

Trademark Fame and Corpus Linguistics

Jake Linford* & Kyra Nelson†‡

ABSTRACT

Trademark law recognizes that the same word can mean different things in different commercial contexts. Legal protection might extend to two or more owners who use the same symbol (like Delta) to indicate different sources of disparate goods or services, such as airlines and faucets. Generally, only those uses that threaten to confuse consumers—the use of similar symbols on identical or related goods—are subject to legal sanction.

But the law extends special protection to famous trademarks, not only against confusing use, but also against dilution: non-confusing use that blurs or tarnishes the distinctiveness of the famous mark. The result of protection against blurring is that the law treats the famous mark as if the sole proper use of the term in the commercial context is to designate goods and services from the famous mark's owner.

Protection against dilution extends only to famous marks, but courts and scholars apply differing standards for assessing fame. Nonetheless, the trend over time has been to treat fame as a threshold requiring both sufficient renown—the famous mark must be a household name—and relatively singular use.

This article argues that corpus linguistic analysis can provide evidence of whether a mark is sufficiently prominent and singular to qualify for anti-dilution protection. Corpus linguistics detects language patterns and meaning from analyzing actual language use. This article draws data primarily from two large, publicly accessible

* Loula Fuller & Dan Myers Professor, Florida State University.

† Brigham Young University Law School.

‡ The authors thank Suneal Bedi, Barton Beebe, Christopher Buccafusco, Sarah Burstein, Rebecca Schoff Curtin, Dave Fagundes, James Hicks, Sapna Kumar, Joe Miller, Janewa Osei-Tutu, Amelia Rinehart, Alexandra Jane Roberts, Guy Rub, Andres Sawicki, Zahr Said, Felix Wu, my colleagues at the Florida State University College of Law, and participants and sponsors of workshops and presentations at the University of Houston Law Center Institute for Intellectual Property & Information Law, the Trademark Scholars Roundtable, the 2020 BYU Corpus Linguistics Symposium, Cardozo Law's 2020 Intellectual Property and Information Law Colloquium, IPSC 2020, and the 17th Annual Works in Progress in Intellectual Property at Santa Clara University School of Law.

© 2022 Linford & Nelson. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction, provided the original author and source are credited.

databases (corpora) to investigate whether litigated trademarks are both prominent and unique. Courts and parties can consider frequency evidence to establish or refute prominence, and contextual evidence like concordance and collocation to establish relative singularity.

Corpus evidence has some advantages over standard methods of assessing fame. Corpus evidence is cheaper to generate than survey evidence but may be equally probative. Corpus analysis can help right-size dilution litigation: A litigant could estimate the prominence and singularity of an allegedly famous mark using corpus evidence prior to discovery and better predict whether the mark should qualify for anti-dilution protection. Judges should be able to rely on the results of corpus analysis with reasonable confidence. Additionally, corpus evidence can show use of a mark over time, providing courts with tools to assess when a mark first became famous, a question that a survey generated for litigation cannot readily answer.

INTRODUCTION

Trademark law recognizes and embraces an inherent homonymy in commercial communication.¹ Legal protection might extend to two or more owners who use the same symbol (like Delta) to indicate different sources of disparate goods or services, such as airlines and faucets. Only the use of a symbol that threatens to confuse consumers—the use of a similar symbol on identical or related goods or services—is subject to legal sanctions for trademark infringement. Thus, neither owner can use trademark law to police the other because source confusion is unlikely.

But there are exceptions to this homonymous structure of trademark law. The law extends special protection to famous trademarks. Famous marks receive protection not only against confusing use, but also against non-confusing use that tarnishes the famous mark or blurs its distinctiveness.² For example, Eastman Kodak, the well-known seller of cameras and film, could prevent the use of its famous Kodak mark to sell pianos if it could persuade a court that use on pianos would reduce the distinctiveness of the Kodak mark. It could similarly prevent the use of Kodak as a mark for a pornographic film studio if that use was deemed likely to harm the

1. Homonyms are words that look or sound the same but have different meanings from unrelated sources. For example, *bank* means both the land by a river and a financial institution, but those meanings developed independently. See Jake Linford, *The False Dichotomy Between Suggestive and Descriptive Trademarks*, 76 OHIO ST. L.J. 1367, 1397 (2015) [hereinafter Linford, *False Dichotomy*]. See also *infra* note 111 and accompanying text.

2. As briefly defined in the Lanham Act, the federal statute that governs trademarks and unfair competition, “‘dilution by tarnishment’ is association arising from the similarity between a mark or trade name and a famous mark that harms the reputation of the famous mark.” 15 U.S.C. § 1125(c)(2)(C). Blurring receives an equally brief definition, “association arising from the similarity between a mark or trade name and a famous mark that impairs the distinctiveness of the famous mark,” but with a helpful checklist of factors. 15 U.S.C. § 1125(c)(2)(B). See also *infra* Part I.A.

reputation of the Kodak mark.³ Non-confusing use may trigger legal sanctions if a court concludes the plaintiff's mark is famous. The end result of protection against blurring is that the famous mark is treated as if it were legally monosemous (i.e., as if the sole proper use of the term in the commercial context is to designate goods and services from the famous mark's owner).⁴ Extending this protection to the famous mark makes it a de jure monoseme, at least in the commercial context.

Trademark scholars are skeptical of extending protection against blurring, perhaps because the concept runs starkly counter to the homonymous structure of trademark law. In the absence of likely confusion due to wide disparity in the goods sold, consumers may not care whether they misapprehend the source of Kodak pianos. Competitors are nonetheless barred from using that term to sell goods the famous mark owner will never sell or license. Potential harm to consumers is unclear, but the harm to competitors—and perhaps to the competitive marketplace—is clearer. Criticism of anti-dilution protection abounds,⁵ but Congress seems unlikely to alter its scope.⁶

One way to diminish the potential cost of anti-dilution protection is to ensure that it applies only to a truly famous mark. Federal law currently defines a famous mark as “widely recognized by the general consuming public of the United States as a designation of source of the goods or services of the mark's owner.”⁷ This language lacks specificity, and courts have struggled to define fame. Blurring ostensibly threatens the ability of a mark to serve as a unique identifier, so perhaps for a mark to qualify for protection against blurring, it must already be relatively unique in the

3. Cf. *Eastman Kodak Co. v. Rakow*, 739 F. Supp. 116, 118 (W.D.N.Y. 1989) (granting injunction for violation of New York state dilution statute against comedian who used Kodak as his stage name and whose act included vulgar humor about bodily functions and sex).

4. A monoseme is a word with only one meaning. See Linford, *False Dichotomy*, *supra* note 1, at 1392–93. For more on how linguists define monosemy and how the concept might shed light on the singularity benchmark, see *infra* Part I.C.

5. See, e.g., Lisa P. Ramsey, *Free Speech Challenges to Trademark Law After Matal v. Tam*, 56 HOUS. L. REV. 401, 401 (2018) (“[T]rademark dilution law [works] an unconstitutional regulation of nonmisleading commercial expression.”); Sandra L. Rierson, *The Myth and Reality of Dilution*, 11 DUKE L. & TECH. REV. 212, 213 (2012) (“[T]he social and transaction costs imposed by dilution law still outweigh the harm that it is designed to avert.”); Rebecca Tushnet, *Gone in Sixty Milliseconds: Trademark Law and Cognitive Science*, 86 TEX. L. REV. 507, 516–17, 525–27, 537 (2008) (critiquing the empirical case for dilution by blurring); Daniel Klerman, *Trademark Dilution, Search Costs, and Naked Licensing*, 74 FORDHAM L. REV. 1759 (2006) (critiquing the search cost account for blurring); Christine Haight Farley, *Why We Are Confused About the Trademark Dilution Law*, 16 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 1175, 1185 (2006) (arguing that dilution is unreal and unprovable); Maureen Morrin & Jacob Jacoby, *Trademark Dilution: Empirical Measures for an Elusive Concept*, 19 J. PUB. POL. & MKTG. 265, 274 (2000) (reporting research indicating famous marks appear immune to blurring); Robert N. Klieger, *Trademark Dilution: The Whitling Away of the Rational Basis for Trademark Protection*, 58 U. PITT. L. REV. 789, 795 (1997) (“[D]ilution protection . . . pose[s] an anticompetitive threat to market efficiency and consumer welfare.”).

6. Cf. Mark A. Lemley, *Fame, Parody, and Policing in Trademark Law*, 2019 MICH. ST. L. REV. 1, 17 n.53 (2019); Mark P. McKenna, *Dilution and Free Speech in the U.S., Reprise*, draft at 1, 9, 17 (March 13, 2019), <https://perma.cc/JGF3-3A8H> (arguing that dilution may be unconstitutional following the Supreme Court's decision in *Matal v. Tam*, but demurring from predicting courts will so hold).

7. 15 U.S.C. § 1125(c)(2)(A).

commercial lexicon. To that end, Barton Beebe argued that protection against blurring should be extended only to a fanciful or coined mark (like Xerox or Kodak) that is also “truly renowned.”⁸ Setting a sufficiently high benchmark to qualify for anti-dilution protection limits the scope of and potential costs imposed by extending that protection.

This Article proposes a slightly less rigid framework than Beebe’s for determining which marks might reasonably merit protection from blurring. First, the mark must have achieved an extraordinary level of prominence or renown in the market. As courts and commentators often repeat, a famous mark must be a “household name.”⁹ Second, the mark must be relatively singular as a source signifier across the spectrum of consumer goods or services. A term used by a variety of sellers for different products would fail to qualify.¹⁰ Essentially, use of the famous mark should signify a single source, irrespective of the differences in goods or services offered.

Prominence and singularity benchmarks are critical components of a properly functioning anti-dilution system. For example, the presumption undergirding fame is that the Kodak mark has acquired significant prominence and familiarity to the vast majority of consumers, even those who do not regularly shop for cameras or film. Should a seller of pianos use the Kodak mark to identify its goods, consumers might conclude that the seller of Kodak film also offers these other Kodak goods, even though these new offerings are dissimilar from those offered by Kodak. If the mark is not prominent, or its use is not singular, courts risk overprotecting the mark in a manner that hampers competition without commensurate benefit.

