstrably *could have* chosen as means to the end of "inform[ing] those skilled in the art about the scope of the invention with reasonable certainty,"[31] one is not in a position to judge fairly the propriety of the means a given patentee actually *did* choose to that end. The upshot of the transformation *Nautilus* and *Teva* have wrought, then, is that we need much more usage-facts grist for the definiteness-inquiry mill, relative to current practice.

This paper examines new ways the Patent Office and the courts can reliably obtain, organize, and bring to bear a robust, orderly factual record on usage patterns in a given technological art at the time a patent is, or was, filed. Usage patterns in prior art English-language[32] patents, the focus of this paper, would provide the most highly relevant data. Treating all prior art U.S. patents (or perhaps all prior art English language patents) as texts in a single corpus for computer-based analysis, the patent system can use the tools of corpus linguistics—"the aim [of which is] the analysis and description of *language use, as realised in text(s)*"[33]—to establish an art's prevailing usage patterns as of a given date. By processing a large collection of texts digitally, corpus linguistics simply lets us achieve a much-needed level of rigor and reliability in our conclusions about language usage—conclusions of the sort we make informally, and virtually instantaneously, every day as native speakers of a language.[34] The rigorously demonstrable usage

30. Teva Pharm. USA, Inc. v. Sandoz, Inc., 135 S. Ct. 831, 838 (2015) (quoting Graver Tank & Mfg. Co. v. Linde Air Prods. Co., 339 U.S. 605, 610 (1950)).

31. *Nautilus*, 134 S. Ct. at 2129.

32. Applications for U.S. utility patents must be in English. 37 C.F.R. § 1.52(b)(1)(ii), (d)(1) (2016).

33. Elena Tognini Bonelli, *Theoretical Overview of the Evolution of Corpus Linguistics*, *in* THE ROUTLEDGE HANDBOOK OF CORPUS LINGUISTICS 14, 18–19 (Anne O'Keefe & Michael McCarthy eds., 2010); *see also* TONY MCENERY & ANDREW HARDIE, CORPUS LINGUISTICS: METHOD, THEORY AND PRACTICE 1–2 (2012) ("We could reasonably define corpus linguistics as dealing with some set of machine-readable texts which is deemed an appropriate basis on which to study a specific set of research questions. The set of texts or *corpus* dealt with is usually of a size which defies analysis by hand and eye alone within any reasonable timeframe.").

34. Utah Supreme Court Associate Chief Justice Thomas Lee recently explained this in a concurring opinion in a case that turned on the best construction of a state criminal statute:

patterns established by corpus linguistics, however, can provide the proper evidentiary foundation for judging claim definiteness objectively.

Because patent examination and patent litigation shape and process inputs differently, with different consequences, one needs to consider these settings individually. This paper does that. What is true across the board is that the patentee is ultimately responsible for claim definiteness as the author of the claim language, both as a claim is presented in the originally filed application[35] and as a claim is amended in the course of patent examination.[36] The patentee's responsibility for the claim language has important implications for the process of proof in resolving a claim definiteness challenge, as the Federal Circuit's *Packard* decision indicates.[37] To point to the most salient analogy in recent Supreme Court patent jurisprudence: The patentee's ultimate responsibility for the chosen claim language shaped the new proof process for nonliteral infringement that the Supreme Court created in *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co.*[38] *Festo*'s focus on the patentee as the actor most accountable for chosen claim language echoes strongly in the Court's analysis of claim definiteness in *Nautilus*. To reliably implement the new *Nautilus* "reasonable certainty" standard, then, it is best to begin with *Festo*.

> In this age of information, we have ready access to means for testing our resolution of linguistic ambiguity. Instead of just relying on the limited capacities of the dictionary or our memory, we can access large bodies of real-world language to see how particular words or phrases are actually used in written or spoken English. Linguists have a name for this kind of analysis; it is known as *corpus linguistics*.
>
> The fancy Latin name makes this enterprise seem esoteric and daunting. It is not. We all engage in it even if we don't attach the technical label to it. A *corpus* is a body, and *corpus linguistics* analysis is no more than a study of language employing a body of language. When we communicate using words we naturally access a large corpus—the body of language we have been exposed to during our lifetimes—to decode the groups of letters or sounds we encounter. The most basic corpus linguistics analysis involves our split-second effort to access the body of language in our heads in our ongoing attempt to decode words or phrases we may be uncertain of. We all do that repeatedly every day.

State v. Rasabout, 2015 UT 72, ¶¶ 57–58, 356 P.3d 1258 (Lee, Associate C.J., concurring in part and concurring in the judgment) (footnote omitted).

35. *See* 35 U.S.C. § 112(b) (2012); 37 C.F.R. §§ 1.51(b)(1), 1.75 (2016).

36. *See* 35 U.S.C. § 132(a) (2012); 37 C.F.R. §§ 1.111(c), 1.121(c) (2016).

37. *See In re* Packard, 751 F.3d 1307, 1313 (Fed. Cir. 2014) ("Given the role of the applicant in the process, it is a reasonable implementation of the examination responsibility, as applied to § 112(b), for the USPTO, upon providing the applicant a well-grounded identification of clarity problems, to demand persuasive responses on pain of rejection. That approach decides this case, because Mr. Packard did not offer a satisfactory response to well-grounded indefiniteness rejections in this case.").

38. *See* 535 U.S. 722, 740 (2002) ("The patentee, *as the author of the claim language*, may be expected to draft claims encompassing readily known equivalents.") (emphasis added).

II. SCOUTING THE NEW TERRAIN

To establish the scope of one's patent claim with reasonable certainty, the patent writer must both select apt terms and arrange those terms well. The quality of selection and arrangement depends, in part, on the options the patent writer has available in the commonly accepted terminology in one's art at the time one writes the claims and the disclosure that supports them.[39] In this sense, the stock of familiar vocabulary and common usage in art *A* at time *T*—a matter of fact, as *Teva* establishes[40]—is the backdrop for one's claim-drafting choices. Those commonly accepted words and phrases are also the point of departure for considering whether a gap between the nature of one's new invention and familiar terms' settled meanings requires one to craft a bespoke definition of one's own for a given claim term. A semantic gap of this sort, if it occurs, is hardly surprising: "The dictionary does not always keep abreast of the inventor. It cannot. Things are not made for the sake of words, but words for things."[41] This freedom to provide an explicit definition for a claim term—to be, as the cases say, one's "own lexicographer"[42]—is open to *all* patentees.

To claim, then, is to choose. The patentee chooses for the claim those words and phrases that best establish the full scope of the invention. The patentee further chooses whether to rely on the ordinary meaning that is known in the art or, instead, to specially define any of

---

39. Another key determinant, of course, is the explanatory support the patent writer provides in the disclosure that supports the claims—disclosure that can set forth express definitions for claim terms, working or prophetic examples of the claimed invention, and drawings. *See* MUELLER, *supra* note 4, at 91–92, 151–58 (describing these features of a patent's supporting disclosure, or "specification"). The PTO's regulation on the basic content of a patent document, including the disclosure that supports the claims, is set forth at 37 C.F.R. § 1.51(b) (2016).

40. *See* Teva Pharm. USA, Inc. v. Sandoz, Inc., 135 S. Ct. 831, 842–43 (2015) (discussing the trial record on the meaning of "molecular weight" in a given piece of art at a given time).

41. Autogiro Co. of Am. v. United States, 384 F.2d 391, 397 (Ct. Cl. 1967).

42. *See* Nautilus, Inc. v. Biosig Instruments, Inc., 134 S. Ct. 2120, 2129 (2014) (quoting Hormone Research Found., Inc. v. Genentech, Inc., 904 F.2d 1558, 1563 (Fed. Cir. 1990)); Vitronics Corp. v. Conceptronic, Inc., 90 F.3d 1576, 1582 (Fed. Cir. 1996) ("Although words in a claim are generally given their ordinary and customary meaning, a patentee may choose to be his own lexicographer and use terms in a manner other than their ordinary meaning . . . ."); *Autogiro*, 384 F.2d at 397 ("To overcome this lag, patent law allows the inventor to be his own lexicographer."). The judicial phrase, if not the principle, seems to originate from a 1921 decision of the U.S. Court of Appeals for the Seventh Circuit. *See* Advance Rumley Co. v. John Lauson Mfg. Co., 275 F. 249, 251 (7th Cir. 1921) ("This court has frequently and consistently recognized the patentee's right to be his own lexicographer . . . ."). "Lexicographer" is simply a posh word for "an author or compiler of a dictionary." WEBSTER'S THIRD NEW INT'L DICTIONARY 1301 (1966). Importantly, "[i]n such cases, the inventor's lexicography governs." Phillips v. AWH Corp., 415 F.3d 1303, 1316 (Fed. Cir. 2005) (en banc).

those claim words or phrases, setting forth the definition(s) in the supporting written disclosure. (Of course, a patentee who mistakenly believes a given term to have a settled meaning to the ordinary artisan may fail to provide a bespoke definition when, in truth, the ordinary artisan would need a definition to take the meaning the patentee intends.) And the patentee's choices are consequential. As the Supreme Court made clear more than a decade ago in *Festo*, the patent system requires the patentee, *not* the broader public, to shoulder the cost of uncertainty that arises from those claim drafting choices.[43] This basic demand on the patentee, established in *Festo*, frames the prudent reading of *Nautilus* and *Teva*.

A. The Festo *Spirit*

Patent infringement is a type of trespass: If one makes, uses, sells, offers to sell, or imports the invention claimed in a utility patent, one directly infringes that claim.[44] Importantly, as a matter of judge-made law since the 1850s, one can infringe a claim even if the accused item does not meet every claim limitation literally. So long as the limitation that is missing literally is present equivalently, *i.e.*, by way of an insubstantially different substitute, the accused item triggers liability.[45]

In *Festo*, its most recent decision on equivalent infringement, the Supreme Court justified this departure from the strictures of literal claim scope—and the increased uncertainty that this departure inevitably creates—as the unavoidable price we pay for making patents sufficiently attractive rewards to induce inventors to apply for them. Taking as a given that a patent "is a property right," and thus, "like any property right, its boundaries should be clear," the Court could not help but acknowledge that, "[u]nfortunately, the nature of language makes it impossible to capture the essence of a thing in a patent application."[46] At the same time, "[i]f patents were always interpreted by their literal terms [alone], their value would be greatly diminished."[47] The solution has long been, the Court held, to extend the patentee's exclusion right: "The

---

43. *See* Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., 535 U.S. 722, 740 (2002) (noting that the burden is on the patentee "to draft claims encompassing readily known equivalents").

44. *See* 35 U.S.C. § 271(a) (2012) (defining utility-patent infringement); MUELLER, *supra* note 4, at 661–63 (describing literal patent infringement).

45. *See* MUELLER, *supra* note 4, at 663–72 (describing infringement under "the doctrine of equivalents," as well as the insubstantial-difference standard of equivalency).

46. *Festo*, 535 U.S. at 730–31.

47. *Id.* at 731.

scope of a patent is not limited to its literal terms but instead embraces all equivalents to the claims described."[48] That the "doctrine of equivalents renders the scope of patents less certain," while regrettable, is "the price of ensuring the appropriate incentives for innovation."[49]

Equivalent infringement is not, however, a license to engage in a freewheeling infringement analysis that is heedless of the particular way the patentee wrote the claim and navigated Patent Office review. First, "the doctrine of equivalents must be applied to individual elements of the claim, not to the invention as a whole."[50] Every claim limitation the patentee chose to recite is significant. Second, if the patentee obtained the claim from the Patent Office only by narrowing its initial scope to overcome a legal hurdle, that narrowing amendment stops the patentee from relying on the doctrine of equivalents to recapture the sacrificed scope as if the amendment had not occurred. This constraint on equivalent infringement, known as "prosecution history estoppel," was the very constraint at issue in *Festo*.[51] Specifically, the Court confronted a challenge to the Federal Circuit's conclusion "that when estoppel arises, it bars suit against every equivalent to the amended claim element."[52] The Federal Circuit, for its part, had tired of the unpredictability of crafting individually tailored estoppels based on the particulars of each case, concluding that its longstanding "case-by-case approach ha[d] proved unworkable."[53]

The Supreme Court rejected the Federal Circuit's complete-bar rule, on the ground that a complete bar imprudently demands perfection in language, in just the same way that a literal-infringement-only approach would demand perfection. The very existence of the doctrine of equivalents recognizes language's imperfection, and so, in turn, must prosecution history estoppel recognize it. At the same time, it is true that this estoppel, even as it fully respects the shortcomings of language, targets a case where the patentee had *not* lacked for the requisite phrasing but instead chose narrower words advisedly:

> Where the original application once embraced the purported equivalent but the patentee narrowed his claims to obtain the patent or to protect its validity, the patentee cannot assert that he lacked the words to describe the subject matter in question. . . . In that instance the

48. *Id.* at 732 (citing Winans v. Denmead, 56 U.S. (15 How.) 330, 347 (1854)).
49. *Id.*
50. Warner-Jenkinson Co. v. Hilton Davis Chem. Co., 520 U.S. 17, 29 (1997).
51. *See* 535 U.S. at 733–34.
52. *Id.* at 727.
53. *Id.* at 730.

> prosecution history has established that *the inventor turned his attention to the subject matter* in question, knew the words for both the broader and narrower claim, *and affirmatively chose the latter*.[54]

Treating every estoppel as a complete foreclosure of equivalents, by contrast, would effectively deny the axiom that language captures an invention's essence only imperfectly:

> By amending the application, the inventor is deemed to concede that the patent does not extend as far as the original claim. It does not follow, however, that the amended claim becomes so perfect in its description that no one could devise an equivalent. *After amendment, as before, language remains an imperfect fit for invention*.[55]

As a result, "there is no more reason for holding the patentee to the literal terms of an amended claim than there is for . . . holding every patentee to the literal terms of the patent."[56] The Federal Circuit's absolute bar could not stand.

Just as clearly, however, the Supreme Court opted against a simplistic return to the unworkable *status quo ante* that, in the Federal Circuit's view, had "l[ed] to excessive uncertainty and burden[ed] legitimate innovation."[57] Specifically, adapting a strategy it had used in an equivalents case just five years earlier,[58] the Court restructured the process of proof by creating a new rebuttable "presumption that prosecution history estoppel bars a finding of equivalence."[59] If the patentee secured the asserted claim from the Patent Office only by amending the claim's text to narrow its scope, thus sacrificing coverage,

> *the patentee should bear the burden* of showing that the amendment does not surrender the particular equivalent in question. . . . The patentee, *as the author of the claim language*, may be expected to draft

---

54. *Id.* at 734–35 (emphasis added); *see also id.* at 737–38 ("[T]he purpose of applying the estoppel in the first place [is] to hold the inventor to the representations made during the application process and to the inferences that may reasonably be drawn from the amendment.").

55. *Id.* at 738 (emphasis added).

56. *Id.*

57. *Id.* at 737.

58. *See* Warner-Jenkinson Co. v. Hilton Davis Chem. Co., 520 U.S. 17, 33 (1997) ("In our view, holding that certain reasons for a claim amendment may avoid the application of prosecution history estoppel is not tantamount to holding that the *absence* of a reason for an amendment may similarly avoid such an estoppel. Mindful that claims do indeed serve both a definitional and a notice function, we think the better rule is to place the burden on the patent holder to establish the reason for an amendment required during patent prosecution. . . . Where no explanation is established, however, the court should presume that the patent applicant had a substantial reason related to patentability for including the limiting element added by amendment.").

59. *Festo*, 535 U.S. at 741.