Another way to reduce the potential harm of anti-dilution protection is to ensure courts and litigants have accurate evidence of fame to assess whether a mark hits the proper benchmarks. Courts will accept survey evidence of the fame or commercial strength of the mark,¹¹ but circumstantial evidence in the form of sales and advertising statistics are also consistently accepted.¹²

8. Barton Beebe, *The Semiotic Analysis of Trademark Law*, 51 UCLA L. REV. 621, 694–95 (2004). *See also id.* at 698 (conceding that “[i]t is at least arguable that only truly renowned marks require, in order to protect their renown, the wide grant of property rights called for by antiblurring protection” but arguing anti-tarnishment protection should not be so limited). The inherent distinctiveness or protectability of a trademark is typically described as falling along a spectrum based on the relationship between the mark and the product sold. Generic terms like Computer for computers are unprotectable. Descriptive terms like Tasty for salad dressing can acquire distinctiveness to qualify for protection. Inherently distinctive marks are protectable from first use and increase in strength along the spectrum: suggestive terms like Penguin for air conditioners; arbitrary terms like Apple for computers; and fanciful terms like Kodak for cameras. *See, e.g.*, Jake Linford, *Are Trademarks Ever Fanciful?*, 105 GEO. L.J. 731, 738 (2017) (describing and analyzing the spectrum of conceptual trademark strength most famously articulated in *Abercrombie & Fitch Co. v. Hunting World, Inc.*, 537 F.2d 4, 9–11 (2d Cir. 1976)).

9. *See, e.g.*, *Coach Servs., Inc. v. Triumph Learning LLC*, 668 F.3d 1356, 1373 (Fed. Cir. 2012) (internal citations omitted). *See also infra* Part II.B.

10. Throughout the text, when we say product, we mean good(s) or service(s) but will use product (singular) where helpful to simplify.

11. *See, e.g.*, *Visa Intern. Service Ass’n v. JSL Corp.*, 590 F. Supp. 2d 1306, 1315 (D. Nev. 2008), *aff’d on other grounds*, 610 F.3d 1088 (9th Cir. 2010).

12. *See, e.g.*, *Bose Corp. v. QSC Audio Prods., Inc.*, 293 F.3d 1367 (Fed. Cir. 2002) (“[W]e have consistently accepted statistics of sales and advertising as indicia of fame . . .”).

But it may be possible to provide an earlier, accurate estimate of likely fame. The science of linguistics, broadly defined, might help bring trademark law into better agreement with consumer perception of and interactions with ostensibly famous marks.¹³ Corpus analysis can help uncover evidence of the prominence and singularity of purported famous marks. This Article proposes a methodology for using various corpus linguistic tools to distinguish famous from non-famous marks prior to discovery or the gathering of survey evidence.

A corpus is “a body or database of naturally occurring language.”¹⁴ That text can be gathered from formal and informal written and spoken sources, including magazines, newspapers, television programming, academic articles, e-mail messages, or transcripts of spoken language. Modern data storage and computing power allows analysis of massive corpora to uncover evidence of language use that was imperceptible to previous generations of scholars. A corpus can be analyzed for the statistical frequency with which a word appears. If the corpus has wide-ranging source material, analysis can show changes in the frequency or meaning of words over time. Comparing the statistical frequency of a mark to other litigated marks and to standard measures of word frequency can help courts and litigants quickly determine whether a mark is truly prominent.¹⁵

Corpus linguistic analysis often employs collocation and concordance to reveal the context of word usage. Collocation “is the tendency of words to be biased in the way they co-occur.”¹⁶ As used in corpus analysis, “collocates” are the words that appear near the target word. For example, top collocates of *bread* include loaf, slice, crumb, butter, white, fresh, eat, bake, and break. Collocational analysis rests on the

13. For an introduction to other scholarship applying linguistic analysis to trademark controversies, see Jake Linford, *A Linguistic Justification for Protecting “Generic” Trademarks*, 17 YALE J. L. & TECH. 110, 112 n.1 (2015) [hereinafter Linford, *Linguistic*].

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

15. Other scholars propose empirical measures of trademark fame. See, e.g., Suneal Bedi & Mike Schuster, *Towards an Objective Measure of Trademark Fame*, 54 U.C. DAVIS L. REV. 431 (2020) (proposing a measure rooted in consumer recognition speed assessed through a product recall method); Adam Omar Shanti, Comment, *Measuring Fame: The Use of Empirical Evidence in Dilution Actions*, 5 MARQ. INTELL. PROP. L. REV. 177 (2001). Other articles engage in efforts to empirically test blurring or tarnishment. See, e.g., Suneal Bedi & David Reibstein, *Measuring Trademark Dilution by Tarnishment*, 95 IND. L.J. 683 (2020) (advancing an “experimental survey methodology to empirically show that tarnishment can exist under certain conditions [where] the key is increasing the number of exposures to the harmful mark”); Barton Beebe, Roy Germano, Christopher Jon Sprigman & Joel H. Steckel, *Testing for Trademark Dilution in Court and the Lab*, 86 U. CHI. L. REV. 611 (2019) (criticizing the response time surveys often used to show dilution and proposing an association strength test); Hannelie Kruger & Christo Boshoff, *The Influence of Trademark Dilution on Brand Attitude: An Empirical Investigation*, 24 MGMT. DYNAMICS 50 (2015); Maureen Morrin, Jonathan Lee & Greg M. Allenby, *Determinants of Trademark Dilution*, 33 J. CONSUMER RSCH. 248 (2006); Julie Manning Magid, Anthony D. Cox & Dena S. Cox, *Quantifying Brand Image: Empirical Evidence of Trademark Dilution*, 43 AM. BUS. L.J. 1, 4 (2006) (outlining a randomized experimental protocol for proving actual dilution); David J. Franklyn, *Debunking Dilution Doctrine: Toward a Coherent Theory of the Anti-Free-Rider Principle in American Trademark Law*, 56 HASTINGS L.J. 117, 131 (2004) (questioning the empirical evidence of dilution); Morrin & Jacoby, *supra* note 5, at 274.

16. SUSAN HUNSTON, CORPORA IN APPLIED LINGUISTICS 68 (2002).

idea that two words frequently occurring together can reveal something meaningful about those terms. Collocation allows interpretation of words in the semantic context where they are found, “providing useful information about the range of possible meanings.”¹⁷

Concordance results instead present the trademark in its surrounding context. Concordance results provide a fragment of the text that appears immediately before and after the target word. For example, a concordance analysis of *bread* could return a string like the following:

. . . of the Soviet collapse. I remember reading once that the prices of grain and **bread** were so out of whack in communist Poland that farmers were not buying pig feed . . .¹⁸

Examining collocates and concordances of purportedly famous marks can quickly illustrate if a mark is used in sufficiently singular fashion in the marketplace.

Consider Microsoft, seen by many as the prototypical famous mark—a term with a meaning that has not slipped beyond its corporate bounds.¹⁹ Corpus analysis confirms the intuition that Microsoft is prominent and singular. We analyzed the use of Microsoft in COCA, the Corpus of Contemporary American English.²⁰ Microsoft appears approximately twenty-four times per million words, a frequency that places it among the top 3,200 most frequently used words in the corpus.²¹ The top twenty collocates of Microsoft—the words that most frequently appear within four words of Microsoft in the corpus—relate to Microsoft Corporation’s goods and services, its corporate leadership, its competitors, or its litigating posture. Concordance results also indicate the use of Microsoft is similarly singular. Nearly every concordance result we reviewed shows brand use that points back to the Microsoft Corporation. Microsoft’s fame is reflected both in its singularity and its prominence. A mark that boasts a similar level of contextual singularity and high frequency use as Microsoft should readily qualify as a famous mark.

Part I of this Article explains the standard tests for famous marks and explains why those tests suggest a famous mark must be both singular and prominent. Part II explains how corpus evidence might help courts and litigants assess singularity and prominence. Singularity can be investigated via concordance and collocation. Prominence can be measured by frequency of the mark’s occurrence in a corpus. We demonstrate this process by assessing trademarks previously litigated in dilution cases using collocation, concordance, frequency, and diachronic measures. Our

17. Lee & Mouritsen, *supra* note 14, at 832.

18. Mark Sumner, *Owners of the World, Unite*, DAILY KOS (May 22, 2011), <https://perma.cc/Y5RF-U9VR>.

19. DAVID CRYSTAL, HOW LANGUAGE WORKS 191 (2005).

20. CORPUS OF CONTEMPORARY AMERICAN ENGLISH (COCA), <https://perma.cc/N429-DT34> (last visited Nov. 16, 2021). COCA is a corpus of 600 million words from American sources, equally divided among spoken, fiction, popular magazines, newspapers, and academic texts, with 20 million words from each year from 1990 to 2019. Our study also considers evidence from GloWbE, the Corpus of Global Web-Based English, and COHA, the Corpus of Historical American English (COHA). *See infra* Part II.

21. *See infra* notes 189–190 and accompanying text.

assessment highlights marks that may have been incorrectly deemed famous and others that were incorrectly denied protection. In addition, the evidence also indicates that just as marks may ascend to fame over time, some traditionally famous marks may descend to mere distinctiveness. Part III highlights how a court might use corpus evidence to dispose of an anti-dilution claim due to lack of fame at an appropriately early stage of litigation. We then consider some implications of and limitations on the use of corpus linguistics to improve fame determinations.

I. FAME: PROMINENCE AND SINGULARITY

The common law recognized the right of trademark owners to prevent confusing or fraudulent uses of a mark.²² But the owner of a famous mark can also acquire relief against dilution even when consumers are unlikely to be confused by the use.²³ This Part describes the development of anti-dilution protection in the United States and its connection to the concept of fame.

The Lanham Act, the law governing federal trademark protection, extends federal anti-dilution protection only to famous marks, but what constitutes fame has shifted over time. Courts and scholars are approaching consensus that anti-dilution protection is properly extended only to those marks that have reached both a significant level of singularity and prominence in the market. This history suggests that courts should require a plaintiff to establish both significantly singular use of the mark and a high level of prominence or renown to qualify for anti-dilution protection.

A. A BRIEF HISTORY OF ANTI-DILUTION LAW IN THE UNITED STATES

Frank Schechter was the first scholar to propose that trademark law might appropriately discourage non-confusing uses of certain prominent marks.²⁴ Massachusetts was the first state to provide anti-dilution protection, but that protection was not limited to famous marks.²⁵ A majority of states subsequently enacted statutory anti-dilution protection.²⁶

22. Keith Aoki, *Authors, Inventors and Trademark Owners: Private Intellectual Property and the Public Domain Part II*, 18 COLUM.-VLA J.L. & ARTS 191, 236 (1994).

23. 15 U.S.C. § 1125(c)(1). *See also* Paul J. Heald, *Filling Two Gaps in the Restatement (Third) of Unfair Competition: Mixed-Use Trademarks and the Problem with Vanna*, 47 S.C. L. REV. 783, 801 (1996) (critiquing anti-dilution statutes because they create liability in the absence of likely confusion).

24. *See infra* Part I.B.

25. Act of May 2, 1947, ch. 307 § 7a, 1947 Mass. Acts 300 (current version at MASS. GEN. LAWS ANN. CH. 110B, § 12).

26. ALA. CODE § 8-12-17; ALASKA STAT. § 45.50.180; ARIZ. REV. STAT. § 44-1448.01; ARK. CODE ANN. § 4-71-213; CAL. BUS. & PROF. CODE § 14247; CONN. GEN. STAT. § 35-11i(c); DEL. CODE ANN. tit. 6 § 3313; FLA. STAT. ANN. § 495.151; GA. CODE ANN. § 10-1-451; HAW. REV. STAT. § 482-32; IDAHO CODE § 48-513; ILL. COMP. STAT. ANN. § 1036/65; IOWA CODE ANN. § 548.113; KAN. STAT. ANN. § 81-214; LA. REV. STAT. ANN. § 51:223.1; ME. REV. STAT. ANN. tit. 10, § 1530; MASS. GEN. LAWS ANN. CH. 110B, § 12; MINN. STAT. § 333.285; MISS. CODE ANN. § 75-25-25; MO. REV. STAT. § 417.061; MONT. CODE ANN. § 30-13-334; NEB. REV. STAT. § 87-140; NEV. REV. STAT. ANN. § 600.435; N.H. REV. STAT.

State anti-dilution laws generally target two types of harmful behavior: blurring and tarnishment.²⁷ As described in one state law case, *Deere & Co. v. MTD Prods., Inc.*, blurring occurs when “the defendant uses or modifies the plaintiff’s trademark to identify the defendant’s goods and services, raising the possibility that the mark will lose its ability to serve as a unique identifier of the plaintiff’s product.”²⁸ In the Kodak example from above, blurring could occur if a seller began offering pianos under the Kodak mark, even though consumers are unlikely to be confused by that sale.²⁹

In *Deere*, the court also stated that tarnishment arises “when the plaintiff’s trademark is linked to products of shoddy quality, or is portrayed in an unwholesome or unsavory context likely to evoke unflattering thoughts about the owner’s product[s].”³⁰ Thus, for example, the Kodak mark might be tarnished if an adult performer used that mark as part of their *nom de scène*.³¹

The Lanham Act was amended in 1995 to extend federal protection against dilution to famous marks.³² That amendment, the Federal Trademark Dilution Act (FTDA), defines dilution as the lessening of the capacity of a famous mark to identify and distinguish goods or services, irrespective of the presence or absence of competition between sellers, likely confusion, mistake, or deception.³³ The FTDA did not include language creating a cause of action for tarnishment.³⁴ Nonetheless, the legislative history indicates Congress intended to provide a mechanism to discourage tarnishing uses.³⁵ Courts read a tarnishment cause of action into the statute,³⁶ building on state law doctrines.³⁷

ANN. § 350-A:12; N.J. STAT. ANN. § 56:3-13.20; N.M. STAT. § 57-3B-13; N.Y. GEN. BUS. LAW. § 360-1; OR. REV. STAT. § 647.107; 54 PA. CONSOL. STAT. ANN. § 1124; R.I. GEN. LAWS § 6-2-12; S.C. CODE ANN. § 39-15-1165; TENN. CODE ANN. § 47-25-513; TEX. BUS. & COM. CODE ANN. § 16.29 16.103; WASH. REV. CODE ANN. § 19.77.160; W. VA. CODE § 47-2-13; WYO. STAT. ANN. § 40-1-115; *see also* Sarah L. Burstein, *Dilution by Tarnishment: The New Cause of Action*, 98 TRADEMARK REP. 1189, 1192–96 (2008) (identifying two broad categories of state dilution statutes).

27. Clarisa Long, *Dilution*, 106 COLUM. L. REV. 1029, 1036 (2006).

28. *Deere & Co. v. MTD Prods., Inc.*, 41 F.3d 39, 43 (2d Cir. 1994) (applying New York dilution law).

29. *See supra* Introduction.

30. *Deere*, 41 F.3d at 43; *see also* Gerard N. Magliocca, *One and Inseparable: Dilution and Infringement in Trademark Law*, 85 MINN. L. REV. 949, 962 (2001).

31. *See, e.g.*, *Eastman Kodak Co. v. Rakow*, 739 F. Supp. 116, 118 (W.D.N.Y. 1989).

32. Federal Trademark Dilution Act of 1995, PL 104–98, January 16, 1996, 109 Stat 985.

33. *Id.*

34. Beebe, *supra* note 8, at 698.

35. H.R. REP. NO. 104–374, at 2 (1995), reprinted in 1995 U.S.C.C.A.N. 1029, 1029 (“The purpose of H.R. 1295 [the FTDA bill] is to protect famous trademarks from subsequent uses that blur the distinctiveness of the mark or tarnish or disparage it . . .”).

36. Beebe, *supra* note 8, at 697 (criticizing the phenomenon).

37. *Jews for Jesus v. Brodsky*, 993 F. Supp. 282, 306 (D.N.J. 1998), *aff’d*, 159 F.3d 1351 (3d Cir. 1998) (footnotes omitted). Plaintiffs also brought early FTDA cases to seek relief against cybersquatting prior to the passage of the Anticybersquatting Consumer Protection Act (ACPA), 15 U.S.C. § 1125(d). Xuan-Thao N. Nguyen, *A Circus Among the Circuits: Would the Truly Famous and Diluted Performer*

Congress amended the Lanham Act's anti-dilution provisions in 2006 when it passed the Trademark Dilution Revision Act (TDRA). Amendments included revising the definition of fame and specifying that niche fame no longer qualified a mark for anti-dilution protection;³⁸ resetting the burden of proof for famous mark owners;³⁹ and clearly articulating specific causes of action for blurring and tarnishment.⁴⁰ The TDRA still governs federal anti-dilution law.

B. PROMINENCE AND SINGULARITY UNDER THE LAW

Federal anti-dilution protection extends only to famous marks. Scholars question whether courts applying those laws sufficiently restrict that protection to only the truly famous, and how to define fame. In this Section, we explain how the TDRA's language and construction builds on the understanding developed at common law and pursuant to state and federal legislation. State anti-dilution laws and courts have inconsistently required evidence of both prominence and singularity for a mark to qualify as famous. After the passage of the FTDA, courts made uneven strides toward requiring both prominence and singularity. While sometimes erring in providing dilution protection to marks of questionable fame, courts applying the TDRA increasingly set both prominence and singularity benchmarks for marks to qualify as famous.

1. *The Rational Basis for Preventing Dilution*

In 1927, Frank Schechter first proposed that some marks should be protected against what we now call dilution.⁴¹ Schechter's proposal built on a trend in then-

Please Stand Up: The Federal Trademark Dilution Act and Its Challenges, 1 CHI.-KENT J. INTELL. PROP. 158, 170–71 (1999) (reporting early cases in which courts extended protection to marks that were arguably not famous in order to curb the use of deceptive domain names).

38. See, e.g., *Bentley v. NBC Universal, LLC*, CV 16-03693, 2016 WL 10570587 (C.D. Cal. Sept. 28, 2016) (granting motion to dismiss dilution claims; allegation that plaintiff's marks "are recognized by people interested in football, sports performance, and fitness training" held insufficient to support element of fame); see also 6 CALLMANN ON UNFAIR COMPETITION, TRADEMARKS & MONOPOLIES § 22:22 (4th ed., 2020 update). But see Alexandra J. Roberts, *New-School Trademark Dilution: Famous Among the Juvenile Consuming Public*, 100 TRADEMARK REP. 1021 (2010) (stating that courts have accepted fame among juveniles as sufficient but arguing against the practice because it encourages harmful persuasive advertising aimed at children); Katherine D. Jochim, *A Place for Famous Market Niche Trade and Service Marks Within Federal Trademark Dilution Law*, 47 SAN DIEGO L. REV. 545 (2010) (arguing that federal dilution law should protect coined marks with niche fame); Natalya Y. Belonozhko, *Famous Trademarks in Fashion*, 51 WILLAMETTE L. REV. 365 (2015) (arguing in favor of dilution protection for marks with niche fame to protect small businesses).

39. Jeremy N. Sheff, *The (Boundedly) Rational Basis of Trademark Liability*, 15 TEX. INTELL. PROP. L.J. 331, 343 (2007) (citing H.R. REP. NO. 109-23, at 5 (2005), as reprinted in 2006 U.S.C.C.A.N. 1091) ("The TDRA was consciously drafted as a rejection" of the Supreme Court's insistence in *Moseley v. V Secret Catalogue, Inc.*, 537 U.S. 418 (2003), that the plaintiff provide "proof of actual (as opposed to likely) dilution.").