> claims encompassing readily known equivalents. A patentee's *decision to narrow* his claims through amendment *may be presumed to be a general disclaimer* of the territory between the original claim and the amended claim.[60]

The patentee, rather than the accused infringer (or, more generally, the public), thus bears the immediate cost of the uncertainty that the claim scope change produces. Of course, the patentee has the opportunity to rebut the presumption that it surrendered the scope it now seeks to enforce. To do so, "[t]he patentee must show that at the time of the amendment one skilled in the art could not *reasonably* be expected to have drafted a claim that would have literally encompassed the alleged equivalent."[61] We thus look to the choices a skilled patentee *could have made*, to judge the quality of the choice this specific patentee *did make*. In particular, the patentee can excuse his or her failure to have recited (literally) the scope it now seeks (equivalently) by showing that the equivalent was "unforeseeable at the time of the application."[62]

By directing this new presumption against the patentee and prescribing means for rebutting it, the Court underscored the need to keep uncertainty costs closely tethered to the patentee's specific claim language choices. Patentees doubtless grumbled, in the wake of *Festo*, that it is difficult to make such pivotal choices about claim language before all the market consequences of the choice are apparent. But of course the gravity of the choices, and prudently allocating the cost of mistaken choices, is precisely the point: The flexibility the law affords, with doctrines such as equivalent infringement, is rooted in reasonably unavoidable imperfections, not studied, self-serving equivocations. The spirit of *Festo* is that a patentee has no right to use avoidably uncertain claim language (to preserve *post hoc* enforcement flexibility) that the patent system is bound to respect.

---

60. *Id.* at 740 (emphases added).

61. *Id.* at 741 (emphasis added). Note the use of "reasonably": We neither demand perfection nor accept sloth.

62. *Id.* at 740. Foreseeability thus helps limit the duty to a reasonable scope, much as it does in negligence. *See* DAN B. DOBBS ET AL., THE LAW OF TORTS § 159 (2d ed. 2016) ("[F]oreseeability of harm, though not sufficient, is necessary to show negligence. No actor can be counted as negligent unless he either actually foresaw, or a reasonable person in a similar position would have foreseen that harm to someone's interests was an unreasonably likely outcome of his conduct."); RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL AND EMOTIONAL HARM § 3 (AM. LAW INST. 2010) ("A person acts negligently if the person does not exercise reasonable care under all the circumstances. Primary factors to consider in ascertaining whether the person's conduct lacks reasonable care are the foreseeable likelihood that the person's conduct will result in harm, the foreseeable severity of any harm that may ensue, and the burden of precautions to eliminate or reduce the risk of harm.").

B. The Nautilus *Standard*

The issue in *Nautilus, Inc. v. Biosig Instruments, Inc.*[63]—the proper standard for judging claim definiteness, which is a validity doctrine—was different from the issue in *Festo*—the scope of prosecution history estoppel, which is an infringement doctrine. But the Court analyzed the *Nautilus* issue in terms remarkably similar to the *Festo* issue. Indeed, *Festo* was the very first decision the *Nautilus* Court cited.[64]

The patent claim in dispute in the case recited a heart-rate monitor in the form of a handle for exercise equipment, such as a treadmill.[65] The disputed claim term, which pertained to a set of detection electrodes arrayed on the handle, required that the electrodes be "in spaced relationship with each other."[66] But the claim did not specify further just what that "spaced relationship" was, either in absolute numerical terms or in relative terms.[67] The trial court had found the claim limitation indefinite, and a Federal Circuit panel—applying the quite patentee-forgiving, uncertainty-tolerant "insolubly ambiguous" standard—had reversed that judgment.[68] The Supreme Court rejected the Federal Circuit standard, although it did not go further to apply the new standard—"inform[ing] with reasonable certainty"—to the disputed claim language in the case.[69]

As the Court noted, several settled claim construction principles were common ground among the parties: One assesses claim definiteness "from the perspective of someone skilled in the relevant art," not from the point of view of a generalist judge or juror; and in assessing claim clarity, one construes the claim language "in light of the patent's specification and prosecution history."[70] But how to resolve the parties' competing contentions about the level of clarity the Patent Act demands?

The Court framed the inquiry as a matter of balancing the critical need for clarity against the stubborn reality of language's imperfections. And in doing so, it drew directly from *Festo*. "On the one hand, the definiteness requirement must take into account the inherent limitations

---

63. 134 S. Ct. 2120 (2014).

64. *Id.* at 2124 ("'Th[at] monopoly is a property right,' and 'like any property right, its boundaries should be clear.'") (quoting *Festo*, 535 U.S. at 730).

65. *Id.* at 2122.

66. *Id.*

67. *See id.* at 2127.

68. *Id.* at 2127–28.

69. *Id.* at 2131.

70. *Id.* at 2128.

of language. Some modicum of uncertainty, the Court has recognized, is the 'price of ensuring the appropriate incentives for innovation.'"[71] On the other hand, however, "a patent must be precise enough to afford clear notice of what is claimed . . . [o]therwise there would be '[a] zone of uncertainty which enterprise and experimentation may enter only at the risk of infringement claims.'"[72] Faced with this unavoidable tension between the need for clarity and the ineliminable imprecision of language, the Court described its task as "reconcil[ing] concerns that tug in opposite directions."[73]

The reconciling construct the Court chose is *reasonably* certain notice. The Court read § 112 of the Patent Act "to require that a patent's claims, viewed in light of the specification and prosecution history, inform those skilled in the art about the scope of the invention with reasonable certainty."[74] By settling on reasonableness, thereby moderating between extremes,[75] the Court "mandate[d] clarity, while recognizing that absolute precision is unattainable."[76] We can, in turn, satisfy this mandate for reasonable clarity with verifiable data about the language usage of those skilled in the relevant art at the relevant time.[77] Just as we did in *Festo*, here again we look to the choices a skilled patentee *could have made* to judge the quality of the choice this specific patentee *did make*.

Admittedly, the *Nautilus* case does not fully set forth a process of proof for establishing or assessing such a factual foundation about language usage. It does, however, point strongly to a presumption framework that focuses on the patentee, familiar from *Festo*. In particular, the Court couched the clear-notice imperative as an important structural response to a patentee's basic incentive to retain vague claim language if possible:

> And absent a meaningful definiteness check, we are told, patent applicants face powerful incentives to inject ambiguity into their claims. See Brief for Petitioner 30–32 (citing patent treatises and drafting guides). See also Federal Trade Commission, The Evolving IP

71. *Id.* (citation omitted) (quoting *Festo*, 535 U.S. at 732).

72. *Id.* at 2129 (quoting United Carbon Co. v. Binney & Smith Co., 317 U.S. 228, 236 (1942)).

73. *Id.*

74. *Id.*

75. *See supra* note 29 and accompanying text.

76. *Nautilus*, 134 S. Ct. at 2129.

77. Usage at the time the patent was written and filed is key: "[T]he definiteness inquiry trains on the understanding of a skilled artisan at the time of the patent application, not that of a court viewing matters *post hoc*." *Id.* at 2130.

> Marketplace: Aligning Patent Notice and Remedies With Competition 85 (2011) (quoting testimony that patent system fosters "an incentive to be as vague and ambiguous as you can with your claims" and "defer clarity at all costs"). Eliminating that temptation is in order, and "the patent drafter is in the best position to resolve the ambiguity in . . . patent claims." *Halliburton Energy Servs., Inc. v. M–I LLC*, 514 F.3d 1244, 1255 (Fed. Cir. 2008). *See also Hormone Research Foundation, Inc. v. Genentech, Inc.*, 904 F.2d 1558, 1563 (Fed. Cir. 1990) ("It is a well-established axiom in patent law that a patentee is free to be his or her own lexicographer . . . .").[78]

The echo from *Festo* is unmistakable: The patentee captains the claims, and it is thus fair to keep on the patentee's side of the ledger the reasonably foreseeable consequences of the patentee's own claim drafting choices. Equally vital is the Court's explicit recognition, quoting *Hormone Research*, that every patentee can set forth a claim boundary not only by using the familiar words or phrases established in the art, but also by inventing new words—or, more likely, new definitions for existing words—if that is what the patentee thinks is best under the circumstances. There is, as a result, simply no basis at all for indulging the patentee's "temptation" to "inject ambiguity into [his or her] claims."[79]

At the time the Court decided *Nautilus*, the Federal Circuit remained internally divided over the question whether claim construction entailed findings of fact that were entitled to deferential review on appeal.[80] In January 2015, in *Teva Pharmaceuticals USA, Inc. v. Sandoz, Inc.*,[81] the Court turned to *that* question.

C. The Teva *Determination*

The Court's 1996 decision in *Markman v. Westview Instruments, Inc.* put to rest the contention that the Seventh Amendment's jury-trial right required that juries answer claim construction questions—it does not.[82] What *Markman* left open, however, was the appellate standard of review for district court claim construction rulings. In *Teva*—a case, like *Nautilus*, centered on an indefiniteness challenge to a patent asserted in

78. *Id.* at 2129 (alterations in original) (footnote omitted).

79. *Id.*

80. *See supra* note 11 (discussing the Federal Circuit's post-*Markman* en banc cases on the standard of review for claim construction determinations).

81. 135 S. Ct. 831 (2015).

82. 517 U.S. 370, 372 (1996) ("We hold that the construction of a patent, including terms of art within its claim, is exclusively within the province of the court.").

litigation—the Court held that the "clearly erroneous" standard in Federal Rule of Civil Procedure 52(a)(6) "appl[ies] when a court of appeals reviews a district court's resolution of subsidiary factual matters made in the course of its construction of a patent claim."[83]

The Court further held that the *ultimate* construction of a disputed claim term is a legal question, subject to plenary review on appeal, setting up a two-step hybrid standard: "The district judge, after deciding the factual dispute, will then interpret the patent claim in light of the facts as he has found them. This ultimate interpretation is a legal conclusion. The appellate court can still review the district court's ultimate construction of the claim *de novo*."[84]

Not *all* claim construction disputes will involve subsidiary fact disputes: "We recognize that a district court's construction of a patent claim . . . often requires the judge only to examine and to construe the document's words without requiring the judge to resolve any underlying factual disputes."[85] On the other hand, "[i]n some cases . . . the district court will need to look beyond the patent's intrinsic evidence and to *consult extrinsic evidence* in order to understand, for example, the background science or *the meaning of a term in the relevant art during the relevant time period*."[86] What's more, "in some instances, a factual finding may be close to dispositive of the ultimate legal question of the proper meaning of the term in the context of the patent."[87] The distinction between intrinsic and extrinsic evidence that the Court flagged, which has become a core fixture of the post-*Markman* era,[88] helps one discern the way that *Teva* highlights *all* the resources that are available in the "reasonable certainty" inquiry that *Nautilus* mandates. Those resources include extrinsic evidence about the meaning of a term in art *A* at time *T*, an evidentiary category about semantic usage within which corpus linguistics data comfortably fall.

---

83. *Teva*, 135 S. Ct. at 836.

84. *Id.* at 841. The Court framed that ultimate legal question this way: "whether a skilled artisan would ascribe that same meaning to that term *in the context of the specific patent claim under review*." *Id.*

85. *Id.* at 840–41.

86. *Id.* at 841 (emphases added).

87. *Id.* at 841–42.

88. *See* Phillips v. AWH Corp., 415 F.3d 1303, 1314–19, 1321–24 (Fed. Cir. 2005) (en banc) (describing the variety and weight of intrinsic and extrinsic sources for construing disputed claim terms); Vitronics Corp. v. Conceptronic, Inc., 90 F.3d 1576, 1581–83 (Fed. Cir. 1996) (same); *see also* MUELLER, *supra* note 4, at 636–38 (summarizing the intrinsic/extrinsic distinction). The Federal Judicial Center's *Patent Case Management Judicial Guide*, canvassing twenty years of practice, provides a concise chart of the major intrinsic and extrinsic sources. PETER S. MENELL ET AL., PATENT CASE MANAGEMENT JUDICIAL GUIDE 5-36 (3d ed. 2016) (specifically Chart 5.1).

The facts of *Teva* make the point. In *Nautilus*, the parties had focused on intrinsic sources.[89] In *Teva*, by contrast, the parties relied heavily on extrinsic evidence in the form of expert testimony.[90] To be sure, the dispute in *Teva* about which of the three possible meanings to choose for the claim term "molecular weight"[91] had a critical intrinsic hook in the patent's written disclosure—namely, the measurement data depicted in figure 1 of the patent:

> Teva argued . . . that the term "molecular weight" in the patent meant molecular calculated in the first way . . . . Sandoz, however, argued that figure 1 of the patent showed that Teva could not be right . . . That figure, said Sandoz, helped to show that the patent term did *not* refer to the first method of calculation.[92]

But the dispute over the figure's meaning spilled out into a dispute between the parties' conflicting experts:

> Teva's expert testified that a skilled artisan would understand that converting data from a chromatogram to molecular weight distribution curves like those in figure 1 would cause the peak on each curve to shift slightly; this could explain the difference between the value indicated by the peak of the curve (about 6.8) and the value in the figure's legend (7.7). Sandoz's expert testified that no such shift would occur. The District Court credited Teva's expert's account, thereby rejecting Sandoz's expert's explanation. The District Court's finding about this matter was *a factual finding*—about *how a skilled artisan would understand the way in which a curve created from chromatogram data reflects molecular weights*. Based on that factual finding, the District Court reached the legal conclusion that figure 1 did not undermine Teva's argument that molecular weight referred to the first method of calculation (peak average molecular weight).[93]

The disputed portion of the patent may have been a chromatogram, rather than solely a word or phrase, but the extrinsic evidence was important

---

89. *See* Nautilus, Inc. v. Biosig Instruments, Inc., 134 S. Ct. 2120, 2127, 2131 (2014) (reviewing parties' arguments about intrinsic materials); Biosig Instruments, Inc. v. Nautilus, Inc., 783 F.3d 1374, 1382, 1384 (Fed. Cir. 2015) (deciding, on remand, to "revisit the intrinsic evidence here to make clear that a skilled artisan would understand with reasonable certainty the scope of the invention," and concluding that "[i]n this case, a skilled artisan would understand the inherent parameters of the invention as provided in the intrinsic evidence"), *cert. denied*, 136 S. Ct. 569 (2015).

90. *See Teva*, 135 S. Ct. at 843.

91. *Id.* at 836, 842 (describing the three possible meanings of this claim term).

92. *Id.* at 842. The figure itself is in an Appendix to the majority's opinion. *Id.* at 844.

93. *Id.* at 843 (emphasis added) (citations omitted); *see also* Teva Pharm. USA, Inc. v. Sandoz, Inc., 789 F.3d 1335, 1341–42 (Fed. Cir. 2015) (discussing, on remand, extrinsic evidence about figure 1).

because it helped answer a question about *meaning*, about semantic usage (in the relevant art at the relevant time).

Of course, an expert witness may "testify in the form of an opinion" only if, among other things, "the testimony is based on sufficient facts or data."[94] The materials on which an expert witness relies to form opinions about the meaning of a claim term can also form the basis for a technology tutorial at the claim construction stage of enforcement litigation,[95] which has increasingly become the stage for adjudicating indefiniteness issues as well.[96] In any event, the objective basis for the opinion about semantic usage is critical: "an expert's *subjective* understanding of a patent term is irrelevant."[97] To gain any traction in a claim construction dispute, an expert witness must ground his or her opinion about usage and meaning in the relevant art with independent evidence.[98] Patents and other published documents from before the filing of the patent under review, known collectively by the shorthand *prior art references*,[99] are a key type of extrinsic evidence for experts, and the courts, to consider:

> [C]ourts are free to consider the prior art when ruling on claim construction. . . . Even apart from prior art recited in the patent and the prosecution history, it is important for trial courts to have the context of other prior art that will form the basis of an invalidity defense. Those prior art references may play as large a role in shaping the claim-construction dispute as does the accused device.[100]

---

94. FED. R. EVID. 702(b). Those underlying materials need not be otherwise admissible evidence in the case, so long as "experts in the particular field would reasonably rely on those kinds of facts or data in forming an opinion on the subject." FED. R. EVID. 703.