40. See *supra* note 2 and accompanying text.

41. Frank I. Schechter, *The Rational Basis of Trademark Protection*, 40 HARV. L. REV. 813, 825–26 (1927).

recent cases to extend trademark protection against non-competing uses of a senior user's mark.⁴² In Schechter's view, allowing non-confusing uses of these marks would lead to a "gradual whittling away or dispersion" of the senior mark's distinctiveness in the marketplace.⁴³ Schechter posited that while the use of a well-known mark like Kodak on noncompeting products like bathtubs or cakes would be unlikely to confuse consumers, allowing such use would result in famous marks like Kodak becoming "commonplace words of the language, despite the originality and ingenuity in their contrivance."⁴⁴ Schechter didn't coin the term "dilution," nor did he speak of fame, but instead argued in favor of "preservation of the uniqueness of a trademark."⁴⁵

Schechter's argument for protecting the uniqueness of certain marks from whittling away appears to stem in large part from a German case before a regional trial court concerning the use of Odol as a mark both for plaintiff's mouthwash and defendant's steel products.⁴⁶ The court in *Odol* concluded that the mark Odol for mouthwash was more than merely a trademark. As translated, the court noted that "'Odol' developed into a catchword that has come to distinguish the complainant's goods and has acquired an advertising strength that goes beyond the otherwise typical function of trademarks, such that when anyone reads or hears the word 'Odol,' they think of the complainant's mouthwash."⁴⁷ The court determined on the basis of that heightened distinctiveness "the complainant has the greatest interest in ensuring that

42. *Id.*

43. *Id.*

44. *Id.*; see also Barton Beebe, *Intellectual Property Law and the Sumptuary Code*, 123 HARV. L. REV. 810, 846 (2010). Courts often invoke the unauthorized use of the famous Kodak mark for pianos as a primary example of diluting use. A recent search identified eighty-five dilution cases in which courts invoked Kodak pianos. See, e.g., *Moseley v. V Secret Catalogue, Inc.*, 537 U.S. 418, 431 (2003). Scholars similarly included the example in 232 articles. See, e.g., Deven R. Desai & Sandra L. Rierson, *Confronting the Genericism Conundrum*, 28 CARDOZO L. REV. 1789, 1843 (2007).

45. Schechter, *supra* note 41, at 831.

46. Barton Beebe, *The German Misappropriation Origins of Trademark Antidilution Doctrine: A Translation of the 1924 Odol Opinion of the Elberfeld Landgericht*, in *TRANSITION AND COHERENCE IN INTELLECTUAL PROPERTY LAW* 460, 464 (Niklas Bruun, Graeme B. Dinwoodie, Marianne Levin & Ansgar Ohly eds., 2021); see Schechter, *supra* note 41, at 832–33; Barton Beebe, *The Suppressed Misappropriation Origins of Trademark Antidilution Law: The Landgericht Elberfeld's Odol Opinion and Frank Schechter's "The Rational Basis of Trademark Protection,"* in *INTELLECTUAL PROPERTY AT THE EDGE: THE CONTESTED CONTOURS OF IP 59* (Rochelle Cooper Dreyfuss & Jane C. Ginsburg eds., 2014). Beebe characterizes the *Odol* case as sounding in misappropriation rather than trademark infringement. Beebe, *supra*, at 75. *But see* Sheff, *supra* note 39, at 333–34 (arguing that trademark infringement and dilution are best conceived as "related and overlapping categories" of "commercial behavior that manipulates the cognitive biases of consumers, and as such threatens to render their heuristic judgments persistently inaccurate").

47. See *Odol*, *supra* note 46, ¶ 4. The original German reads as follows:

„Odol“ ein Schlagwort geworden ist, das für die Waren der Klägerin kennzeichnend geworden ist und eine über die sonstige Bedeutung von Warenzeichen hinausgehende Werbekraft erlangt hat, so daß jedermann, wenn er das Wort Odol liest oder hört, an das klägerische Mundwasser denkt.

Our thanks to Barton Beebe for sharing this translation.

its mark not become diluted; it would lose in advertising strength and become devalued, if it is used for other, wholly different goods.”⁴⁸

Both Schechter and the *Odol* court viewed the marks that qualified for protection against dilution as noteworthy because of their prominence and uniqueness. Indeed, Schechter was so bold as to state that “the only rational basis” for trademark protection was to protect the “uniqueness and singularity” of the mark.⁴⁹ Schechter’s proposal was limited to marks that were also what we would today call fanciful or arbitrary,⁵⁰ “added to . . . the human vocabulary by their owners” and associated “from the very beginning . . . in the public mind with a particular product, not with a variety of products.”⁵¹ It took twenty years for Schechter’s proposal to gain legislative traction, but the majority of states now provide anti-dilution protection,⁵² although not every statute limits protection to marks that are famous.⁵³

2. Fame Under the TDRA

In 2006, Congress enacted the appropriately named Trademark Dilution Revision Act, which revised the fame standard.⁵⁴ The TDRA defines a famous mark as one that is “widely recognized by the general consuming public of the United States as a designation of source of the goods or services of the mark’s owner.”⁵⁵ The legislative history of the TDRA identified Congress’s purpose as protecting “only the very famous trademark” from blurring or tarnishment.⁵⁶ That history also articulates an intent to “tighten[] the definition of what is necessary to be considered a famous mark” compared to the ostensibly looser FTDA standard.⁵⁷

48. *See id.* ¶ 5:

Die Klägerin hat infolgedessen das größte Interesse daran, daß ihr Zeichen nicht verwässert wird; es würde an Werbekraft einbüßen, wenn jedermann es zur Bezeichnung seiner Waren verwenden würde. Das Warenzeichen wird für die Waren der Klägerin entwertet, wenn es für andere gänzlich verschiedene Waren verwendet wird.

49. Schechter, *supra* note 41, at 831.

50. An arbitrary mark is a symbol with preexisting meaning unconnected to the good or service offered in association with it. A fanciful mark is an invented word, “coined specifically for use as a trademark.” Linford, *False Dichotomy*, *supra* note 1, at 1405–06; *see also* Abercrombie & Fitch Co. v. Hunting World, Inc., 537 F.2d 4, 9–11 (2d Cir. 1976).

51. Schechter, *supra* note 41, at 829; *see also* J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION, § 24:87 (5th ed. 2019) (summarizing Schechter).

52. *See supra* note 26.

53. *See, e.g.*, MASS. GEN. LAWS ANN. Ch. 110H, § 3 (requires only that a mark is “registered under this chapter, or a mark valid at common law, or a trade name valid at common law”).

54. Passage of the TDRA was driven in large part by a response to the Supreme Court’s holding in *Moseley v. V Secret Catalogue, Inc.*, 537 U.S. 418, 431 (2003). *See supra* note 39 and accompanying text.

55. 15 U.S.C. § 1125(c)(2)(A). *But see* Graeme B. Dinwoodie & Mark D. Janis, *Dilution’s (Still) Uncertain Future*, 105 MICH. L. REV. FIRST IMPRESSIONS 98, 100 (2006) (positing that the TDRA fame factors might “trigger judicial resistance to blurring causes of action”).

56. H.R. REP. NO. 109-23, at 25 (2005) : (statement of Rep. Howard L. Berman) (“Dilution should once again be used sparingly as an ‘extraordinary’ remedy, one that requires a significant showing of fame.”).

57. *Id.*

The TDRA offers four factors for courts to consider in assessing “whether a mark possesses the requisite degree of recognition”:

- i. The duration, extent, and geographic reach of advertising and publicity of the mark, whether advertised or publicized by the owner or third parties.
- ii. The amount, volume, and geographic extent of sales of goods or services offered under the mark.
- iii. The extent of actual recognition of the mark.
- iv. Whether the mark was registered⁵⁸

To establish the first two factors, courts take evidence from mark owners regarding the duration of use of the purportedly famous mark, advertising expenditures, amount, volume, and extent of sales.⁵⁹ Parties often present survey evidence to establish the third factor, the extent of actual recognition among consumers.⁶⁰ Parties also offer evidence of unsolicited media references to the mark.⁶¹ Unfortunately, gathering this evidence can be prohibitively expensive for some litigants.⁶²

Courts have held that a famous mark must point more or less exclusively to the goods of the mark owner. Protection against blurring is extended to the famous mark on the presumption that it can “uniquely identify a single source” irrespective of the goods or services offered.⁶³ Plaintiffs will often boast exclusive use when attempting to establish fame.⁶⁴ But courts applying the TDRA have not required absolutely

58. 15 U.S.C. § 1125(c)(2)(A). The FTDA did not define fame, but it did articulate eight factors that courts could weigh in assessing fame. *See* 15 U.S.C. § 1125(c)(1)(A)–(H) (Supp. V 1999) (repealed).

59. *See, e.g.,* Clearly Food & Beverage Co. v. Top Shelf Beverages, Inc., 102 F. Supp. 3d 1154, 1176–77 (W.D. Wash. 2015) (concluding plaintiff failed to establish its Clearly Canadian mark had acquired sufficient fame to raise a question of fact regarding its dilution claim); Burberry Ltd. v. Euro Moda, Inc., No. 08 Civ. 5781, 2009 WL 1675080, at *12–13 (S.D.N.Y. June 10, 2009) (analyzing evidence that Burberry’s marks are famous under both the TDRA and FTDA standards).

60. *See, e.g.,* Maker’s Mark Distillery, Inc. v. Diageo N. Am., Inc., 703 F. Supp. 2d 671, 698 (W.D. Ky. 2010), *aff’d*, 679 F.3d 410 (6th Cir. 2012) (“[T]he lack of survey evidence leaves the Court with little to rely on as to the third factor—the extent of actual recognition of the mark.”). *But see* adidas-Am., Inc. v. Payless Shoesource, Inc., 546 F. Supp. 2d 1029, 1063 n.12 (D. Or. 2008) (“Given the extensive evidence adidas submitted as to each of the statutory ‘fame’ factors, its failure to conduct a fame survey is not dispositive.”).

61. *Paramount Farms Int’l LLC v. Keenan Farms Inc.*, No. 2:12-CV-01463-SVW-E, 2012 WL 5974169, at *14 (C.D. Cal. Nov. 28, 2012) (noting 300,000 “likes” on brand’s Facebook page “lends credence to [] evidence that the trade dress has become famous”). *But see* Moore v. Weinstein Co., LLC, No. 3:09-CV-00166, 2012 WL 1884758, at *44 (M.D. Tenn. May 23, 2012), *aff’d*, 545 F. App’x 405 (6th Cir. 2013) (holding that in absence of evidence under the other factors, unsolicited third party mentions were insufficient to establish the fame of the mark).

62. *See* Jake Linford, *Democratizing Access to Survey Evidence of Distinctiveness*, 225, 229–32, in *TRADEMARK LAW AND THEORY: REFORM OF TRADEMARK LAW* (Graeme B. Dinwoodie & Mark D. Janis, eds., 2021) [hereinafter Linford, *Democratizing Access*].

63. *Louis Vuitton Malletier S.A. v. Haute Diggity Dog, LLC*, 507 F.3d 252, 265 (4th Cir. 2007).

64. *Id.* at 265–66 (4th Cir. 2007) (discussing “the substantial goodwill VUITTON has established in its famous marks through more than a century of *exclusive use*”).

exclusive use from a mark owner.⁶⁵ As one district court concluded, “[t]he law does not require that use of the famous mark be absolutely exclusive, but merely ‘substantially exclusive.’”⁶⁶

Courts applying the TDRA have also recognized that establishing prominence and renown takes time and requires more than just secondary meaning or even substantially exclusive use. A famous mark is “part of the collective national consciousness.”⁶⁷ To that end, the TDRA ostensibly “restrict[s] dilution cases of action to those few truly famous marks.”⁶⁸ Indeed, a famous mark must be a “household name.”⁶⁹ As the Federal Circuit observed, “[i]t is well-established that dilution fame is difficult to prove.”⁷⁰ Thus, Thomas McCarthy concluded, only “truly eminent and widely recognized marks” should qualify as famous.⁷¹ As Barton Beebe summarizes it, the “TDRA is simply not intended to protect trademarks whose fame is at all in doubt.”⁷²

This is something of a change compared to the FTDA. Courts were under the impression that the FTDA’s fame threshold was “rigorous.”⁷³ But scholars criticized early fame determinations under the FTDA, noting that some courts appeared lax in applying the fame factors.⁷⁴ Other courts were critiqued for making fame determinations “on an intuitive basis.”⁷⁵ For example, a court held famous the

65. The same was true under the FTDA. *See, e.g.,* *Tiffany & Co. v. Classic Motor Carriages Inc.*, 10 U.S.P.Q. 2d 1835, 1989 WL 281893, at *3 (T.T.A.B. Mar. 22, 1989) (holding Tiffany was famous under the FTDA notwithstanding the existence of some limited third party uses); *see also* Shanti, *supra* note 15, at 188 (citing RICHARD L. KIRKPATRICK, *LIKELIHOOD OF CONFUSION IN TRADEMARK LAW* 3–7 (1999)).

66. *Nike, Inc. v. Nikepal Intern., Inc.*, No. 2:05-cv-1468-GEB-JFM, 2007 WL 2782030, at *7 (E.D. Cal. Sept. 18, 2007); *cf. L.D. Kichler Co. v. Davoil, Inc.*, 192 F.3d 1349, 1352 (Fed. Cir. 1999) (holding that under the FTDA in the trademark context, “substantially exclusive” use does not mean totally exclusive use).

67. *Apple, Inc. v. Samsung Elecs. Co., Ltd.*, No. 11–cv–01846–LHK, 2012 WL 2571719, at *7 (N.D. Cal. June 30, 2012).

68. *Bd. of Regents, Univ. of Texas Sys. ex rel. Univ. of Texas at Austin v. KST Elec., Ltd.*, 550 F. Supp. 2d 657, 679 (W.D. Tex. 2008).

69. *See, e.g.,* *Schutte Bagclosures Inc. v. Kwik Lok Corporation*, 193 F. Supp. 3d 245, 283 (S.D.N.Y. 2016), *aff’d*, 699 Fed. Appx. 93 (2d Cir. 2017); *Rosetta Stone Ltd. v. Google, Inc.*, 676 F.3d 144, 171 (4th Cir. 2012); *Kibler v. Hall*, 843 F.3d 1068, 1083 (6th Cir. 2016), *cert. denied*, 138 S. Ct. 91 (2017); *Blumenthal Distributing, Inc. v. Herman Miller, Inc.*, 963 F.3d 859, 870 (9th Cir. 2020); *Coach Servs., Inc. v. Triumph Learning LLC*, 668 F.3d 1356, 1373 (Fed. Cir. 2012). FTDA cases also used the “household name” formulation. *See, e.g.,* *Thane Int’l, Inc. v. Trek Bicycle Corp.*, 305 F.3d 894, 911 (9th Cir. 2002); *TCPIP Holding Co. v. Haar Commc’ns, Inc.*, 244 F.3d 88, 99 (2d Cir. 2001); *McNeil Consumer Brands, Inc. v. U.S. Dentek Corp.*, 116 F. Supp. 2d 604, 608 (E.D. Pa. 2000).

70. *Coach*, 668 F.3d at 1373 (holding evidence did not prove that COACH for high-end handbags and leather goods was a “famous” mark for purposes of the TDRA).

71. MCCARTHY, *supra* note 51, § 24:104.

72. Barton Beebe, *A Defense of the New Federal Trademark Antidilution Law*, 16 *FORDHAM INTELL. PROP. MEDIA & ENT. L.J.* 1143, 1158 (2006).

73. *Everest Cap. Ltd. v. Everest Funds Mgmt., L.L.C.*, 393 F.3d 755, 763 (8th Cir. 2005); *see* Long, *supra* note 27, at 1029 (noting that initial enthusiasm for enforcement under the FTDA dwindled over time).

74. Nguyen, *supra* note 37, at 187.

75. Shanti, *supra* note 15, at 182.

Panavision mark for theatrical motion picture, television camera, and photographic equipment.⁷⁶ *Panavision* was decided in the early days of the FTDA, when courts used the statute to police cybersquatting.⁷⁷ As Xuan-Thao Nguyen argues, had the court accurately applied the FTDA's test for fame, it "could not have concluded that 'Panavision' is a famous mark within the meaning of the Dilution Act."⁷⁸

In many pre-FTDA cases applying state anti-dilution statutes, courts consider both whether the allegedly diluted mark is generally known (i.e., whether its use was sufficiently prominent or renowned or it was a household name), and whether the claimant's use of the mark is relatively singular or exclusive.⁷⁹ Beverly Pattishall summarizes the late-twentieth-century view on what types of marks qualified for anti-dilution protection, identifying them as marks that had acquired "uniqueness, arbitrariness, fame and celebrity."⁸⁰ Pattishall articulates the connection between dilution and fame, as dilution reduces the ability of the mark to signify "anything unique, singular or particular," or to stand as "a single thing coming from a single source."⁸¹

Under the TDRA, uniqueness and prominence feed into one another. The TDRA's test for blurring connects the harm of blurring to the uniqueness of the mark. The statute articulates six non-exclusive factors for courts to consider when determining whether defendant's use blurs the ostensibly famous mark, including the extent to which the owner engages in "substantially exclusive use" of the famous mark and the "degree of recognition" of that famous mark.⁸² A mark that is not

76. *Panavision Int'l, L.P. v. Toeppen*, 945 F. Supp. 1296, 1302–03 (C.D. Cal. 1996), *aff'd*, 141 F.3d 1316 (9th Cir. 1998).

77. *See* Nguyen, *supra* note 37, at 170 (reporting cases where courts extended protection to non-famous marks in order to curb the use of deceptive domain names).

78. *Id.* at 187. We return to the question of Panavision's fame in Part II.

79. *See* *Accuride Int'l, Inc. v. Accuride Corp.*, 871 F.2d 1531, 1539 (9th Cir. 1989) (applying California anti-dilution statute) (affirming the district court's holding that widespread third party use of elements of a trademark helped to defeat a dilution claim). *Id.* at 1536 ("There is no evidence that the term 'ACCURIDE,' whether used as a trademark or trade name, is strongly and *uniquely associated* with AII's business. The district court correctly concluded that 'ACCURIDE' is not a strong mark/name.") (emphasis added).

80. Beverly W. Pattishall, *Dawning Acceptance of the Dilution Rationale for Trademark-Trade Identity Protection*, 74 TRADEMARK REP. 289, 301 (1984); *see also* Nancy S. Greiwe, *Antidilution Statutes: A New Attack on Comparative Advertising*, 61 B.U. L. REV. 220, 223 (1981).

81. Pattishall, *supra* note 80, at 308. For cases illustrating the theme, *see, e.g.*, *Accuride Int'l, Inc. v. Accuride Corp.*, 871 F.2d 1531, 1536 (9th Cir. 1989) (holding plaintiff's mark was not "strongly and uniquely associated" with plaintiff's goods or business); *Dreyfus Fund Inc. v. Royal Bank of Canada*, 525 F. Supp. 1108, 1125 (S.D.N.Y. 1981) (noting that plaintiff's marks were "not so unique and arbitrary as to deserve protection in fields totally unrelated to [plaintiff's] activities" but that the statute should be read to protect against dilution in limited areas of use). Other courts applying state anti-dilution statutes did not interpret them to require uniqueness or ubiquity. *See, e.g.*, *Wedgewood Homes, Inc. v. Lund*, 659 P.2d 377, 380 (Or. 1983) (refusing to limit application of the Oregon anti-dilution statute to "marks which are coined, unique or truly famous").

82. 15 U.S.C. § 1125(c)(2)(B)(iii)–(iv).

widely recognized in its substantially exclusive use cannot be harmed by nonexclusive use.⁸³

Survey evidence on fame is used in federal trademark cases to establish sufficient prominence by testing for association. “[S]urveys showing that a large percentage of the general public recognizes the brand . . . would provide evidence of fame.”⁸⁴ Courts have not come to an agreement on a minimum level of recognition required to establish fame, although under the TDRA, niche fame will not suffice.⁸⁵ Commentators have offered their own reasoned opinions. McCarthy once argued that a mark is not famous unless more than 50% of defendant’s customers associate the mark with plaintiff’s goods,⁸⁶ but he has since raised his proposed threshold to 75%.⁸⁷ Shanti instead proposes that 70% of consumers should recognize the mark.⁸⁸ Nguyen proposed a more lenient 40% benchmark.⁸⁹

As these sources illustrate, courts have fumbled toward a metric that requires both sufficient prominence and satisfactorily singular use of the famous mark to signify a unique source for marked products. Those dual benchmarks directly relate to the remedial power extended to the owner of a famous mark. The famous mark is treated as commercially monosemous in scope, reaching any product offered for sale under the mark irrespective of confusion and no matter how different from those offered by the famous mark’s owner. In the next Section, we discuss whether that practically infinite reach should require a showing of virtually unique meaning and consider some of the conceptual difficulties for the singularity requirement in more detail.