95. *See* MENELL ET AL., *supra* note 88, at 5-16–5-18 (describing such tutorials).

96. *See id.* at 5-84.

97. Gen. Protecht Grp., Inc. v. Int'l Trade Comm'n, 619 F.3d 1303, 1310 (Fed. Cir. 2010) (emphasis added).

98. *See* Symantec Corp. v. Computer Assocs. Int'l, Inc., 522 F.3d 1279, 1289 n.3 (Fed. Cir. 2008) (according no weight to an expert's opinion on claim meaning that is not grounded in evidence showing a claim term's accepted meaning in the relevant field); Sinorgchem Co., Shandong v. Int'l Trade Comm'n, 511 F.3d 1132, 1137 n.3 (Fed. Cir. 2007) (same). In short, "[e]xpert opinions should be grounded both in the intrinsic evidence and by support in other independent, reliable sources." MENELL ET AL., *supra* note 88, at 5-43.

99. *See* MUELLER, *supra* note 4, at 229 (noting that "prior art **references** . . . are documents such as patents, articles from scientific journals, and other technical literature," establishing a baseline for what is unpatentably old under 35 U.S.C. § 102—"the catalog or universe of prior art that potentially can be cited by the USPTO in rejecting a patent applicant's claims").

100. MENELL ET AL., *supra* note 88, at 5-23–5-24; *see also id.* at 5-39 (noting that "extrinsic evidence can be useful [to claim construction], and *Phillips* and *Teva* confirm that district courts are free to consider extrinsic evidence, including expert testimony, dictionaries, treatises, and other such sources").

This art-specific material from about the time the patent was drafted and filed, which reflects usage among patent claimants and others in the art at that time, justifies the expert's opinion and educates the trial judge. The judge can then make a well-supported finding in light of that material when its significance has been explained.

A recent Federal Circuit case illustrates the vital role such extrinsic usage evidence can play in resolving fact disputes in an indefiniteness case. In *Icon Health & Fitness, Inc. v. Polar Electro Oy*,[101] the district court found the claim terms "in-band" and "out-of-band" fatally indefinite under the *Nautilus* "reasonable certainty" standard, and the Federal Circuit affirmed that result.[102] Interestingly, though it had already conducted a hearing on the parties' claim construction disputes, the district court held a new hearing and considered more briefing in the case after the Supreme Court decided *Nautilus*.[103] Moreover, at the district court's request, the parties "retain[ed] experts in order to ascertain the perspective of persons skilled in the art" and then "the district court held an evidentiary hearing to hear expert testimony."[104] Polar Electro, the accused infringer, prevailed in its indefiniteness defense.[105]

On appeal, affirming that result, the Federal Circuit emphasized the central role of the extrinsic evidence in the case. Patentee "Icon's and its expert's position ha[d] been that 'in-band' and 'out-of-band' communications are different from each other," which "difference is alone sufficient to render the claims definite and capable of construction."[106] But Polar Electro had shown that, in the prior art, numerous authors used these terms with greater clarity[107]—clarity that Icon could readily have achieved. The gap in the patent disclosure, Polar's expert testified, was that "[t]he 'relationship' between 'in-band' and 'out-of-band' . . . is completely unspecified."[108] Crucially, that gap was demonstrably out of step with the customary usage that prior art references reflected:

> To support its position that a [point of] reference is required to provide context and give terms 'in-band' and 'out-of-band' meaning, Polar's

101. 656 F. App'x 1008 (Fed. Cir. 2016).
102. *Id.* at 1010.
103. *Id.* at 1012.
104. *Id.* at 1012–13.
105. *Id.* at 1016.
106. *Id.* at 1013.
107. *See id.* at 1015.
108. *Id.* at 1014.

> expert proffered *ten extrinsic prior art patents and text books*, each of which 'defines a reference that allows the reader to differentiate in-band from out-of-band in relation to that reference.'[109]

Icon had no rebuttal,[110] and thus no explanation for why it failed to write its patent claims at least as clearly as several prior art references had been written. As the Federal Circuit highlighted, "[t]his is precisely the type of extrinsic evidence upon which a district court may rely in analyzing the record before it when construing claim terms."[111] Using the prior art to see the choice Icon *could have* made, and thus to judge the quality of the choice Icon *did* make, both the trial and appellate courts concluded that Icon fell fatally short.

Polar Electro shows that one can successfully evaluate reasonable certainty against the backdrop of a sample of prior art documents that establish how, at a given time, artisans in a field used words and phrases to achieve a given level of clarity. What if, in that case, Polar Electro had gathered not a group of ten prior art references, but 100? Or 1,000? Or 100,000? And could both show and explain the upshot of that collection without unduly burdening the court? These larger aggregations are the stuff of corpus linguistics. And though they differ greatly in magnitude from the prior part bundle in *Polar Electric*, at bottom the aggregations serve to prove the same basic facts about usage at a given time and place. Before examining corpus linguistics in more detail in Part IV, however, it is helpful to explore the meaning of reasonable care in tort law, as well as the rigors of reasonable notice under the Due Process clause.

III. DETERMINING REASONABLY CERTAIN NOTICE

Nautilus, abjuring perfect clarity and condemning studied vagueness, demands that a patent claim "inform those skilled in the art about the scope of the invention with reasonable certainty."[112] How does one determine, in a given case, whether a claim has achieved a level of clarity that is *reasonable*, rather than too low? (Clarity greater than the reasonable level is no cause for objection.) Generally speaking, reasonableness connotes a good fit between means and ends amid varied options and cross-cutting goals, a point of moderation that

---

109. *Id.* at 1015 (emphasis added) (quoting the district court's *Markman* decision in the case).
110. *See id.* at 1015 n.1.
111. *Id.* at 1015.
112. Nautilus, Inc. v. Biosig Instruments, Inc., 134 S. Ct. 2120, 2129 (2014).

accommodates competing values and avoids extremes.[113] After discussing tort law's reasonable care standard and the evidentiary value of custom, I consider litigation procedures that mandate reasonable notice, where the duty rolls with the available means. Both suggest that we judge reasonable certainty against available options.

A. Reasonable Care

"Negligence, the foundation of tort law, turns on a reasonableness standard."[114] In particular, "[t]he premise of negligence law is that we owe everyone a duty to exercise reasonable care in the conduct of our affairs."[115] The current *Restatement* thus provides that "[a] person acts negligently if the person does not exercise reasonable care under all the circumstances."[116] The prior two *Restatements* were to the same effect.[117] The standard is an open one, "permit[ting] negligence to function as an all-purpose accident tort."[118] By demanding *reasonable* care, the negligence tort moderates between the untenable extremes of perfect security for plaintiffs and perfect freedom for defendants. It also refracts this moderation through a mutuality lens, "entrench[ing] the principle that being reasonable requires constraining one's conduct by reference to the perils one creates for others."[119]

H.L.A. Hart highlighted the negligence tort's synthesis of moderation and mutuality, though he did so with an eye toward understanding how the law deploys general standards to govern varied circumstances. After discussing the use of general standards in

---

113. *See* MacCormick, *supra* note 29, at 1587 ("[W]hat is presupposed in any resort to reasonableness as a standard is that there is some topic or focus of concern to which, in accordance with variable circumstances, various factors are relevant, these having to be set in *an overall balance of values* one way or the other.") (emphasis added); REASONABLENESS AND LAW, *supra* note 29, at xi ("As a normative criterion . . . the reasonable . . . is structured by a core meaning that consists in its calling to take into account different claims and reasons so as to find among them a common ground and an equilibrium . . . .").

114. Zaring, *supra* note 26, at 536; *see also id.* at 539 (observing that "reasonableness has proven to be . . . so useful [a concept in torts] that it has become the jurisprudential underpinning of that entire field of law").

115. *Id.* at 537.

116. RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL AND EMOTIONAL HARM § 3 (AM. LAW INST. 2010).

117. *See* RESTATEMENT (SECOND) OF TORTS § 283 (AM. LAW INST. 1965) ("Unless the actor is a child, the standard of conduct to which he must conform to avoid being negligent is that of a reasonable man under like circumstances."); RESTATEMENT OF TORTS § 283 (AM. LAW INST. 1934) ("Unless the actor is a child or an insane person, the standard of conduct to which he must conform to avoid being negligent is that of a reasonable man under like circumstances.").

118. JOHN C.P. GOLDBERG & BENJAMIN C. ZIPURSKY, TORTS 82 (2010).

119. Zipursky, *supra* note 29, at 2162.

administration,[120] he turned to the use of such standards in adjudication. Concluding that adjudication is fruitful "where the sphere to be controlled is such that . . . the range of circumstances, though very varied, covers familiar features of common experience," he observed that in such cases "common judgments of what is 'reasonable' can be used by the law."[121] To illustrate the strategy, Hart turned to "[t]he most famous example of this technique in Anglo-American law"—the reasonable care standard in negligence.[122] He described the standard's target as an intermediate point of balance respecting two parties' interests:

> [S]anctions may be applied to those who fail to take reasonable care to avoid inflicting physical injuries . . . . But what is reasonable or due care in a concrete situation? . . . What we are striving for in the application of standards of reasonable care is to ensure (1) that precautions will be taken which will avert substantial harm, yet (2) that the precautions are such that the burden of proper precautions does not involve too great a sacrifice of other respectable interests.[123]

Or as Professor MacCormick put it, exploring Hart's analysis, it is "the striking of a balance between the two values of relative security from harm and relative liberty to do as you like."[124] A plaintiff's plea is for security, and a defendant's is for liberty. Reasonable care vindicates both, in an equilibrium that credits both in each case and across the run of cases. "Being reasonable requires a curtailment of what one might really want in order to accommodate others' legitimate interests."[125]

---

120. H.L.A. HART, THE CONCEPT OF LAW 131 (2d ed. 1994) (explaining that "the legislature sets up very general standards and then delegates to an administrative, rule-making body acquainted with the varying types of case, the task of fashioning rules adapted to their special needs").

121. *Id.* at 132.

122. *Id.*

123. *Id.* at 132–33.

124. MacCormick, *supra* note 29, at 1585; *see also id.* at 1584–85 ("On the one hand, we set value upon the security of persons and their property and their economic interests from damage resulting from others' acts. On this account, we think it right and proper that each person take care to avoid inflicting bodily harm on others or damaging their property or economic well-being. On the other hand, we set value upon the freedom of individuals to pursue their own activities and way of life without having to undertake an intolerable burden of precautions against the risks of damage to others. The law has to express a balance between these values in general terms . . . by prescribing that such care has to be taken as would be taken by a reasonable and prudent person.") (footnotes omitted). Courts have also made this point. *See, e.g.*, Ray v. Am. Nat'l Red Cross, 696 A.2d 399, 404 (D.C. 1997) (observing that "our modern rule of negligent liability proceeds from a balance struck between an actor's freedom of choice and another's security in person and property").

125. Zipursky, *supra* note 29, at 2162. Professor Ripstein similarly described the mutuality that reasonable care instantiates: "Reasonableness tests are not a proxy for some other measure of responsibility; they are constitutive of responsibility, understood in terms of the ways in which people are accountable to each other." Arthur Ripstein, *Reasonable Persons in Private Law*, *in* REASONABLENESS AND LAW, *supra* note 29, at 256.

Transposing this to the claim definiteness context, an accused infringer's plea is for more clarity, and a patentee's is for more charity. Reasonable certainty vindicates both, in an equilibrium that credits both in each case and across the run of cases. Being reasonable requires a patentee to curtail its interest in *ex post* enforcement flexibility, in order to accommodate the public's legitimate interest in *ex ante* planning.[126] At the same time, it requires an accused infringer to curtail its interest in perfect clarity, in order to accommodate the patentee's legitimate interest in a fair appraisal of the clarity its claim drafting efforts achieved, given the options that were available in the art at the time.[127] In this sense, reasonable certainty is less a proxy for claim definiteness than it is constitutive of it, understood as a way that patentees and the public are accountable to one another for the clarity the claim language provides.[128]

Returning to the negligence tort, the reasonable care standard is not simply an abstract commitment to moderation and mutuality, though they are critically important precepts.[129] In concrete terms, demonstrating that one failed to act with reasonable care (or, rather, successfully hewed to it) includes providing an evidentiary context for evaluating the accused tortfeasor's conduct. Evidence of customary practice, in particular, plays a central role. "[T]he rule for the greater part of negligence law" is that "evidence that the defendant complied with customary understandings of what ordinary care requires is *highly relevant* to the breach issue, but *not*

---

126. *See* Nautilus, Inc. v. Biosig Instruments, Inc., 134 S. Ct. 2120, 2129 (2014) ("[A] patent must be precise enough to afford clear notice of what is claimed, thereby apprising the public of what is still open to them. Otherwise there would be a zone of uncertainty which enterprise and experimentation may enter only at the risk of infringement claims. And absent a meaningful definiteness check, we are told, patent applicants face powerful incentives to inject ambiguity into their claims. Eliminating that temptation is in order, and the patent drafter is in the best position to resolve the ambiguity in patent claims.") (internal alterations, citations, footnotes, and quotation marks omitted).

127. *See id.* at 2128 ("[T]he definiteness requirement must take into account the inherent limitations of language. Some modicum of uncertainty, the Court has recognized, is the price of ensuring the appropriate incentives for innovation. One must bear in mind, moreover, that patents are not addressed to lawyers, or even to the public generally, but rather to those skilled in the relevant art." (internal citations and quotation marks omitted)).

128. *Cf.* Ripstein, *supra* note 125, at 256; *see also* Zipursky, *supra* note 29, at 2169 ("The mutuality conception of reasonableness is found in other parts of the law—such as nuisance law—and throughout ordinary conversation as well. A reasonable person in this sense is not wholly insensitive to the fact that, like her, others have wishes, desires, demands, and needs. She has a sense of mutuality. This is a significant part of what the reasonable person or the reasonably prudent person standard highlights . . . .").

129. *See* Zipursky, *supra* note 29, at 2169 ("Reasonableness in negligence law relates both to the idea of moderate care and to the idea of mutuality as to others' needs. *Moderation and mutuality* is the watchword of negligence law.").

dispositive."[130] The current *Restatement* frames the value of custom evidence from both directions, providing as follows:

> (a) An actor's compliance with the custom of the community, or of others in like circumstances, is evidence that the actor's conduct is not negligent but does not preclude a finding of negligence.
>
> (b) An actor's departure from the custom of the community, or of others in like circumstances, in a way that increases risk is evidence of the actor's negligence but does not require a finding of negligence.[131]

The prior *Restatement* is to like effect,[132] and both draw on the approach Judge Learned Hand took in a famed admiralty case against two tugboats that lost their barges in a storm that took them unawares (for lack of radios on board), *The T.J. Hooper* case.[133] As Judge Hand observed in the case, rejecting the contention that industry practice was dispositive, "in most cases reasonable prudence is in fact common prudence; but strictly it is never its measure; a whole calling may have unduly lagged in the adoption of new and available devices."[134] Custom is relevant, but not dispositive. Of course, there may be no relevant customary practice that throws light on a given defendant's conduct. In such a case, the parties and the court may find it useful to consider the circumstances of the accident in light of the burden of greater precautions, the likelihood of harm in their absence, and the magnitude of that harm should it occur—in other words, the framework Judge Hand developed in another renowned admiralty case, *United States v. Carroll Towing Co.*,[135] decided 15 years after *The T.J. Hooper*.[136]

---

130. GOLDBERG & ZIPURSKY, *supra* note 118, at 144. The "courts have crafted an exception for . . . claims for professional negligence—i.e., malpractice—and particularly medical malpractice." *Id.* at 145. This exception is not germane here, inasmuch as claim indefiniteness is an attack on a particular patent claim's validity, not necessarily an attack on a patent drafter's professional competence in writing that claim in a given way.

131. RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL AND EMOTIONAL HARM § 13 (AM. LAW INST. 2010).

132. *See* RESTATEMENT (SECOND) OF TORTS § 295A (AM. LAW INST. 1965) ("In determining whether conduct is negligent, the customs of the community, or of others under like circumstances, are factors to be taken into account, but are not controlling where a reasonable man would not follow them.").

133. 60 F.2d 737 (2d Cir. 1932).

134. *Id.* at 740.

135. 159 F.2d 169 (2d Cir. 1947).

136. Professors Goldberg & Zipursky critique the mathematicization of Judge Hand's *Carroll Towing* approach by law and economics scholars and judges, GOLDBERG & ZIPURSKY, *supra* note 118, at 149–51, maintaining that "Hand himself never supposed that his formula could be calculated or applied with any sort of precision . . . claim[ing instead that] it had value primarily in focusing attention on the sorts of considerations that arguably matter most to the determination of fault." *Id.*

In a claim definiteness dispute, the analog to custom evidence in a negligence case is semantic-usage evidence about words or phrases in the relevant technological field at the time the patentee drafted the claim under review. That usage evidence demonstrates, objectively, the range of established semantic expressions on which patentees customarily drew to convey a specific idea, as well as the backdrop against which a patentee could draft a bespoke definition of its own (if that specially defined claim language would better capture the invention, in the patentee's estimation). Corpus linguistics methodologies can provide us reliable access to objective usage data on a tailored, targeted basis.

Negligence law's reasonable care standard, then, provides neither that one may utterly ignore risks to others nor that one must compensate others for any and all harms. Reasonableness moderates between these extremes. Evidence of customary practice, when available, goes a long way toward resolving whether a particular negligence defendant acted reasonably under the relevant circumstances. The same is true for claim definiteness: Perfect clarity is not obtainable, nor is self-serving ambiguity tolerable. We require reasonable certainty to moderate between these extremes. Evidence of customary semantic usage, which we should endeavor to make more readily available, goes a long way toward resolving whether a particular patentee chose clear claim language for the relevant art at the relevant time. Critically, what this perspective rules out of bounds is an accused infringer's contention that a given claim term is indefinite simply by virtue of the fact that—in court, long after the patent was drafted—one can show a word or phrase that, had *it* been used in the claim instead, would have prevented the semantic uncertainty on which the accused infringer's attack focuses. Of course one can confect, with the benefit of hindsight, an imagined alternative text that makes the actual claim term seem less clear. But one can just as easily confect alternative language that makes the actual claim term seem clearer. Both parties, in short, can play an "it's not as good [bad] as X" game. In a reasonableness framework that game isn't worth the candle.

---

at 149. This question of deep tort theory is not one that I can, or need to, resolve here. The key point, for present purposes, is that in actual tort cases custom evidence can be quite relevant to, but not dispositive on, the core question whether the accused tortfeasor has fallen below the standard of reasonable care—a point that even the most law-and-economics-focused jurists readily accept. *See, e.g.*, Rodi Yachts, Inc. v. Nat'l Marine, Inc., 984 F.2d 880, 888 (7th Cir. 1993) (Posner, J.) (discussing the role of custom evidence on the question of breach). Indeed, *Carroll Towing* itself adverts to the heft that custom evidence likely would have had in the case, had the defendant's employee been absent only overnight, rather than during working hours. *Carroll Towing*, 159 F.2d at 173 ("We need not say whether, even in such crowded waters as New York Harbor a bargee must be aboard at night at all; *it may be that* the custom is otherwise . . . and that, if so, *the situation is one where custom should control*. We leave that question open . . . .") (emphasis added).

The central question to illuminate, instead, is whether the claim term under review, when contextualized with actual usage evidence from the relevant art and time, embodied a drafting choice well calculated to provide clear notice as of that time. The proper inquiry thus requires that we shine a brighter light on the textual alternatives that were demonstrably available in the relevant field at the relevant time. It is helpful to recognize, in this connection, that the demand for reasonable efforts at providing prior notice of a judicial proceeding, under the Constitution's Due Process clause, has *also* come to focus on the demonstrably available means for providing such notice in a given place and time. This reasonable-notice jurisprudence merits attention here.

B. Reasonable Notice

The law predicates the sound adjudication of duties and freedoms on reasonably certain notice to the parties affected, not only in the patent system but also in the basic fabric of any binding judicial process. It is fundamental to U.S. constitutional law that the state cannot deprive one of life, liberty, or property without due process of law.[137] The Supreme Court elaborated on the constitutional floor for adequate notice of judicial process, more than sixty-five years ago, in *Mullane v. Central Hanover Bank & Trust Co.*,[138] and *Mullane* remains the governing Due Process standard.[139] Efforts at notice sometimes fail, and a failure of

---

137. U.S. CONST. amend. V, cl. 3 ("No person shall be . . . deprived of life, liberty, or property, without due process of law[.]"); U.S. CONST. amend. XIV, § 1, cl. 3 ("No state shall . . . deprive any person of life, liberty, or property, without due process of law[.]").

138. 339 U.S. 306 (1950).

139. *See* United Student Aid Funds, Inc. v. Espinosa, 559 U.S. 260, 272 (2010) ("Due process requires notice 'reasonably calculated, under all the circumstances, to apprise interested parties of the pendency of the action and afford them an opportunity to present their objections.' *Mullane v. Central Hanover Bank & Trust Co.*, 339 U.S. 306, 314 (1950).") (concluding that a defect in service of process did not rise to the level of an unconstitutional deprivation of Due Process). The U.S. Courts of Appeals routinely invoke *Mullane*'s Due Process standard on notice questions. *See, e.g.*, DPWN Holdings (USA), Inc. v. United Air Lines, Inc., 747 F.3d 145, 150 (2d Cir. 2014); Alderwoods Grp., Inc. v. Garcia, 682 F.3d 958, 972 (11th Cir. 2012); Gates v. City of Chicago, 623 F.3d 389, 401 (7th Cir. 2010). And *Mullane* has remained the central touchstone for conceptualizing what Due Process requires. *See, e.g.*, Dusenbery v. United States, 534 U.S. 161, 168 (2002) ("Since *Mullane* was decided, we have regularly turned to it when confronted with questions regarding the adequacy of the method used to give notice."); Nelson v. Adams USA, Inc., 529 U.S. 460, 466 (2000) ("[T]he proceedings did not comply with Rule 15, and neither did they comport with due process. See*, e.g., Mullane v. Central Hanover Bank & Trust Co.*, 339 U.S. 306, 314 (1950) . . . ."); Cleveland Bd. of Educ. v. Loudermill, 470 U.S. 532, 542 (1985) ("An essential principle of due process is that a deprivation of life, liberty, or property 'be preceded by notice and opportunity for hearing appropriate to the nature of the case.' *Mullane v. Central Hanover Bank & Trust Co.*, 339 U.S. 306, 313 (1950)."); 4A CHARLES ALAN WRIGHT ET AL., FEDERAL PRACTICE & PROCEDURE § 1074 (4th ed. 2016) (stating that *Mullane* "generally is recognized as the keystone of the modern

actual notice does not, by itself, constitute a Due Process defect.[140] What, then, is an adequate notice mechanism? What would fall below the Due Process minimum?

The process that is due turns on the specific facts.[141] The particular question in *Mullane* was the constitutionally minimally adequate notice due to a common trust's beneficiaries before a binding judicial process could cut off their right to bring further claims against the trustee for a prior period's management and fees.[142] The trust company, complying with the process specified by the New York state statute that allowed for common trusts, gave notice of the judicial proceeding to the beneficiaries "by publication in a local newspaper."[143] The trustee did so even though, for some of the affected beneficiaries, it had names and mailing addresses on file; indeed, it had mailed information to the beneficiaries at the time it created the common fund in question.[144] Although the Court could accept publication notice—"a customary substitute . . . where it is not reasonably possible or practicable to give more adequate warning"—for the beneficiaries "whose interests or whereabouts could not with due diligence be ascertained,"[145] it could *not* accept publication notice for the beneficiaries the trustee knew by name and address: "we find no tenable ground for dispensing with a serious effort to inform them personally of the [judicial] accounting, at least by ordinary mail to the record addresses."[146] The trustee bank's reliance on a newspaper announcement fell short of the "elementary and fundamental requirement of due process": namely, "notice *reasonably* calculated, under all the circumstances, to apprise interested parties of the pendency of the action."[147]

---

philosophy regarding the due process aspects of a notice requirement and the importance of the case should not be underestimated").

140. *See Dusenbery*, 534 U.S. at 170. Perfect notice is not required—just as perfect security is not required in negligence law, and perfect clarity is not required of patent claims.

141. *See* Walker v. City of Hutchinson, 352 U.S. 112, 115 (1956) (noting "the impossibility of setting up a rigid formula as to the kind of notice that must be given; notice required will vary with circumstances and conditions").

142. 339 U.S. at 311 ("We understand that every right which beneficiaries would otherwise have against the trust company, either as trustee of the common fund or as trustee of any individual trust, for improper management of the common trust fund during the period covered by the accounting is sealed and wholly terminated by the decree.").

143. *Id.* at 309.

144. *Id.* at 310, 318.

145. *Id.* at 317.

146. *Id.* at 318.

147. *Id.* at 314 (emphasis added).

Critically, the Court contrasted a genuine attempt at actual notice with an inadequate "process which is a mere gesture,"[148] framing the inquiry as a comparison of a notifier's chosen method with the other methods then available:

> *The means* employed must be such as *one desirous of actually informing* the absentee [person] *might reasonably adopt* to accomplish it. The *reasonableness* and hence the constitutional validity of any chosen method *may be defended on the ground that* it is *in itself reasonably certain* to inform those affected *or*, where conditions do not reasonably permit such notice, that *the form chosen is not substantially less likely* to bring home notice *than other of the feasible and customary substitutes*.[149]

In other words, in a pattern analogous that which we see in *Festo* regarding equivalent infringement[150] and in *Nautilus* regarding indefiniteness,[151] the Court looks to the means of notice such a party *could have used* to judge the quality of the means of notice this party *did use*. And, consistent with a reasonableness approach that strikes a balance between extremes as it vindicates the legitimate interests of both sides, Due Process neither demands perfection (which would unfairly hold the sender to an unachievable goal) nor accepts sloth (which would unfairly disregard the recipient's right to heard). The trustee bank in *Mullane* fell short because, in the particular circumstances of that case, it used the plainly less effective of two feasible options: "Where the names and post office addresses of those affected by a proceeding are at hand, the reasons disappear for resort to means less likely than the mails to apprise them of its pendency."[152] Given that personal service of written notice meets any test,[153] one who *actually* strove to approximate it would mail notice to known names and addresses, not drop an ad in the local classifieds.

---

148. *Id.* at 315.

149. *Id.* (emphases added) (citations omitted).

150. *See supra* notes 61–62 and accompanying text.

151. *See supra* notes 74–79 and accompanying text.

152. 339 U.S. at 318; *see also id.* at 319 ("The statutory notice to known beneficiaries is inadequate, not because in fact it fails to reach everyone, but because under the circumstances it is not reasonably calculated to reach those who could easily be informed by other means at hand. However it may have been in former times, the mails today are recognized as an efficient and inexpensive means of communication. Moreover, the fact that the trust company has been able to give mailed notice to known beneficiaries at the time the common trust fund was established is persuasive that postal notification at the time of accounting would not seriously burden the plan.").

153. *Id.* at 313 ("Personal service of written notice within the jurisdiction is the classic form of notice always adequate in any type of proceeding.").

Mullane provides a framework for inquiry, not a "send notice by mail" rule: because the adequacy of notice turns on the details of a given case, notice by mail, though required in *Mullane*, may fall short in a different case. In *Jones v. Flowers*,[154] for example, the Court concluded that a state lands commissioner's use of certified mail fell below the Due Process minimum. The Court in *Jones* acknowledged that, at least in the case "when someone is home to sign for the letter," this more formal type of mail "makes actual notice more likely, because requiring the recipient's signature protects against misdelivery."[155] The problem for the commissioner, however, was that both his certified letters to Jones, the person whose property was sold to cover unpaid taxes, had twice been returned to the commissioner, unopened and marked "unclaimed."[156] Jones learned of the State's forced sale of his property only after it had been sold to Flowers, and Flowers "had an unlawful detainer notice delivered to the property."[157] Applying *Mullane*, the Court held that "when mailed notice of a tax sale is returned unclaimed, the State must take additional reasonable steps to attempt to provide notice to the property owner before selling his property, if it is practicable to do so."[158] The returned, unclaimed letter necessitates more effort because that is what, in the words of *Mullane*, "one desirous of actually informing" the other party would recognize:

> We do not think that a person who actually desired to inform a real property owner of an impending tax sale of a house he owns would do nothing when a certified letter sent to the owner is returned unclaimed. . . . [W]hen a letter is returned by the post office, the sender will ordinarily attempt to resend it, if it is practicable to do so. This is especially true when, as here, the subject matter of the letter concerns such an important and irreversible prospect as the loss of a house. Although the State may have made a reasonable calculation of how to reach Jones, it had good reason to suspect when the notice was returned that Jones was "no better off than if the notice had never been sent." Deciding to take no further action is not what someone "desirous of actually informing" Jones would do; such a person would take further reasonable steps if any were available.[159]

---

154. 547 U.S. 220 (2006).

155. *Id.* at 234.

156. *Id.* at 223–24.

157. *Id.* at 224. "The notice was served on Jones' daughter, who contacted Jones and notified him of the tax sale." *Id.*

158. *See id.* at 223, 225.

159. *Id.* at 229–30 (internal citations and quotation marks omitted); *see also id.* at 238 ("[W]e conclude, at the end of the day, that someone who actually wanted to alert Jones that he was in

In this case, the further steps that were reasonably available included "resend[ing] the notice [to Jones] by regular mail, so that a signature was not required," or "post[ing a] notice on the front door," or "address[ing] otherwise undeliverable mail to 'occupant.'"[160] The lesson of *Jones* is that reasonable notice takes account of both the available options and the notifier's best information about how effective the available mechanisms are likely to be at providing actual notice to the intended recipient. One must use the available means of notice that, objectively, would satisfy the well-informed person who *actually* strove to give *real* notice.

Tort law's reasonable care standard and procedural law's reasonable notice standard, then, fill out and reinforce the reasonable certainty inquiry that *Nautilus* mandates, including the careful consideration of facts about semantic usage that *Teva* underscores. Evidence of custom, whether to prove that one behaved with reasonable care (*e.g.*, the lack of radios in *The T.J. Hooper*) or that one provided reasonable notice of a judicial process (*e.g.*, the use of publication notice in *Mullane*), is relevant, but not dispositive. For patent claims, the relevant custom is the semantic usage of those in the relevant art at the relevant time.[161] Reasonableness vindicates the interests of opposing parties in each case, and across cases, by avoiding the extremes that would wholly sacrifice one side's interests to those of the other. We anchor that point of balance in a fact-sensitive assessment of the demonstrably available means that the responsible party—patentee, tort defendant, litigation plaintiff—could have used, to fairly judge the means that party did use. Specifically, following *Mullane*, to judge the responsible party's past conduct objectively, it helps to ask what a party actually striving to achieve the desired goal would have done in those circumstances with those available means. Whether the goal is a patent claim with a reasonably certain boundary, a course of conduct undertaken with reasonable care, or a means of notice reasonably calculated to apprise an interested party of a judicial process, we reject a hindsight-driven standard of perfection in favor of a fact-driven standard of good-faith, if imperfect, effort.