C. SINGULARITY, MONOSEMY, AND COMPETING STANDARDS OF FAME

Trademark rights have historically been coupled with injunctive relief.⁹⁰ Two theories are advanced for the centrality of the injunction in trademark cases. First, courts exhibit some reluctance to grant damages in passing off and unfair competition cases because the harms are difficult to establish.⁹¹ Second, damages are characterized as inadequate to remedy the reputational harms inflicted by trademark

83. In many state dilution cases, courts connected remedy to right, noting that anti-dilution protection is extended to preserve the uniqueness of a trademark. Thus, the New York Court of Appeals noted that the danger facing a diluted mark is the loss of “its ability to serve as a unique identifier of the plaintiff’s product.” *Deere & Co. v. MTD Prods., Inc.*, 41 F.3d 39, 43 (2d Cir. 1994).

84. *Thane Int’l, Inc. v. Trek Bicycle Corp.*, 305 F.3d 894, 912 (9th Cir. 2002).

85. *See supra* note 38 and accompanying text.

86. MCCARTHY, *supra* note 51, § 24:92.

87. *Id.* § 24:106.

88. Shanti, *supra* note 15, at 203.

89. Xuan-Thao N. Nguyen, *The New Wild West: Measuring and Proving Fame and Dilution Under the Federal Trademark Dilution Act*, 63 ALB. L. REV. 201, 234 (1999).

90. Jake Linford, *The Path of the Trademark Injunction*, in RESEARCH HANDBOOK ON THE LAW & ECONOMICS OF TRADEMARKS (Glynn Lunney, ed., forthcoming 2022).

91. *See* Mark A. Thurmon, *Confusion Codified: Why Trademark Remedies Make No Sense*, 17 J. INTEL. PROP. L. 245, 268 (2010) (explaining courts at law rarely granted damages in trademark cases unless plaintiff could establish defendant’s fraudulent intent and that courts of equity frequently required the same showing for an accounting of profits).

infringement.⁹² Mark owners were historically entitled to an injunction only in those cases where the allegedly infringing use amounted to fraud or was likely to confuse consumers.⁹³ But the primacy of injunctive relief over damages suggests that the right protected by trademark law is the right of the mark owner not to face confusing competition from sellers of the same or similar products using marks identical or similar to that of the owner.

Like trademark infringement, diluting uses frequently trigger an injunctive response, and damages in rare cases.⁹⁴ To the extent that the menu of remedies for dilution mirror those offered to rectify trademark infringement, protection against infringement and dilution look like equivalent rights. But as explained in Part I.A, protection against dilution reaches a broader category of potential trespasses; diluting uses are subject to sanction irrespective of evidence of confusion.⁹⁵

Consider then what the law communicates by granting an injunction in cases where the alleged infringer uses the owner's mark or a close approximation on widely disparate goods. Whether we think of the remedy as stronger or the right as broader, anti-dilution protection reaches farther than protection against trademark infringement. Indeed, if the allegedly diluting mark is sufficiently similar to the famous mark, anti-dilution protection reaches every point along a virtually infinite spectrum of potential products.

Federal and state anti-dilution laws treat a famous mark as if it were legally monosemous (i.e., as if the sole proper use of the term in the commercial context is to designate goods and services from the famous mark's owner). Such a generous shield is naturally attractive to the mark owner. If every valid trademark qualifies both for protection against passing off and trademark dilution, the rational mark owner will plead and pursue dilution remedies in every case. Protection against dilution fences off a broader swath of the commercial lexicon for the claimant than protection against passing off or unfair competition. Scholars therefore query which—if any—subset of marks should qualify for such robust protection.⁹⁶ The greater the number of marks that qualify for anti-dilution protection, the greater the potential for overprotection.⁹⁷ The lower the fame thresholds, the greater the number

92. *Hogg v. Kirby* (1803), 32 Eng. Rep. 336, 339 (Ch.) (“[A] Court of Equity in these cases is not content with an action for damages; for it is nearly impossible to know the extent of the damage.”).

93. *See supra* note 22 and accompanying text.

94. The Lanham Act's provision for profits, damages and costs, and attorney fees reaches only willful diluting uses. 15 U.S.C. § 1117(a). But courts have power to grant injunctive relief for any use of the famous mark likely to dilute it. 15 U.S.C. § 1116(a).

95. *See* Laura A. Heymann, *The Law of Reputation and the Interest of the Audience*, 52 B.C. L. REV. 1341, 1400 (2011); Jessica Litman, *Breakfast with Batman: The Public Interest in the Advertising Age*, 107 YALE L.J. 1717, 1724 (1999).

96. *See supra* note 5 and accompanying text.

97. *See, e.g.,* *Thane Int'l, Inc. v. Trek Bicycle Corp.*, 305 F.3d 894, 910 (9th Cir. 2002) (recognizing the importance of limiting anti-dilution protection to a narrow subset of marks to avoid “a crippling effect on the marketing of products, as more and more marks would be off limits for use in any market”); *see also* MCCARTHY, *supra* note 51, § 24:104 (citing Milton W. Handler, *Are the State Antidilution Laws*

of marks that qualify. Conversely, higher fame thresholds result in narrower rights for that same set of marks. Thus, as Rebecca Tushnet argues, the virtue of a fame requirement is that “[h]igh standards for fame [and] uniqueness . . . provide limits on the concept of dilution and may allow courts to bar objectionable free riding without expensive and uncertain evidentiary battles.”⁹⁸

The common law presumes that the remedy defines the right.⁹⁹ If an ostensibly famous mark is treated like a *de jure* monoseme, perhaps the right to that remedy should turn on whether the mark is a *de facto* monoseme, or something quite like it.¹⁰⁰ Requiring exclusive use of a famous mark would lead us back to Schechter’s approach, later embraced by Beebe, that anti-dilution protection should extend only to coined marks or words in new combinations. A coined or fanciful mark enters the commercial lexicon as a term with only one meaning—to identify the source of a given product—having literally been created for that singular, source-signifying purpose.¹⁰¹ Trademark doctrine treats these monosemous fanciful terms as inherently source signifying because at first use consumers are deemed unlikely to attribute any other meaning to them.¹⁰²

Courts and litigants in trademark cases might be vexed by the pursuit of a truly monosemous mark. Some linguists strictly reserve monosemy “for words with a

Compatible with the National Protection of Trademarks?, 75 TRADEMARK REP. 269, 280 (1985) (“Without a requirement that the plaintiff’s mark be very strong or famous, an antidilution statute becomes a rogue law that turns every trademark, no matter how weak, into an anti-competitive weapon.”).

98. Tushnet, *supra* note 5, at 566 (citing *I.P. Lund Trading ApS v. Kohler Co.*, 163 F.3d 27, 45–50 (1st Cir. 1998)); see also Jerre B. Swann, Sr., *Dilution Redefined for the Year 2002*, 92 TRADEMARK REP. 585, 602 (2002) (“[T]he undisciplined application of a dilution remedy might foster” claims on “property rights in words.”).

99. Helge Dedek, *The Relationship Between Rights and Remedies in Private Law: A Comparison Between the Common and the Civil Law Tradition* 65, 69, in *TAKING REMEDIES SERIOUSLY* (Robert J. Sharp & Kent Roach eds., 2009); *Kingdom of Spain v. Christie, Manson & Woods Ltd.* [1986] 1 WLR 1120, 1129 (Ch.) (“[Under] English law . . . , a man’s legal rights are in fact those which are protected by a cause of action [and not] separate from the remedy given to the individual.”).

100. Jeremy Sheff argues that Schechter’s proposal camouflaged two theories in tension with one another. Sheff, *supra* note 39, at 338. The “uniqueness theory . . . holds that trademark liability should be imposed to provide manufacturers with the means and incentive to create and preserve consumer goodwill.” *Id.* at 341. The “free-riding theory” instead imposes liability “to prevent second comers from misappropriating the consumer goodwill generated by another’s trademark.” *Id.* But Sheff also recognizes a feedback loop between the potential harm to the famous mark’s uniqueness and the free-riding the law might be focused on preventing. “It is precisely the dual nature of trademark harms—the tendency of *ex ante* manipulations to ripen into point-of-sale manipulations, and vice-versa, in a self-amplifying loop—that allowed Schechter to carry out this theoretical *pas de deux*.” *Id.* at 379.

101. Linford, *Linguistic*, *supra* note 13, at 134–35.

102. Linford, *False Dichotomy*, *supra* note 1, at 1411; MCCARTHY, *supra* note 51, § 11:4. See also Linford, *supra*, at 1405–06:

[C]reating a fanciful mark requires forging a new word from existing morphemes, the smallest building blocks of language. The process of creating a fanciful mark is essentially monosemous. The mark owner adds an entirely new entry into the lexicon by coining a new word from raw linguistic material to identify its product. A fanciful mark is inherently distinctive because it has no other meaning, at least when initially coined.

single meaning, a single referent, and a single grammatical function.”¹⁰³ Others suggest that while a lexicographer might separate core meanings into separate dictionary entries for sake of clarity,¹⁰⁴ those meanings might share a common, singular, monosemous core.¹⁰⁵ A semanticist might thus reasonably consider a word monosemous or quasi-monosemous even though a lexicographer might divide that word into different dictionary entries to make it more useful for the reader.¹⁰⁶ Consider nouns like *window*,¹⁰⁷ or *mouth*.¹⁰⁸ Those words may have metaphoric extensions—*window to the soul*, *mouth of a cave*—but nonetheless retain a core meaning.¹⁰⁹

Most words, however, have multiple meanings. Sometimes those meanings are related to one another, or polysemous.¹¹⁰ Sometimes they stem from different root

103. Rosamund Moon, *Monosemous Words and the Dictionary*, in *DICTIONARY AND THE LANGUAGE LEARNER* 173, 173 (Anthony P. Cowie ed., 1987); see also Dirk Geeraerts, *Reclassifying Semantic Change*, 4 *QUADERNI DI SEMANTICA* 217, 227 (1983). Monosemy goes by other names, including univocality, D.A. CRUSE, *LEXICAL SEMANTICS* 62 (1986), oligosemy, eurysemy, and stenosemy. J.C. Catford, “*Insects are Free*”: *Reflections on Meaning in Linguistics*, 33 *LANGUAGE LEARNING* 13, 24 (1983). A famous mark might exhibit stenosemy in the commercial context, “having a restricted or narrow-ranging meaning.” Catford, *supra*, at 24. Charles Ruhl posits that when a word is monosemous, having a single meaning, there is an accompanying absence of ambiguity. CHARLES RUHL, *ON MONOSEMY: A STUDY IN LINGUISTIC SEMANTICS* (1989).