---

danger of losing his house would do more when the attempted notice letter was returned unclaimed, and there was more that reasonably could be done.").

160. *Id.* at 234–35.

161. Teva Pharm. USA, Inc. v. Sandoz, Inc., 135 S. Ct. 831, 841 (2015) (recognizing that, in a patent enforcement case, a district court judge may need "to consult extrinsic evidence in order to understand . . . the meaning of a term in the relevant art during the relevant time period").

IV. JUDGING DEFINITENESS IN PATENT CLAIMS

Both reasonable care and reasonable notice highlight the need, in the case of clashing interests, for moderation, mutuality, and a focus on a rich factual record about customary practice and readily available alternatives at a given place and time. All of this is powerful guidance for implementing *Nautilus*'s reasonable certainty standard for claim definiteness and *Teva*'s endorsement of extrinsic usage evidence. The Federal Circuit's 2016 *Polar Electro* decision strongly suggests the direct relevance of these concepts, given its reliance on extrinsic evidence of customary semantic usage in the prior art when judging whether a challenged claim is fatally indefinite.[162] It is important, in view of this guidance, to further detail—in both process and evidence terms—the optimal way to assess whether a given claim term is,[163] or was,[164] a reasonably certain means to the end of clear notice.

To judge definiteness objectively, we require evidence of customary usage in the art from the time the patent application was filed, which helps establish the range of available established words and phrases that a patent writer could have used, or against the backdrop of which a patent writer could have provided an explicit, application-specific definition (*i.e.*, been his or her "own lexicographer"). How, then, do we amplify the factual materials? We need to get more usage facts into the mix, with methods that are robust, orderly, and readily applicable to the specific questions that arise in concrete disputes. The tools developed in corpus linguistics can provide the facts the patent system needs, as the materials below demonstrate. Because enforcement litigation and patent examination tap into and process evidence differently, it is helpful to begin by considering, separately, these two contexts for deploying the evidence that corpus linguistics can provide.

A. The Litigation Process

In patent litigation, an accused infringer can raise the defense that the asserted patent claim is invalid for indefiniteness.[165] Every claim enjoys

---

162. *See supra* notes 101–11 and accompanying text.

163. During patent prosecution, per *In re Packard*, it is a present-tense question. *See supra* notes 17–21 and accompanying text.

164. During patent enforcement litigation, per *Nautilus*, it is a past-tense question. *See supra* note 77 and accompanying text.

165. *See* 35 U.S.C. § 282(b)(3)(A) (2012). The accused infringers in *Teva* and *Nautilus* did just that. *Teva*, 135 S. Ct. at 835–36; Nautilus, Inc. v. Biosig Instruments, Inc., 134 S. Ct. 2120, 2127 (2014).

a presumption of validity,[166] including a subsidiary presumption of definiteness,[167] and thus it falls to the accused infringer[168] to provide clear and convincing evidence to the contrary.[169] In addition, to help ensure that judging purported indefiniteness does not make tacit resort to a hindsight-driven demand for perfect clarity that unduly credits *post hoc* allegations of vagueness confected for litigation,[170] it is appropriate to require the accused infringer to initiate its indefiniteness defense with concrete evidence—not conjecture—that the clarity-enhancing language now proffered was demonstrably available to patent writers in the relevant art when the patent in suit was sought at the Patent Office. We can use the process of proof that *Festo* established as a template for an indefiniteness inquiry that focuses on this *ex ante* assessment; to do so comports with *Nautilus*'s focus on the critical policy tension, so central to *Festo*, between the need for clarity and the imperfection of language.

The process of proof patterned on *Festo*,[171] mutatis mutandis, is as follows:

• First, the accused infringer must come forward with evidence showing one or more concrete words or phrases, actually used in one or more English-language prior art patents from the same[172] technological art, that would have more clearly delineated the scope of the claim for

---

166. 35 U.S.C. § 282(a) (2012) ("A patent shall be presumed valid. Each claim . . . shall be presumed valid independently . . . .").

167. *See Nautilus*, 134 S. Ct. at 2130 n.10 (observing that § 282's "presumption of validity . . . incorporates th[e § 112] definiteness requirement by reference").

168. 35 U.S.C. § 282(a) (2012) ("The burden of establishing invalidity of a patent or any claim thereof shall rest on the party asserting such invalidity.").

169. *See* Microsoft Corp. v. i4i Ltd. P'ship, 564 U.S. 91, 102–03 (2011) (rejecting Microsoft's effort to overturn the case law requiring clear and convincing evidence of invalidity).

170. Assessing purported indefiniteness appears to include a risk of hindsight bias, akin to the risk of hindsight bias that can affect the assessment of purported obviousness. The risk of hindsight bias in an obviousness inquiry has long been recognized in cases and commentaries. *See* Joseph Scott Miller, *Remixing Obviousness*, 16 TEX. INTELL. PROP. L.J. 237, 239–50 (2008) (discussing these cases and commentaries). But I have not seen recognition of this indefiniteness-related risk of hindsight in any case or commentary. *Nautilus*, in critiquing the Federal Circuit's insoluble ambiguity standard, applied as of the time of the enforcement litigation, did take pains to instruct that the proper measure is "the understanding of a skilled artisan *at the time of the patent application*, not that of a court viewing matters *post hoc*." 134 S. Ct. at 2130 (first emphasis added). The critique was thus alert to the way that the Federal Circuit standard's *substantive criterion*, "tolerat[ing] imprecision just short of that rendering a claim 'insolubly ambiguous,'" could "diminish the definiteness requirement's public-notice function . . . ." *Id.* The risk of hindsight bias suggests, in addition, that the Federal Circuit standard's *timing* could diminish a patentee's legitimate flexibility in crafting claim language, even as it diminished public notice from the other direction.

171. *See supra* notes 57–62 and accompanying text.

172. There are varied ways to think about sameness in this context. One proxy to consider is a patent's assigned Patent Office class. *See infra* note 232.

the skilled reader, for the same or a substantially similar feature, than do the actual words or phrases used in the claim under review. To ensure that the prior patents establish language usage that was genuinely available to the patentee, the focal category of prior art should be patents from more than a year before the application for the patent in suit was filed.[173]

• Second, to rebut the accused's evidence of more clearly delineative, actually available claim language from the relevant art and time, the patentee must come forward with evidence showing that, at the time the patent was drafted, one skilled in the art could not reasonably be expected to have better reduced the risk of reader confusion, either by (a) using the better-informing word or phrase from the prior art, or by (b) expressly defining the challenged word or phrase that was actually used, such that the resulting claim language (and any patentee-drafted definition supporting it) would have been at least as clearly informative as the word or phrase from the prior art.

• Third, the district court, making subsidiary findings to resolve any genuine issues of material fact about usage, resolves the ultimate legal question whether the actual claim term is fatally indefinite, given the alternative established usages or definition-writing options available to the patentee at the time the claim was written and the patent application was filed.

It bears noting that this process of proof, though modeled on *Festo*, would fully accommodate the evidence and reasoning in the Federal Circuit's 2016 *Polar Electro* decision using factual findings about evidence of customary usage from the prior art to ground an indefiniteness judgment.[174] Indeed, the more tailored process just sketched is an improvement over more haphazard approaches that may allow for the sound evidentiary process in *Polar Electro* but do not

173. All such patents are unquestionably prior art, both under the 1952 Patent Act that prevailed before 2011, 35 U.S.C. § 102(b) (2006), and the new Patent Act as amended in 2011 by the America Invents Act, 35 U.S.C. §§ 102(a)(1), 102(b)(1) (2012). *See generally* MUELLER, *supra* note 4, at 253–56 (discussing the old § 102(b)), 317–19 (discussing the new § 102). I have no strong view, at the moment, about whether or not English-language patents from other jurisdictions—such as the U.K., Canada, Australia, or New Zealand—should be included. Language usage in one or more of these countries differs enough from that of the U.S. that I worry that including their patents, although formally eligible as prior art under 35 U.S.C. § 102 (old and new), could skew the indefiniteness inquiry here toward false positives. Given that U.S. patents must be in English, however, 37 C.F.R. §§ 1.52(b)(ii), (d)(1) (2016), it strikes me as necessary to limit the prior art evidence of usage to patent documents written in English as an original matter (even though these foreign-language patents, too, are formally prior art under 35 U.S.C. § 102 (old and new)).

174. *See supra* notes 101–11 and accompanying text.

ensure it. And that would be so even without the more rigorously derived evidence of customary usage that corpus analysis can provide.

B. The Examination Process

The Patent Office receives, at present, from about 500,000 to 600,000 utility patent applications per year,[175] each one of which contains about two or three independent claims.[176] Once issued, a patent can be licensed, or a third party threatened with enforcement, without the patent's validity ever being tested in court.[177] Moreover, during examination—in contrast to litigation—an indefiniteness problem can be corrected,[178] by the simple expedient of a claim amendment or "a separate definition of the unclear language."[179] It is important, then, to ensure that the Patent Office, no less so than the courts, targets the *Nautilus* standard of reasonable certainty.

The Patent Office examines a patent application for compliance with all the formal and substantive requirements of the Patent Act,[180] including claim definiteness.[181] The Federal Circuit's recent *Packard* decision provides the procedural framework: It is the patent examiner's responsibility, in the first instance, to identify words or phrases in the applicant's claims that are unclear and to describe why they are unclear,[182] to "provid[e] the applicant a well-grounded identification of [any] clarity problems."[183] After that, "if the applicant does not adequately respond to that prima facie case," the examiner can "confirm

175. *US Patent Activity Years 1790 to the Present*, USPTO, https://www.uspto.gov/web/offices/ac/ido/oeip/taf/h_counts.htm (last visited Sept. 8, 2017).

176. *See* Dennis Crouch, *Independent Patent Claims*, PATENTLYO (June 24, 2015), https://patentlyo.com/patent/2015/06/independent-patent-claims.html; Dennis Crouch, *Average Number of Independent Claims Per Patent*, PATENTLYO (Jan. 25, 2014), https://patentlyo.com/patent/2014/01/average-independent-patent.html.

177. *See generally* Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U. L. REV. 1495, 1500–06 (2001) (discussing, broadly, "what patentees do with their patents").

178. *See* Halliburton Energy Servs., Inc. v. M–I LLC, 514 F.3d 1244, 1255 (Fed. Cir. 2008) (observing that "it is highly desirable that patent examiners demand that applicants [resolve any ambiguity in a patent claim] in appropriate circumstances so that the patent can be amended during prosecution rather than attempting to resolve the ambiguity in litigation").

179. *In re* Packard, 751 F.3d 1307, 1311 (Fed. Cir. 2014).

180. 35 U.S.C. § 131 (2012); 37 C.F.R. § 1.104(a)(1) (2016).

181. *See Packard*, 751 F.3d at 1311; *see also* MPEP § 2173 (9th ed. Rev. 7, Nov. 2015) (providing guidance to examiners).

182. *See Packard*, 751 F.3d at 1312–13.

183. *Id.* at 1313.

that rejection on the substantive basis of having failed to meet the requirements of § 112(b)."[184]

A critical facet of this interactive process, then, is that the patent examiner puts an unclear word or phrase in the context of the relevant prior art and the patterns of language usage that prior art references reflect. As *Packard* puts it, the examiner scrutinizes claim language for clarity from "the perspective of one of ordinary skill in the art."[185] *Packard*, moreover, casts the clarity standard for the examination stage in much the same terms as *Nautilus*, "invok[ing a] standard of *reasonable* precision in the use of language in the context of the circumstances."[186] Judging reasonable certainty, to those in a given art at a given time, can certainly include evidence of usage in prior art patents.[187] No doubt an experienced examiner, having analyzed many (in some instance, *very* many) claims in the relevant art, provides an important part of this language-usage context simply by dint of long practice. We can also augment the examiner-provided context with the tools of corpus linguistics. Indeed, a great benefit of the scale of computer-analyzed usage corpora is that they provide usage data with greater accuracy, speed, and comprehensiveness than an unaided person can achieve.

C. The Corpus Linguistics Evidence

"Corpus linguistics has," to date, "had a very modest impact on legal practice and scholarship,"[188] though the scholarship landscape is

184. *Id.* at 1312; *see also id.* at 1313 ("Given the role of the applicant in the process, it is a reasonable implementation of the examination responsibility, as applied to § 112(b), for the USPTO, upon providing the applicant a well-grounded identification of clarity problems, to demand persuasive responses on pain of rejection.").

185. *Id.* at 1312.

186. *Id.* at 1313 (emphasis added). Additionally, the *Packard* court expressly rejects the goal of perfect clarity: "The requirement, applied to the real world of modern technology, does not contemplate in every case a verbal precision of the kind found in mathematics. Nor could it do so in a patent system that actually works, in practice, to provide effective protection for modern-day inventions." *Id.*

187. *See* MPEP § 2173.02 ¶ II (9th ed. Rev. 7, Nov. 2015) ("Definiteness of claim language must be analyzed, not in a vacuum, but in light of," among other things, "(B) The teachings of the prior art[.]").

188. Lee J. Strang, *How Big Data Can Increase Originalism's Methodological Rigor: Using Corpus Linguistics to Reveal Original Language Conventions*, 50 U.C. DAVIS L. REV. 1181, 1203 (2017). Indeed, "[t]his subject is so new to legal circles that the first conference on the subject [of law and corpus linguistics] did not occur until Spring, 2016." *Id.* As of early September 2017, the Westlaw JLR database contains only six articles the titles of which contain the phrase "corpus linguistics," the earliest of which is dated 2011. WESTLAW, http://www.westlaw.com (follow "Advanced" hyperlink, search "corpus linguistics" in Name/Title field, and narrow by secondary

changing, especially in the neighboring areas of statutory interpretation[189] and constitutional interpretation.[190] One catalyzing resource has been Google's Ngram viewer,[191] an offshoot of the Google Books Library Project.[192] Lawyers and legal scholars can now see that they have much to gain from the decades of experience that corpus linguistics experts can offer, including in the creation and use of software analysis tools. The fact that these tools and methods can dovetail smoothly with extremely familiar, long-available computer-assisted legal research tools, such as Westlaw and Lexis, makes it that much easier for lawyers and legal scholars to consider the full value of corpus linguistics evidence.[193]

---

sources).

189. *See* Stephen C. Mouritsen, *Hard Cases and Hard Data: Assessing Corpus Linguistics as an Empirical Path to Plain Meaning*, 13 COLUM. SCI. & TECH. L. REV. 156, 204–05 (2011); Stephen C. Mouritsen, *The Dictionary Is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning*, 2010 BYU L. REV. 1915, 1919 (2010); Neal Goldfarb, Words, Meanings, Corpora: A Lawyer's Introduction to Meaning in the Framework of Corpus Linguistics 43–44 (Jan. 26, 2017) (unpublished discussion draft), https://ssrn.com/abstract=2907485; Zachary D. Smith, United They Hold, Divided They Might Fail: A Corpus Linguistics Analysis of the U.S. Supreme Court's Recent Ordinary Meaning Cases 4–5 (Dec. 18, 2015) (unpublished student article), https://ssrn.com/abstract=2781672; Lawrence M. Solan & Tammy A. Gales, Finding Ordinary Meaning in Law: The Judge, the Dictionary or the Corpus? 1 (Oct. 10, 2016) (unpublished article), https://ssrn.com/abstract=2850703.

190. *See* Jennifer L. Mascott, *Who Are 'Officers of the United States'?*, 70 STAN. L. REV. (forthcoming 2018) (manuscript at 15), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2918952; James C. Phillips et al., *Corpus Linguistics & Original Public Meaning: A New Tool to Make Originalism More Empirical*, 126 YALE L.J.F. 20, 26 (2016); Lawrence M. Solan, *Can Corpus Linguistics Help Make Originalism Scientific?*, 126 YALE L.J.F. 57, 57–58 (2016); Strang, *supra* note 188, at 1202.

191. *Ngram Viewer*, GOOGLE BOOKS, https://books.google.com/ngrams (last visited Sept. 8, 2017).

192. *See id.*; *see also About Google Books*, GOOGLE BOOKS, https://www.google.com/intl/en/googlebooks/about/index.html (last visited Sept. 8, 2017). For a key paper describing the technical features, and introducing the scholarly potential, of the Google Ngram viewer project, see Jean-Baptiste Michel et al., *Quantitative Analysis of Culture Using Millions of Digitized Books*, 331 SCIENCE 176 (2011). A group of law professors and computer scientists created a similar pilot project for analyzing the full-text corpus of U.S. Supreme Court decisions from 1791 to 2005. *See* Daniel Martin Katz et al., Legal N-Grams? A Simple Approach to Track the 'Evolution' of Legal Language (Dec. 13, 2011) (unpublished conference proceeding), https://ssrn.com/abstract=1971953. Legal scholars who want to explore the corpus linguistics literature from inside can consult, in addition to the ROUTLEDGE HANDBOOK and MCENERY & HARDIE books already cited, *supra* note 33, the following helpful sources: PAUL BAKER, USING CORPORA IN DISCOURSE ANALYSIS (2006); SUSAN HUNSTON, CORPORA IN APPLIED LINGUISTICS (2002); GRAEME KENNEDY, AN INTRODUCTION TO CORPUS LINGUISTICS (1998).

193. *See* Phillips et al., *supra* note 190, at 23 ("Lawyers use corpora on a daily basis. In a sense, Google and Westlaw or Lexis are corpora."); Solan & Gales, *supra* note 189, at 4 ("Judges and legal scholars are accustomed to searching large banks of information for relevant instances of a word's meaning. That process characterizes the case method generally.").

"[T]he aim" of corpus linguistics is simply "the analysis and description of *language use, as realised in text(s)*."[194] Indeed, "[l]inguists have always used the word *corpus* to describe a collection of naturally occurring examples of language . . . collected for linguistic study."[195] The falling price and rising power of computing technology over the last few decades, however, "brought about a radical change" in the field: "Corpus linguistics is thus now inextricably linked to the computer, which has introduced incredible speed . . . and the ability to handle huge amounts of data."[196] The word *corpus*, as a result, "has been reserved for collections of texts (or parts of text) that are stored and accessed electronically."[197] To set the stage for exploring how the indefiniteness inquiry might avail itself of the fruits of corpus analysis, it is useful to describe the conventional types of corpora and analytical tools that linguists have developed.

"A corpus is defined in terms of both its form and its purpose."[198] In particular, "[t]he corpus is stored in such a way that it can be studied non-linearly, and both quantitatively and qualitatively."[199] Linguists differentiate corpora with three key distinctions. With respect to the texts comprising the corpus, a corpus can be either *general*, a "corpus of texts of many types," or *specialized*, a "corpus of texts of a particular type."[200] The Westlaw SCT database is specialized, comprising only U.S. Supreme Court opinions and orders, whereas the Corpus of Contemporary American English (COCA) is general.[201] With respect to the time period the stored texts reflect, a corpus can be either *static*, fixed in content from a given year or span of years, or *dynamic*, also known as

194. Bonelli, *supra* note 33, at 18–19.

195. HUNSTON, *supra* note 192, at 2; *see also* KENNEDY, *supra* note 192, at 1 ("In the language sciences a corpus is a body of written text or transcribed speech which can serve as a basis for linguistic analysis and description.").

196. KENNEDY, *supra* note 192, at 5.

197. HUNSTON, *supra* note 192, at 2.

198. *Id.*

199. *Id.*

200. *Id.* at 14; *see also* KENNEDY, *supra* note 192, at 19–20 (contrasting general and specialized corpora).

201. Solan & Gales, *supra* note 189, at 4 ("In the past two decades, large corpora of general American English have been made available to the public as a research tool. The most prominent have been developed by linguistic researchers at Brigham Young University[, including] the Corpus of Contemporary American English (COCA) . . . ."); *Search*, CORPUS OF CONTEMPORARY AMERICAN ENGLISH, http://corpus.byu.edu/coca/ ("The Corpus of Contemporary American English (COCA) is the largest freely-available corpus of English, and the only large and balanced corpus of American English. . . . The corpus contains more than 520 million words of text (20 million words each year 1990-2015) and it is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts.") (last visited Sept. 29, 2017).

a *monitor* corpus, receiving new texts on an ongoing basis to reflect ongoing usage change.[202] The Helsinki Corpus, for example, is static, "consist[ing] of texts from 700 [c.e.] to 1700 [c.e.]";[203] whereas the Westlaw SCT database is dynamic, growing each time the Supreme Court issues another opinion or order. Finally, with respect to the processing of texts for further analysis, a corpus can be either *tagged*, "refer[ring] to the addition of a code to each word in a corpus, indicating the part of speech,"[204] or *raw*, "contain[ing] a minimum of annotations."[205] Part-of-speech tags enable some of the more illuminating analyses that corpus linguistics provides,[206] and it is precisely this tagging that is absent from the tools lawyers commonly use: "Westlaw or Google are essentially raw corpora,"[207] having neither part-of-speech tags nor the analytical tools that depend on them. More pertinent here, although Westlaw has a corpus of the texts of issued U.S. patents and published U.S. patent applications—called the "United States Patents & Applications" database—it does not tag the words in these texts for parts of speech. If, for example, one wanted to analyze how U.S. patents issued in 2016 use the word *light* in the Abstract portion of the patent,[208] one could retrieve all the issued U.S. patents that contain the string "light" in their respective Abstracts (using the search query < ISSUEDATE(aft 12-31-2015 & bef 01-01-2017) & ABSTRACT(light) >). There are 17,253 such patents.[209] To group the uses of *light* as a noun (as in U.S. Patent No. 9,341,321, which refers to an "assembly [that] includes a light source and a lens") separately from the uses of *light* as a verb (as in U.S. Patent No. 9,255,677, which refers to "a first beam of light that primarily lights up a first portion"), one must sort the

202. HUNSTON, *supra* note 192, at 16; KENNEDY, *supra* note 192, at 22, 60–62.

203. HUNSTON, *supra* note 192, at 16; KENNEDY, *supra* note 192, at 38 (discussing the Helsinki Corpus).

204. HUNSTON, *supra* note 192, at 18.

205. KENNEDY, *supra* note 192, at 21.

206. *See* HUNSTON, *supra* note 192, at 18–19, 80–83 (providing example analyses and findings that depend upon using a tagged corpus).

207. Phillips et al., *supra* note 190, at 24.

208. A patent applicant must provide an Abstract, which summarizes "the technical disclosure in the [patent] specification . . . as concise[ly] as the disclosure permits," in order "to enable the [Patent] Office and the public generally to determine quickly from a cursory inspection the nature and gist of the technical disclosure." 37 C.F.R. § 1.72(b) (2016).

209. U.S. Patents Issued in 2016 That Use the Word Light, WESTLAW, http://westlaw.com (follow "Intellectual Property" hyperlink, then "Patents & Applications" hyperlink, and then "United States Patents & Applications" hyperlink and search < ISSUEDATE(aft 12-31-2015 & bef 01-01-2017) & ABSTRACT(light) >).

flagged texts manually. Traditional corpus analysis tools, by contrast, automate the task.[210]

Again, corpora are designed not to store texts for later reading (as does an archive or library), but to extract data across the range of texts, thus providing far richer, more reliable descriptions of semantic usage. "[T]he main argument in favour of using a corpus," Professor Hunston urges, "is that it is a more reliable guide to language use than native speaker intuition is."[211] Popular corpus linguistics tools thus "process data from a corpus in three [main] ways: showing frequency, phraseology, and collocation."[212]

Frequency data can help one compare usage of both different words or phrases within a corpus and of the same words or phrases across corpora, as well as describe changes in relative frequency of usage over time.[213] For example, using the Google Ngram viewer, one can use the wildcard query term, *, to see that in the *American English* corpus of books, the five most common 1-grams[214] in the decade spanning 1999 to 2008, inclusive, are—in descending order—*the* (around 4.4% of all 1-grams in the corpus, by year), *of* (around 2.6% of all the 1-grams), *and* (around 2.3% of all the 1-grams), *to* (around 2.0% of all the 1-grams), and *a* (around 1.5% of all the 1-grams). These results are quite consistent with top word counts for general English language corpora.[215] With the wildcard term, the Google Ngram viewer shows the ten most common items. Figure 1, below, graphs the search just described. The usage frequencies are stable over the decade.

---

210. *See, e.g.*, HUNSTON, *supra* note 192, at 81 (contrasting the common collocates of the verb *light* from those of the noun *light* and the adjective *light*).

211. *Id.* at 20.

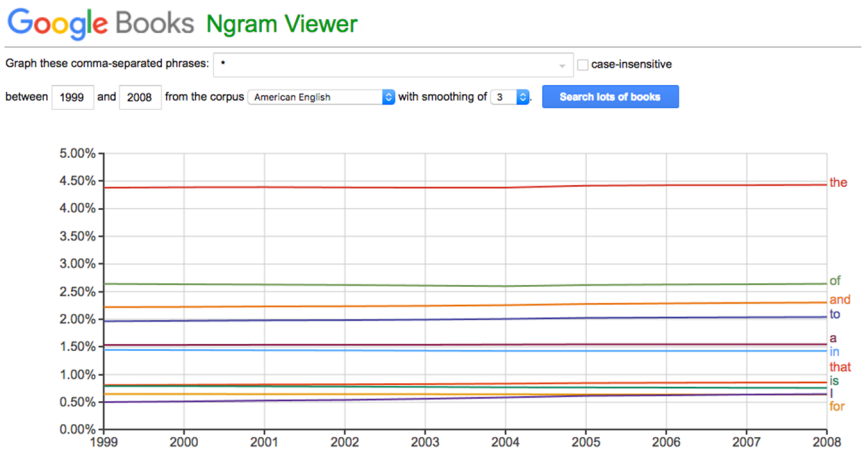
212. *Id.* at 3.

213. *See id.* at 3–9 (discussing use of frequency data and examples); KENNEDY, *supra* note 192, at 97–107 (same).

214. "A 1-gram is a string of characters uninterrupted by a space . . . ." Michel et al., *supra* note 192, at 176. A 1-gram can be a word, such as *light*, or a typo, such as *lihgt*. *Id.* "An *n*-gram is a sequence of 1-grams, such as the phrase[] 'stock market' (a 2-gram) . . . ." *Id.*

215. *See* HUNSTON, *supra* note 192, at 4 (table 1.1); KENNEDY, *supra* note 192, at 98 (table 3.3).

Figure 1. Ten most common 1-grams, 1999-2008, *American English* corpus



Contrast the stability of the most common English words to sharp turns in usage. For example, consider Justice Holmes' opinion rejecting a First Amendment challenge to criminalizing the distribution of Socialist party leaflets, *Schenck v. United States*.[216] The case is famed for establishing a "clear and present danger" test.[217] The Ngram viewer plot for the 3-gram *[adjective] and present*, for the period from 1808 to 2008, shows that the most common such phrase in the *American English* corpus is *past and present*, for two centuries. (See Figure 2, below.) But in the 1930s, the phrase *clear and present* surges up, peaking in 1966. Could this be the influence of *Schenck*'s clear and present danger test? The Ngram viewer plot for the 4-gram *clear and present [word]* is highly suggestive. (See Figure 3, below.) The most common target 4-gram in the corpus, by far, is *clear and present danger*, which appears from nowhere at around the same time as *Schenck*, in 1919, and surges up from the late 1930s through the 1950s. It has the same mid-1960s peak. Usage frequency can take sudden turns, as this example illustrates.

216. 249 U.S. 47 (1919).

217. *Id.* at 52 ("The question in every case is whether the words used are used in such circumstances and are of such a nature as to create *a clear and present danger* that they will bring about the substantive evils that Congress has a right to prevent.") (emphasis added). On the case and its legacy, see Mark Tushnet, *The Hughes Court and Radical Political Dissent: The Cases of Dirk De Jonge and Angelo Herndon*, 28 GA. ST. U. L. REV. 333, 334–36 (2012).

Figure 2. Ten most common 3-grams of the form *[adjective] and present*, 1808-2008, *American English* corpus

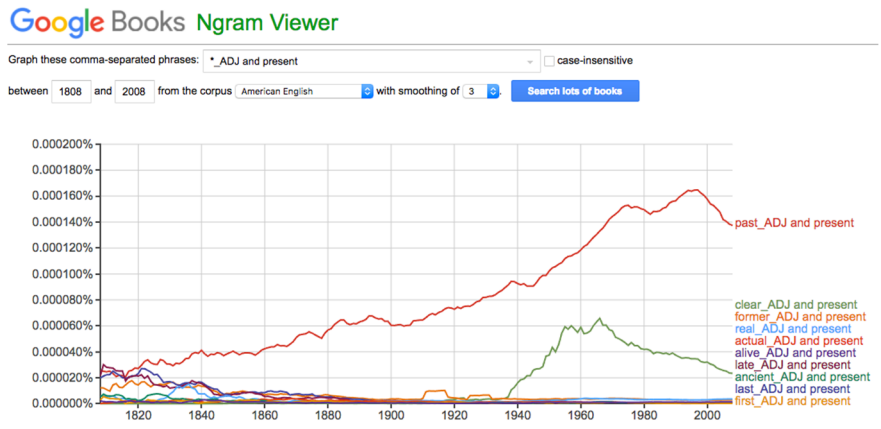
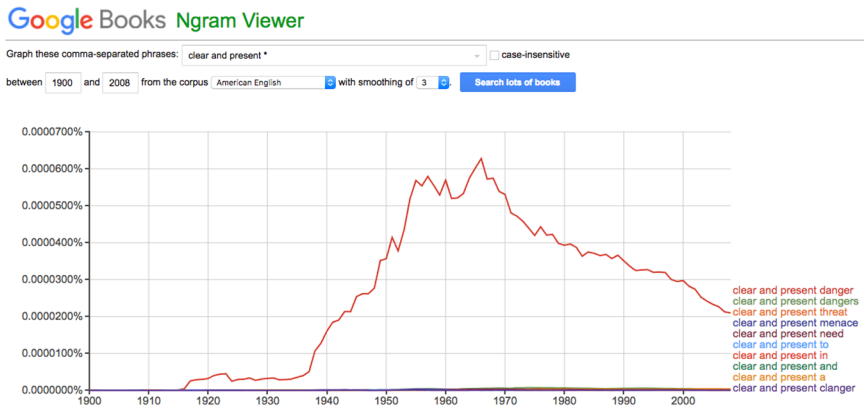


Figure 3. Ten most common 4-grams of the form *clear and present [word]*, 1908-2008, *American English* corpus