104. A key challenge for lexicographers is distinguishing between monosemy and polysemy. Some dictionaries will disagree whether a given word is monosemous or polysemous. See HENRI BÉJOINT, *TRADITION AND INNOVATION IN MODERN ENGLISH DICTIONARIES* 190 (1994) [hereinafter BÉJOINT, *TRADITION*] (citing B.T. Atkins, *Building a Lexicon: The Contribution of Lexicography*, 4 *Int’l J. Lexicography* 167 (1991)).

105. Monosemy might not amount to a singular meaning, but instead, “unicity of meaning,” a word with a “number of meanings” that are closely related. See Henri Béjoint, *Monosemy and the Dictionary*, in *PROCEEDINGS OF THE 3RD EURALEX INTERNATIONAL CONGRESS* at 13, 19 (Tamas Magay & J. Zigany eds., 1988) [hereinafter Béjoint, *Monosemy*].

106. BÉJOINT, *TRADITION*, *supra* note 104, at 230 (citing Moon, *supra* note 103, at 173).

107. *Id.* at 229.

108. *Id.* at 230

109. Charles Ruhl, *Data, Comprehensiveness, Monosemy*, in *SIGNAL, MEANING, AND MESSAGE: PERSPECTIVES ON SIGN-BASED LINGUISTICS*, 171, 171 (Wallis Reid, Ricardo Otheguy & Nancy Stern eds., 2002) (arguing many presumed polysemes are instead monosemes with “potentially infinite variations of a single general meaning”). But see Ronald W. Langacker, *Cognitive Grammar*, in *THE OXFORD HANDBOOK OF COGNITIVE LINGUISTICS* 421, 432–33 (2010) (querying whether any single, monosemous sense would necessarily be so abstract as to be effectively unhelpful).

110. For example, meanings of *bank* that relate to storage, including the financial institution, the building where the institution is housed, institutions that store things (like a blood bank) and as a synonym for “rely upon” (i.e., “you can bank on it”). Linford, *False Dichotomy*, *supra* note 1, at 1395 (citing *inter alia* 1 KEITH ALLAN, *LINGUISTIC MEANING* § 3.3.1 (1986) (defining polysemy as “the property of an expression with more than one meaning”); Ekaterini Klepousniotou & Shari R. Baum, *Disambiguating the Ambiguity Advantage Effect in Word Recognition: An Advantage for Polysemous but Not Homonymous Words*, 20 *J. NEUROLINGUISTICS* 1, 4 (2007) (positing that polysemes have a single core meaning from which interrelated senses are derived); see also BÉJOINT, *TRADITION*, *supra* note 104, at 190.

words and are unrelated, or homonymous.¹¹¹ In truth, monosemy, polysemy, and homonymy are not clearly delineated categories,¹¹² but occur on a spectrum.¹¹³ Many cases will likely fall in the polysomous middle of the continuum between the poles of monosomyous and homonymous conditions (singular and unrelated multiple meanings).

It is challenging to measure monosemy. Many linguists argue one cannot show monosemy because it requires proving the negative: the absence of ambiguity between multiple potential senses of a word.¹¹⁴ It is easier to show the polysemous condition of multiple connected meanings.¹¹⁵

Furthermore, language is generally subject to change, and therefore, monosemy might be inherently unstable.¹¹⁶ Language tends to shift away from monosemous

111. *Bank*, referring to the land next to a river and a financial institution, is a prototypical example. Linford, *Linguistic*, *supra* note 13, at 136. See also 1 ALLAN, *supra* note 110, at 147 (defining homonymy as “the relation between two or more expressions which have the same form but different meanings”); Klepousniotou & Baum, *supra* note 110, at 4 (explaining that homonyms have mutually exclusive meanings, one of which must be selected before further processing can occur).

112. Monosemy and polysemy could be described in terms of “minimalist/maximalist” or “abstractivist/cognivist.” See Brigitte Nerlich & David D. Clarke, *Polysemy and Flexibility: Introduction and Overview*, in POLYSEMY: FLEXIBLE PATTERNS OF MEANING IN MIND AND LANGUAGE 3, 14–16 (Brigitte Nerlich, Zazie Todd, Vimala Herman & David D. Clarke eds., 2003).

113. BÉJOINT, TRADITION, *supra* note 104, at 190 (1994); Béjoint, *Monosemy*, *supra* note 105, at 13, 19; Moon, *supra* note 103, at 176 (“The polysemy—or monosemy—of an individual word is not an absolute.”).

114. See William Croft, *Linguistic Evidence and Mental Representations*, 9 COGNITIVE LINGUISTICS 151, 158–59 (1998) (arguing that a monosemous model of language is undermined by cross-linguistic comparison demonstrating grammatical or semantic idiosyncrasy); Dominiek Sandra, *What Linguists Can and Can’t Tell You About the Human Mind: A Reply to Croft*, 9 COGNITIVE LINGUISTICS 361, 367–68 (1998) (concurring with Croft that the monosemous model is a possible but not necessary given existing evidence and linguists must therefore beware a tendency to presume general models are more likely than idiosyncratic models of language storage); David Tuggy, *Linguistic Evidence for Polysemy in the Mind: A Response to William Croft and Dominiek Sandra*, 10 COGNITIVE LINGUISTICS 343, draft at 11 (1999) (noting that as an endpoint on a continuum of possible language models, arguing for a monosemous model amounts to “postulating a negative” which is “notoriously hard to prove”).

115. Barbara Lewandowska-Tomaszczyk, *Polysemy, Prototypes, and Radial Categories*, in THE OXFORD HANDBOOK OF COGNITIVE LINGUISTICS 139, 155 (Dirk Geeraerts & Hubert Cuyckens eds., 2007) (reporting that language change studies reflect a uniform shift from the concrete to the abstract, i.e., from monosemy to polysemy). See also Tuggy, *supra* note 114, at 11. From one perspective, however, any single unit of meaning is monosemous, if it is not ambiguous in context. Wolfgang Teubert, *Units of Meaning, Parallel Corpora, and Their Implications for Language Teaching*, in APPLIED CORPUS LINGUISTICS: A MULTIDIMENSIONAL PERSPECTIVE 171, 174 (U. Connor & T.A. Upton eds., 2004) (comparing the ambiguous word *fire* with the unambiguous and monosemous unit of meaning *enemy fire*). But see RUHL, MONOSEMY, *supra* note 103 at 4 (arguing that every word contributes a single, monosemous semantic meaning to each and every utterance in which it occurs). See also Ryder A. Wishart, *Monosemy in Biblical Studies: A Critical Analysis of Recent Work*, 6 BIBLICAL & ANCIENT GREEK LINGUISTICS 99, 101–03 (2017) (describing Ruhlian monosemy and applying it to analysis of biblical language).

116. Béjoint, *Monosemy*, *supra* note 105, at 20 (“it has often been observed that monosemy is essentially unstable”); JOHN LYONS, LANGUAGE, MEANING, AND CONTEXT 47 (1981).

meaning, from the concrete to the abstract,¹¹⁷ or from the specific to the general.¹¹⁸ Much of the hydraulic pressure of meaning-making will push away from monosemy toward polysemy. Lawyers too might be skeptical of monosemy. As Justice Holmes once opined, “[a] word is not a crystal, transparent and unchanged; it is the skin of a living thought and may vary greatly in color and content according to the circumstances and time in which it is used.”¹¹⁹

Nonetheless, scholars have advanced some useful tests for monosemy. Henri Béjoint proposed a test that considers evidence of ambiguity or antagonism between interpretations.¹²⁰ Under this test, technical or scientific terms are more likely to be categorized as monosemous because they are more likely than the average term to show no antagonism of interpretation.¹²¹

It is unavoidable that speaking of famous marks as monosemes does some violence to both concepts. It is not clear that linguists and lexicographers would or should advocate for treating monosemy as the key to providing access to legal sanctions. Nor do we argue that one should tie anti-dilution protection to the condition of strict monosemy. Such a test might well be under-inclusive, ignoring marks that many would deem famous.

Moreover, corpus evidence probably cannot show true monosemy. Monosemy is the absence of ambiguity with regard to a given word, and it is difficult to prove the negative. An incident of monosemy thus is not falsifiable. But even for skeptics, the occasional brand stands out as truly monosemous, or nigh unto it. Linguist David Crystal offers the Microsoft trademark as an example of a monosemous technical term.¹²² To date, the mark has resisted broadening change that would add meanings to its original source significance.

If we consider a spectrum of singularity to plurality, Microsoft sits quite close to the singular end of the spectrum. There may be other marks that arguably approach a level of singularity and prominence comparable to Microsoft’s. Those that do are also arguably famous marks. The next Part shows how courts and litigants can use corpus evidence to help establish or call into question the singularity and prominence of other ostensibly famous trademarks.