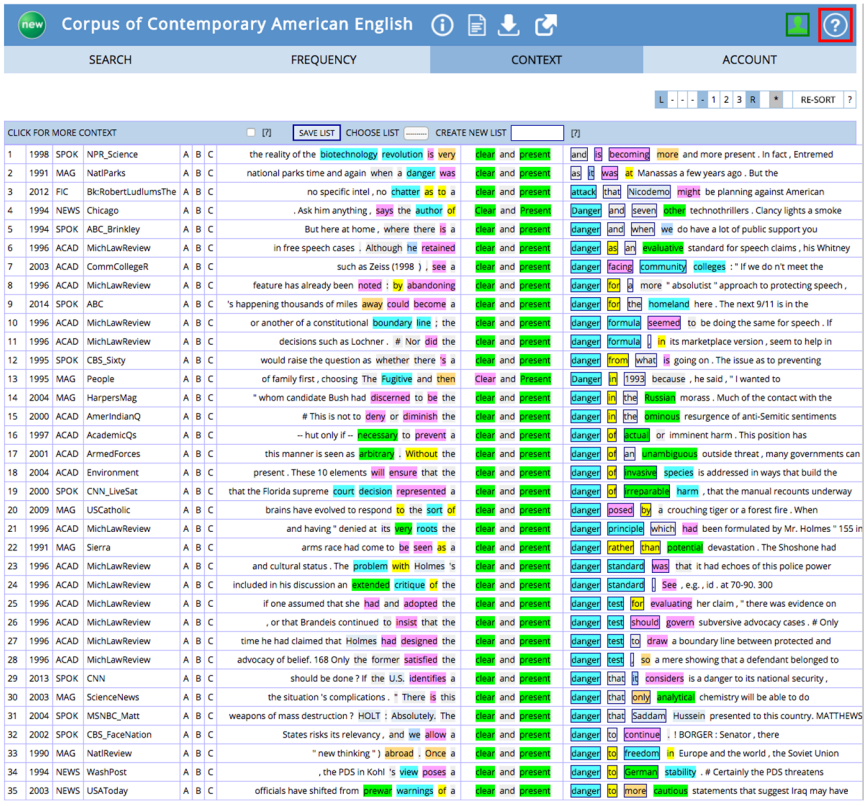


One can also analyze a corpus using concordance lines, which reveal common phraseology.[218] "A concordance is a formatted version or display of all the occurrences or tokens of a particular type in a corpus."[219] Much like a summary list of hits in a Westlaw search, which shows the target word or phrase with some surrounding text, a set of concordance lines shows the target item in context, pulling hits from texts within the corpus.[220] The most common concordance-line format is known as KWIC, for Key Word In Context, which "has the keyword in

218. *See* HUNSTON, *supra* note 192, at 9–12, 38–66 (discussing concordance analysis).
219. KENNEDY, *supra* note 192, at 247.
220. *Id.* at 251.

the centre of the line . . . but with more context [on] each side of the keyword."[221] Figure 4, below, shows some concordance lines for the phrase *clear and present* in the COCA, with the context words to the right of the phrase listed in alphabetical order. The high frequency of the word *danger* as the next in line is unmistakable in this corpus, just as it is in the Google Ngram viewer.

Figure 4. Concordance lines for *clear and present*



1	1998	SPOK	NPR_Science	A	B	C	the reality of the biotechnology revolution is very	clear	and	present	danger	is becoming more and more present . In fact , Entremed
2	1991	MAG	NatlParks	A	B	C	national parks time and again when a danger was	clear	and	present	danger	as it was at Manassas a few years ago . But the
3	2012	FIC	BkcRobertLudlumsThe	A	B	C	no specific intel , no chatter as to a	clear	and	present	danger	attack that Nicodemo might be planning against American
4	1994	NEWS	Chicago	A	B	C	. Ask him anything , says the author of	Clear	and	Present	Danger	and seven other technothrillers . Clancy lights a smoke
5	1994	SPOK	ABC_Brinkley	A	B	C	But here at home , where there is a	clear	and	present	danger	and when we do have a lot of public support you
6	1996	ACAD	MichLawReview	A	B	C	in free speech cases . Although he retained	clear	and	present	danger	as an evaluative standard for speech claims , his Whitney
7	2003	ACAD	CommCollegeR	A	B	C	such as Zeiss (1998) , see a	clear	and	present	danger	facing community colleges : " If we do n't meet the
8	1996	ACAD	MichLawReview	A	B	C	feature has already been noted : by abandoning	clear	and	present	danger	for a more " absolutist " approach to protecting speech ,
9	2014	SPOK	ABC	A	B	C	's happening thousands of miles away could become a	clear	and	present	danger	for the homeland here . The next 9/11 is in the
10	1996	ACAD	MichLawReview	A	B	C	or another of a constitutional boundary line ; the	clear	and	present	danger	formula seemed to be doing the same for speech . If
11	1996	ACAD	MichLawReview	A	B	C	decisions such as Lochner . # Nor did the	clear	and	present	danger	formula , in its marketplace version , seem to help in
12	1995	SPOK	CBS_Sixty	A	B	C	would raise the question as whether there 's a	clear	and	present	danger	from what is going on . The issue as to preventing
13	1995	MAG	People	A	B	C	of family first , choosing The Fugitive and then	Clear	and	Present	Danger	in 1993 because , he said , " I wanted to
14	2004	MAG	HarpersMag	A	B	C	" whom candidate Bush had discerned to be the	clear	and	present	danger	in the Russian morass . Much of the contact with the
15	2000	ACAD	AmerIndianQ	A	B	C	# This is not to deny or diminish the	clear	and	present	danger	in the ominous resurgence of anti-Semitic sentiments
16	1997	ACAD	AcademicQs	A	B	C	-- hut only if -- necessary to prevent a	clear	and	present	danger	of actual or imminent harm . This position has
17	2001	ACAD	ArmedForces	A	B	C	this manner is seen as arbitrary . Without the	clear	and	present	danger	of an unambiguous outside threat , many governments can
18	2004	ACAD	Environment	A	B	C	present . These 10 elements will ensure that the	clear	and	present	danger	of invasive species is addressed in ways that build the
19	2000	SPOK	CNN_LiveSat	A	B	C	that the Florida supreme court decision represented a	clear	and	present	danger	of irreparable harm , that the manual recounts underway
20	2009	MAG	USCatholic	A	B	C	brains have evolved to respond to the sort of	clear	and	present	danger	posed by a crouching tiger or a forest fire . When
21	1996	ACAD	MichLawReview	A	B	C	and having " denied at its very roots the	clear	and	present	danger	principle which had been formulated by Mr. Holmes " 155 in
22	1991	MAG	Sierra	A	B	C	arms race had come to be seen as a	clear	and	present	danger	rather than potential devastation . The Shoshone had
23	1996	ACAD	MichLawReview	A	B	C	and cultural status . The problem with Holmes 's	clear	and	present	danger	standard was that it had echoes of this police power
24	1996	ACAD	MichLawReview	A	B	C	included in his discussion an extended critique of the	clear	and	present	danger	standard .] See , e.g. , id . at 70-90. 300
25	1996	ACAD	MichLawReview	A	B	C	if one assumed that she had and adopted the	clear	and	present	danger	test for evaluating her claim , " there was evidence on
26	1996	ACAD	MichLawReview	A	B	C	, or that Brandeis continued to insist that the	clear	and	present	danger	test should govern subversive advocacy cases . # Only
27	1996	ACAD	MichLawReview	A	B	C	time he had claimed that Holmes had designed the	clear	and	present	danger	test to draw a boundary line between protected and
28	1996	ACAD	MichLawReview	A	B	C	advocacy of belief. 168 Only the former satisfied the	clear	and	present	danger	test] so a mere showing that a defendant belonged to
29	2013	SPOK	CNN	A	B	C	should be done ? If the U.S. identifies a	clear	and	present	danger	that it considers is a danger to its national security ,
30	2003	MAG	ScienceNews	A	B	C	the situation 's complications . " There is this	clear	and	present	danger	that only analytical chemistry will be able to do
31	2004	SPOK	MSNBC_Matt	A	B	C	weapons of mass destruction ? HOLT : Absolutely. The	clear	and	present	danger	that Saddam Hussein presented to this country. MATTHEWS
32	2002	SPOK	CBS_FaceNation	A	B	C	States risks its relevancy , and we allow a	clear	and	present	danger	to continue . ! BORGER : Senator , there
33	1990	MAG	NatlReview	A	B	C	" new thinking ") abroad . Once a	clear	and	present	danger	to freedom in Europe and the world , the Soviet Union
34	1994	NEWS	WashPost	A	B	C	, the PDS in Kohl 's view poses a	clear	and	present	danger	to German stability . # Certainly the PDS threatens
35	2003	NEWS	USAToday	A	B	C	officials have shifted from prewar warnings of a	clear	and	present	danger	to more cautious statements that suggest Iraq may have

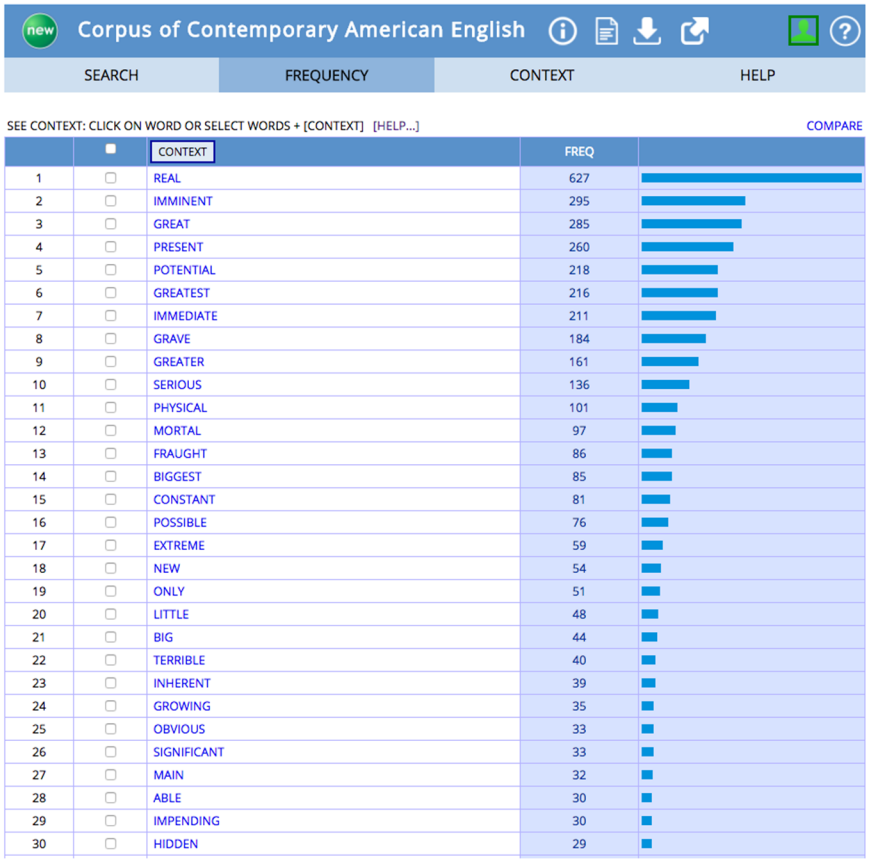
The phrase *clear and present danger* also illustrates the third analytical tool common in corpus linguistics—namely, collocation, "a tendency to occur in the company of other words in certain contexts, e.g. *pouring rain*, *statistically significant*, *intrinsic value*, *strong tendency*."[222] In the COCA, for example, the top 40 adjectives that occur within two

221. *Id.*

222. *Id.* at 108; *see also* HUNSTON, *supra* note 192, at 12–13, 68–79 (discussing collocation data and analysis).

words before *danger* are shown in Figure 5. The word *real* is the top collocate, and it appears within two words before *danger* more than twice as often as the second collocate, *imminent*. When taken together, these three tools—frequency data, KWIC lines, and collocation data—enable a corpus linguist to describe usage, common *and uncommon*, with far greater precision than one could otherwise achieve. And it is uncommon usage, of course, that can undermine the reasonably certain notice that *Nautilus* requires from patent claims.

Figure 5. Adjectival collocates for the noun *danger*, within two words before



		CONTEXT	FREQ
1	☐	REAL	627
2	☐	IMMINENT	295
3	☐	GREAT	285
4	☐	PRESENT	260
5	☐	POTENTIAL	218
6	☐	GREATEST	216
7	☐	IMMEDIATE	211
8	☐	GRAVE	184
9	☐	GREATER	161
10	☐	SERIOUS	136
11	☐	PHYSICAL	101
12	☐	MORTAL	97
13	☐	FRAUGHT	86
14	☐	BIGGEST	85
15	☐	CONSTANT	81
16	☐	POSSIBLE	76
17	☐	EXTREME	59
18	☐	NEW	54
19	☐	ONLY	51
20	☐	LITTLE	48
21	☐	BIG	44
22	☐	TERRIBLE	40
23	☐	INHERENT	39
24	☐	GROWING	35
25	☐	OBVIOUS	33
26	☐	SIGNIFICANT	33
27	☐	MAIN	32
28	☐	ABLE	30
29	☐	IMPENDING	30
30	☐	HIDDEN	29

Let us begin, then, with the premise that a patent claim term that embodies *atypical* usage is—absent a bespoke definition—less likely to inform the person of skill in the art of the claim's scope with reasonable certainty. How does it benefit the definiteness inquiry to look to evidence rooted in corpus linguistics resources? If a patent examiner or

federal judge is to evaluate how typical, or atypical, a bit of claim language is, for a given art at a given time, there are only two choices, broadly speaking: Estimate typicality using introspection and intuition, or estimate typicality using data about usage patterns in an appropriate corpus. The latter is plainly more reliable, as Professor Hunston explains:

> Although a native speaker has experience of very much more language than is contained in even the largest corpus, much of that experience remains hidden from introspection. . . .
>
> Intuition is a poor guide to at least four aspects of language: collocation, frequency, prosody and phraseology. . . .
>
> It is almost impossible to be conscious of the relative frequency of words, phrases and structures except in very general terms (anyone might guess that *take* is a more frequent verb than *disseminate*, but it is difficult to guess whether *fare* or *fantasy* is more frequent[[223]]). . . .
>
> Although native speakers can often recognize if a phraseology is unusual, articulating the nature of the atypicality may be more difficult.[224]

The recent *Polar Electro* case illustrates the benefit to a definiteness inquiry that usage evidence from the prior art provides.[225] Even so, given the fact that the key evidence in the case derived from just ten prior art references,[226] one could wonder whether the outcome reflects anecdata more than it reflects data.

Patent law cannot yet harness, fully, the sources and methods of corpus linguistics. Current corpora straddle the required evidence: The databases of U.S. patents, such as Westlaw's or Google's,[227] do not have tagged content or the analytical tools that presuppose part-of-speech tags; and the corpora with tagged content and the analytical tools that use tags, such as the Google Ngram viewer or the COCA, do not contain U.S. patents. This shortcoming in existing corpora of U.S. patents is not surprising, given that lawyers have not—at least up to now—seen themselves as corpus linguists. "A corpus is always designed for a

---

223. In the Google Ngram viewer's *American English* corpus, it turns out that the answer depends on the time period as to which one asks. An Ngram graph is in the Appendix.

224. HUNSTON, *supra* note 192, at 20–21 (citation omitted).

225. *See supra* notes 101–11 and accompanying text.

226. Icon Health & Fitness, Inc. v. Polar Electro Oy, 656 F. App'x 1008, 1015 (Fed. Cir. 2016) (reporting that "Polar's expert proffered ten extrinsic prior art patents and text books").

227. Google's U.S. patent database is found at https://patents.google.com/.

particular purpose, and the type of corpus will depend on its purpose."[228] Even with the corpora we have now, however, I think we can see clearly enough how much more robust our definiteness inquiries would be if we had, in effect, a Google Ngram viewer plus Google Patents mash-up.

Consider, for example, the patent invalidated for indefiniteness in *Halliburton Energy Services, Inc. v. M-I LLC*,[229] U.S. Patent No. 6,887,832. The Federal Circuit held the asserted claims invalid under its pre-*Nautilus* "insolubly ambiguous" standard.[230] The technology in question was oil field drilling fluids, and the fatally indefinite claim language was the phrase "fragile gel drilling fluid."[231] Because the '832 patent's earliest filing date was December 29, 2000,[232] it is unquestionably the case that all U.S. patents with issue dates before December 29, 1999, are prior art to the '832 patent.[233] How often, in patents in this technological field or in patents generally, did the phrase *fragile gel drilling fluid* occur in patent claims, or anywhere in a patent's text? Westlaw's patent database can provide us precise counts, both among patents in the same Patent Office technology class[234] (in this instance, Class 507, "Earth boring, well treating, and oil field chemistry"[235]) and among patents in all technology classes. Table 1, below, presents the counts, both for the full 4-gram and for selected subparts. The fact that the phrase *fragile gel drilling fluid* appeared in *no* prior art U.S. patent claims or patent texts, in Class 507 or in *any* class, could suggest that, at the very least, the phrase is in need of a robust explicit definition in the '832 patent. The patentee had, in fact, provided

228. HUNSTON, *supra* note 192, at 14.

229. 514 F.3d 1244 (Fed. Cir. 2008).

230. *Id.* at 1249–50.

231. *Id.* at 1246.

232. U.S. Patent No. 6,887,832.

233. *See supra* note 173; 35 U.S.C. § 102(b) (2006) (current version at 35 U.S.C. § 102(b) (2012)).

234. Scholars who conduct empirical patent studies have been critical of the U.S. patent classification system, with justification. *See, e.g.*, John R. Allison et al., *Valuable Patents*, 92 GEO. L.J. 435, 472 (2004) ("Because they are designed to assist in narrowly-tailored prior art searches, the government's classifications focus on the functional rather than the conceptual and do so at very low levels of abstraction."). "When a researcher works with an extremely large dataset such that it is not feasible to study each patent in depth," however, "reliance on PTO classifications or International Patent Classifications (IPCs, which the PTO assigns from a concordance based on the PTO's own classifications) may be an unavoidable shortcut." John R. Allison & Lisa Larrimore Ouellette, *How Courts Adjudicate Patent Definiteness and Disclosure*, 65 DUKE L.J. 609, 633 n.74 (2016). The searches I do here certainly query "an extremely large dataset."

235. *Class Numbers and Titles*, USPTO, https://www.uspto.gov/web/patents/classification/selectnumwithtitle.htm (last visited Sept. 29, 2017).

a definition,[236] but the Federal Circuit concluded that the definition in the body of the '832 patent did not actually make the claims sufficiently clear in scope.[237] Had the drafter of the '832 patent, or the Patent Office during the examination process, had before it data of the form that Table 1 provides, perhaps the explicit definition provided in the '832 patent would have gone to greater lengths to delineate claim scope clearly. (I say "perhaps" to concede that hindsight may color one's view of these data.[238])

Table 1. Frequency of *fragile gel drilling fluid* in U.S. patents issued before Dec. 29, 1999

In Class 507/!	**Claim**	**Anywhere**
fragile gel drilling fluid	0	0
fragile gel drilling	0	0
fragile gel	0	2
fragile	8	139
gel	370	2297
drilling fluid	735	1226
comprising! or consisting![239]	13,997	
In All Classes	**Claim**	**Anywhere**
fragile gel drilling fluid	0	0
fragile gel	5	66
drilling fluid	3555	340,239
comprising! or consisting!	5,330,729	

And consider, in addition, what we might learn if we could interrogate a Google Ngram viewer-style corpus of U.S. patents from the relevant time period, either from the same technology class or from all technologies. Parts of speech would be tagged, and we could use wildcard searches to hone in on the adjectives that tend to co-occur with the word *gel*. We do *not* have that corpus and associated tools, so far as I

236. *Halliburton*, 514 F.3d at 1246–47.

237. *Id.* at 1256 ("Halliburton's proposed definition of that term is not sufficiently definite because it does not adequately distinguish the fragileness of the invention from disclosed prior art, it is ambiguous as to whether an upper bound of fragileness is contemplated, and it is ambiguous as to its requisite ability to suspend drill cuttings.").

238. *See supra* note 170.

239. This query is reported because it provides a proxy for the total number of patents in the group searched, which can serve as a rough denominator for estimating frequency. The search term tracks the number of patents because the words *comprising* and *consisting* are the two standard transition words that patent writers use to connect a claim's preamble (or introductory portion) with its body (or main portion). *See* MUELLER, *supra* note 4, at 113–20 (explaining the basic "anatomy" of a U.S. patent claim).

am aware. We can, however, use the existing Google Ngram viewer to make queries that help us appreciate how such data could help ground an exhaustive, objective definiteness inquiry. For example, taking a six-decade snapshot ending in 1999, the *American English* corpus contains the phrase *fragile gel*, but infrequently: at its highest point in the graph in Figure 6 (which was in 1994), it was 0.0000001020% of all 2-grams in the corpus. How does *fragile* fare among all adjectives immediately preceding *gel* in the corpus? It does not appear among the ten most common adjectives in the *[adjective] gel* 2-gram, which are shown in Figure 7 below. And when plotted alongside the frequency line for the sum of all *[adjective] gel* 2-grams, in Figure 8 below, *fragile gel* effectively flatlines at zero. Had an analysis of a U.S. patents corpus produced frequency data along these lines, the great caution suggested by the counts in Table 1, above, would be greater still.

Figure 6. Frequency of *fragile gel*, 1939-1999, *American English* corpus

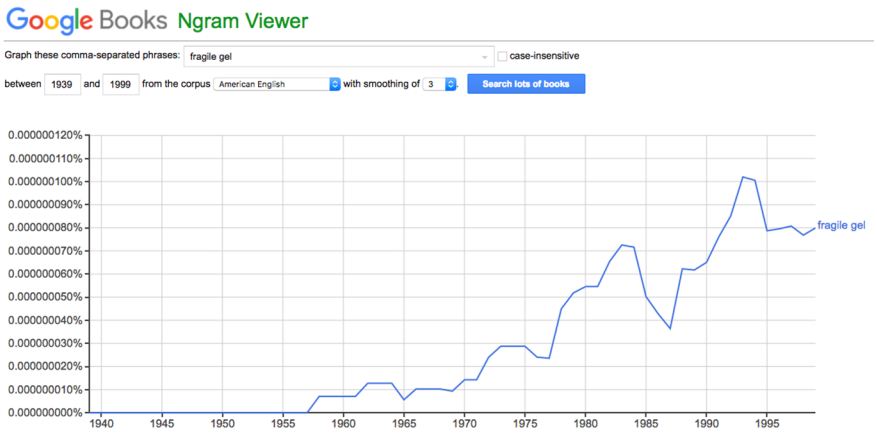


Figure 7. Ten most common 2-grams of the form *[adjective] gel*, 1939-1999, *American English* corpus

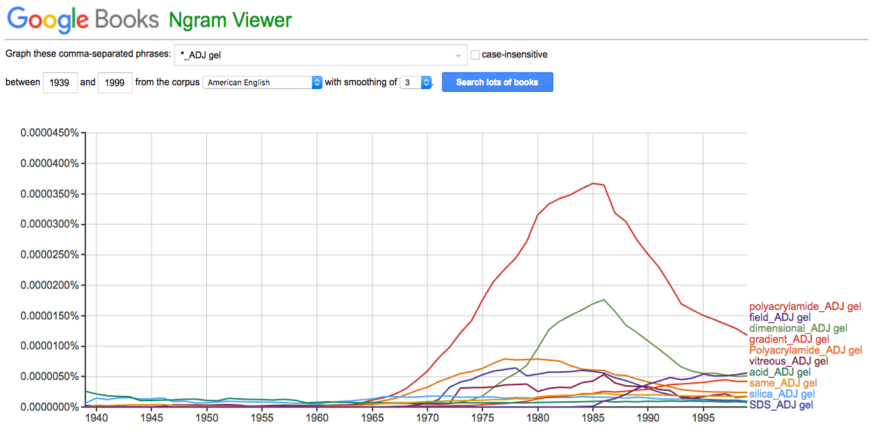
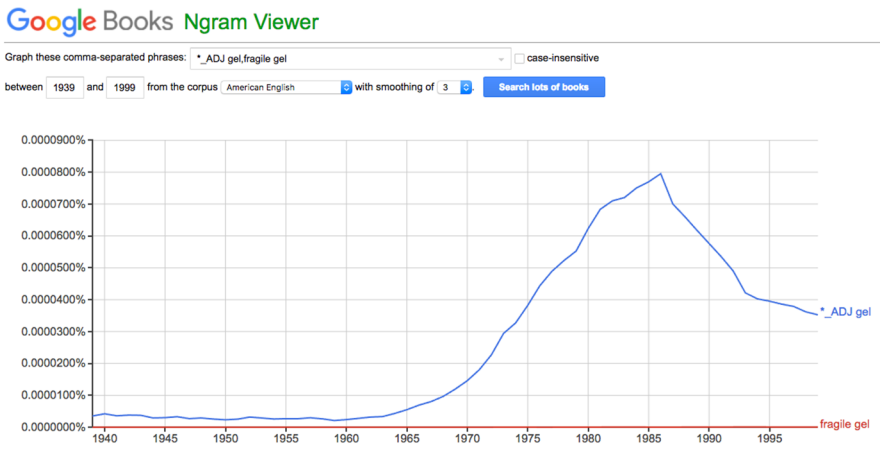


Figure 8. Frequencies of *fragile gel* and **_ADJ gel*, 1939-1999, *American English Corpus*



The frequency data pertaining to the *Nautilus* case tell a different story. In *Nautilus*, the challenged patent, U.S. Patent No. 5,337,753, used the claim term "in spaced relationship with each other" to refer to the spacing of electrodes on a handle that served as a heart monitor.[240] The Supreme Court did not adjudicate the indefiniteness defense on the merits, preferring to send the case back to the Federal Circuit for application of the new "reasonable certainty" standard in the first

240. Nautilus, Inc. v. Biosig Instruments, Inc., 134 S. Ct. 2120, 2125–27 (2014).

instance.[241] On remand, the Federal Circuit, applying the new standard, held—as it had before—that the *spaced relationship* term in the '753 patent did not render the claims fatally indefinite.[242] Because the '753 patent's earliest filing date was June 9, 1992,[243] it is unquestionably the case that all U.S. patents with issue dates before June 9, 1991, are prior art to the '753 patent.[244] Westlaw's patent database can provide us frequency counts for *in spaced relationship with each other* and portions thereof, both among patents in the same technology class (Class 128, "Surgery"[245]) and among patents in all technology classes. Table 2, below, presents the counts, both in and beyond Class 128, and in and beyond the claim language. In contrast to *Halliburton*'s *fragile gel*, the *spaced relationship* of *Nautilus* does not appear quite so atypical. These data do not raise a bright red flag.

Table 2. Frequency of *in spaced relationship with each other* in U.S. patents issued before June 9, 1991

In Class 128/!	**Claim**	**Anywhere**
in spaced relationship with each other	5	44
in spaced relationship with	44	96
in spaced relationship	317	638
spaced relationship	591	1201
relationship with	1354	3233
relationship	4467	16,465
comprising! or consisting![246]	66,605	
In All Classes	**Claim**	**Anywhere**
in spaced relationship with each other	260	560
in spaced relationship	22,180	287,097
spaced relationship	40,206	320,763
comprising! or consisting!	4,483,142	

241. *Id.* at 2131 ("As we have explained, the Federal Circuit invoked a standard more amorphous than the statutory definiteness requirement allows. We therefore follow our ordinary practice of remanding so that the Court of Appeals can reconsider, under the proper standard, whether the relevant claims in the '753 patent are sufficiently definite.").

242. Biosig Instruments, Inc. v. Nautilus, Inc., 783 F.3d 1374, 1376 (Fed. Cir. 2015).

243. U.S. Patent No. 5,337,753.

244. *See supra* note 173; 35 U.S.C. § 102(b) (2006) (current version at 35 U.S.C. § 102(b) (2012)).

245. *Class Numbers and Titles*, USPTO.GOV, https://www.uspto.gov/web/patents/classification/selectnumwithtitle.htm (last visited Sept. 30, 2017).

246. This query is reported because it provides a proxy for the total number of patents in the group searched, which can serve as a rough denominator for estimating frequency.

The foregoing examples illustrate that corpus linguistics sources and methods could provide concrete evidence of demonstrable semantic usage in a given art at a given time, greatly improving the reliability of definiteness analysis in the Patent Office and in the courts. There is, in addition, a corpus linguistics tool that linguists themselves may not find particularly useful, but that the patent law community might very much want. Described functionally, what is needed is a comparator that takes a given patent application (or issued patent), identifies the *n* prior art patents that are most semantically similar in content (e.g., the five most similar, or the ten most similar), and flags in the target patent the claim terms that are the most *unlike* the claim terms in the prior art batch. This tool would generate, in effect, an indefiniteness heat map of the claim terms in a target patent document that merit further attention as potentially atypical, and thus potentially unclear, claim text. Recent developments in the use of semantic analysis to situate all existing U.S. patents in a similarity space, establishing which patents are most similar to which others, may be a large step in the direction I am describing.[247] The Patent Office, for its part, acknowledged two years ago that its own semantic comparator tool, known as "PLUS" (for Patent Linguistic Utility Service), has arguably fallen behind the functionalities available elsewhere.[248]

V. CONCLUSION

To implement the new definiteness standard that *Nautilus* and *Teva* prescribe, both patent examination and patent litigation should shift focus. The lawyer's skill in textual exegesis is helpful, but judging definiteness requires giving greater weight to robust data about the common usage options a patentee had readily available in a given technological field at a given time. Those data, generated with corpus

247. *See* Kenneth A. Younge & Jeffrey M. Kuhn, Patent-to-Patent Similarity: A Vector Space Model (July 30, 2016) (unpublished manuscript), http://ssrn.com/abstract=2709238. This research is in use by an entity called the Patent Research Foundation to create "proximity reports"; the entity's website is at https://www.patrf.org, and Messrs. Younge and Kuhn are two of the Foundation's three listed directors. *The Patent Research Foundation*, PATRF, https://www.patrf.org/foundation (last visited Sept. 8, 2017).

248. Request for Comments on Enhancing Patent Quality, 80 Fed. Reg. 6475 (Feb. 5, 2015) ("Given that computerized searching algorithms and database technologies have advanced significantly in recent years, the USPTO is seeking input on new tools that might be useful to conduct a pre-examination search. For instance, the new tool might utilize a custom extraction routine that enables keyword, stemming, concept-semantic, and relational word searching capabilities. The USPTO's current pre-examination search tool PLUS does not possess these functionalities. Likewise, the new tool might employ more modern natural language search queries, which PLUS also cannot do.").

linguistics tools, provide the context within which to judge the language the patentee actually did choose.

How can one bring about this change of focus? The Patent Office, for its part, should make concerted efforts to develop corpus linguistics tools for use during examination. The goal should be to flag, in every application, all claim words and phrases that render the claim language atypical for that art, given the prior art U.S. patents and published applications with which one can compare it. The Patent Office tools should also be made readily available to potential applicants, who can use them to help avoid from the outset at least some share of indefinite claim language.

Private firms that have numerous, extensive interactions with the patent system, and who also have both expertise in computer science and substantial economic resources—Apple, Google, IBM, Microsoft, and others—should also help drive change. These firms, singly or as a group, should establish projects and prizes for computer scientists, linguists, or others who successfully create corpora and corpus linguistics tools that are tailored to enhance patent claim definiteness analysis.

Appendix

Relative frequencies of *fare* and *fantasy*, 1908-2008, *American English* corpus