117. Lewandowska-Tomaszczyk, *supra* note 115, at 155.

118. *Id.*

119. Towne v. Eisner, 245 U.S. 418, 425 (1918).

120. Béjoint, *Monosemy*, *supra* note 105, at 20. Such ambiguity can be explored using a zeugma test. *Id.* Zeugma is “a rhetorical figure in which a word or phrase is made to apply, in different senses, to two (or more) others.” *zeugma*, OXFORD ENGLISH DICTIONARY (3d ed. 2018), <https://www.oed.com/view/Entry/232821>. For example, “John and his driver’s license expired last Thursday” highlights the ambiguity in the verb *expire*. CRUSE, *supra* note 103, at 62. The zeugmatic structure shows that *expire* is likely not monosemous. A lack of tension conversely may indicate monosemy. Consider the sentence, “My cousin, who is pregnant, was born on the same day as Arthur’s [cousin], who is the father.” *Id.* Béjoint identifies other words for which it is impossible to show ambiguity, including *monarch*, *dahlia*, *cub*, *decline*, *child*, *stallion*, and *horse*. Béjoint, *Monosemy*, *supra* note 105, at 19. As for *horse*, consider the following sentence: John was shooting up horse and accidentally shot his. Perhaps Béjoint is unfamiliar with drug slang.

121. Béjoint, *Monosemy*, *supra* note 105, at 19.

122. CRYSTAL, *supra* note 19, at 191.

II. USING CORPUS ANALYSIS TO ASSESS FAME

In this Part, we consider whether corpus linguistic evidence might help establish whether a mark is used in a satisfactorily singular manner and has acquired a sufficient level of prominence among the general consuming public to qualify as famous. We consider several potential uses of corpus analysis to identify those dual elements of fame. Ideally, as Justice Thomas Lee and Stephen Mouritsen have argued, “Linguistic corpora can perform a variety of tasks that cannot be performed by human linguistic intuition alone.”¹²³ An analysis of relevant corpora can help interested parties identify a pattern of collocation and concordance of a claimed famous term with goods, services, or other indicia consistent with relatively singular trademark use. Corpus analysis should also reveal sufficiently frequent use to support or refute an argument that the mark is widely recognized among the general consuming public. As we outline in the following sections, our corpus analysis of singularity and prominence suggests that some marks held famous under the FTDA and TDRA likely fall short of being widely recognized by the general consuming public.¹²⁴

The following figure may help concretize the takeaway: To qualify as famous and thus receive anti-dilution protection, a trademark should be used in relatively singular fashion at relatively high frequency by the general consuming public. Corpus linguistic evidence from the sources we describe in the following sections can help courts, trademark examiners, and litigants more quickly and cheaply assess both factors. As we explain below, corpus evidence confirms the intuition that Microsoft is both prominent and singular¹²⁵ but demonstrates that Coach, a mark used for handbags, has not acquired the necessary prominence or singularity to qualify as famous.¹²⁶

123. Lee & Mouritsen, *supra* note 14, at 831.

124. For other examples of applying corpus data to trademark questions, see Alexandra J. Roberts, *Mark Talk*, 39 CARDOZO ARTS & ENT. L.J. (forthcoming 2022); Neal A. Hoopes, *Reclaiming the Primary Significance Test: Dictionaries, Corpus Linguistics, and Trademark Genericide*, 54 TULSA L. REV. 407, 410–11 (2019); Barton Beebe & Jeanne C. Fromer, *Are We Running Out of Trademarks? An Empirical Study of Trademark Depletion and Congestion*, 131 HARV. L. REV. 945 (2018); Quentin J. Ullrich, *Corpora in the Courts: Using Textual Data to Gauge Genericness and Trademark Validity*, 108 TRADEMARK REP. 989, 992–93 (2018); Adam Kilgarriff, *Corpus Linguistics in Trademark Cases*, DICTIONARIES: J. DICTIONARY SOC’Y N. AM., 2015, at 100, 101; Lisa Larrimore Ouellette, *The Google Shortcut To Trademark Law*, 102 CALIF. L. REV. 351, 354 (2014); Paul J. Heald & Robert Brauneis, *The Myth of Buick Aspirin: An Empirical Study of Trademark Dilution by Product and Trade Names*, 32 CARDOZO L. REV. 2533, 2574–75 (2011).

125. See *infra* Parts II.A.1, II.B.1, and II.C.1.

126. See *infra* Parts II.A.3, II.B.6, and II.C.1.

	LESS SINGULAR	MORE SINGULAR
HIGH FREQUENCY	<i>Not famous:</i> Prominent, not singular <i>E.g.</i> , Delta	<i>Famous:</i> Both prominent and singular <i>E.g.</i> , Microsoft
LOW FREQUENCY	<i>Not famous:</i> Neither prominent nor singular <i>E.g.</i> , Coach	<i>Not famous:</i> Singular, not prominent <i>E.g.</i> , Panavision

Figure 1. A framework for the singularity and prominence of famous trademarks.

We consider data primarily from two corpora from the Brigham Young University suite of corpora.¹²⁷ The bulk of our analysis draws on the Corpus of Contemporary American English (COCA), which claims that it is “the largest freely-available corpus of English, and the only large and balanced corpus of American English.”¹²⁸ COCA is a corpus of more than one billion words of text, sampled from transcripts of spoken English, fiction, popular magazines, newspapers, and academic texts, with 25 million words added each year from 1990 through 2019. COCA was assembled by Mark Davies and is hosted by BYU. The corpus is updated regularly. COCA is a corpus of “general English” or, in other words, a corpus meant to capture the language experience of a typical speaker of American English, and is an appropriate dataset given our interest in how an ordinary speaker of English would encounter these marks.

We also draw data from the Corpus of Historical American English (COHA). COHA is “the largest structured corpus of historical English.”¹²⁹ COHA is roughly half the size as COCA, containing more than 475 million words of text, but spans two centuries, with text from the 1810s through the 2000s in decade-long tranches. Compared to COCA, COHA has fewer words from a given time period. Text is drawn from fiction, magazines, newspapers, and non-fiction works.

We consider four types of corpus analysis: collocation, concordance, and synchronic and diachronic measures of frequency. Collocation and concordance are two contextual measures that can be used to help establish whether the use of a mark is substantially singular. If the majority of collocates or the top collocates show

127. We ran initial analyses using the Corpus of Global Web-based English (GloWbE), which contains 1.9 billion words from twenty countries over a limited period (2012–13). Given the limited time period, GloWbE provides a synchronic snapshot, a view of usage of these trademarks on the Internet over the space of a year. CORPUS OF GLOBAL WEB-BASED ENGLISH, <https://perma.cc/YP6C-DQEY> (last visited Nov. 12, 2021). After an initial review of several brands, we found that the two datasets gave us similar results. We ultimately decided to use COCA for the bulk of our contemporaneous analyses as COCA represents a wider range of language use and has more up-to-date data included.

128. CORPUS OF CONTEMPORARY AMERICAN ENGLISH (COCA), <https://www.english-corpora.org/coca/> (last visited Nov. 12, 2021).

129. CORPUS OF HISTORICAL AMERICAN ENGLISH (COHA), <https://perma.cc/Z9SP-FWN5> (last visited Nov. 12, 2021) (“COHA contains more than 400 million words of text from the 1810s–2000s (which makes it 50–100 times as large as other comparable historical corpora of English) and the corpus is balanced by genre decade by decade.”). COHA was funded by a grant from the National Endowment for the Humanities.

trademark use, that suggests the mark is used in a comparatively singular manner and that trademark use is leading other uses.¹³⁰ Similarly, concordance lines provide an excerpt of text surrounding the target word and provide more detail about contextual use.¹³¹

We then consider two techniques to measure the prominence of the claimed mark. Corpus analysis is well suited to disclose the raw frequency of the appearance of a term in the corpus. Frequency evidence can give courts and litigants a relatively quick impression of whether a mark is sufficiently prominent to qualify as famous.¹³² We selected several marks that owners have claimed as famous in litigated cases.

The measurement is relatively simple for coined marks like Xerox or Kodak—one could reasonably expect the majority of uses to point to brand use, and thus, a raw frequency measure will reflect the prominence of the mark. But many owners have claimed fame on behalf of marks derived from existing words. Not every instance of such a word in the corpus will point to brand usage. Consider, for example, Top for tobacco. The raw frequency measure for the token *top* may vastly overstate frequency of mark usage. *Top* is a high frequency word by any measure (262.37 frequency per million words, or fpm), but its use as a trademark for tobacco does not mean that every token of *top* in the corpus points to the tobacco seller. For marks like Top, we correct the frequency data by running a concordance to estimate the likely frequency with which an appearance of the word in question will point to trademark use. We then report an estimated frequency per million.¹³³

In the case of *top*, we found no trademark use of *top* in 498 concordance lines, leaving us with an estimated frequency per million (efpm) of 0.0 for the Top trademark. We then supplemented concordance findings with a collocation search. We found thirty instances of *top* collocated with *tobacco*. Reading the concordance lines revealed one reference to Top rolling papers.¹³⁴ We do not assert that the Top mark is not used in commerce, but its use is virtually unattested in these data, meaning the mark has little or no presence in our dataset.¹³⁵ One might reasonably

130. See *infra* Part II.A.

131. See *infra* Part II.B.

132. See Ronald R. Butters, *A Linguistic Look at Trademark Dilution*, 24 SANTA CLARA COMPUT. & HIGH TECH. L.J. 507, 510–11 (2008) (positing that “a linguist could readily construct, using the normal procedures of lexicography, an evaluation metric that would help a trier of fact to evaluate the degree of fame of a particular mark. Adapting the extant lexicographical methodology to compare, for example, the relative frequency of occurrence in a LexisNexis search for instances of Burger King, Apple, and Microsoft versus Kleenex, Xerox, and Frisbee will yield a useful linguistic measure of relative ‘fame.’”).

133. See *infra* Part II.C.

134. The following concordance line was extracted from Kenneth Howe, *American Nomads*, S.F. CHRON., Jan. 14, 1996 (search conducted March 3, 2021):

... from rain. Scattered between them lay a litter of cardboard, discarded packages of **Top** Cigarette tobacco wrappers, Darigold milk cartons, Taco Bell hot sauce packets and heaps ...

135. Lee & Mouritsen, *supra* note 14, at 840:

But if *vehicle* is never used to refer to *bicycle* or *airplane* in the corpus data, then we may end up with an even further extension of our frequency continuum from *possible but rare* to *possible but unattested*.

expect a stronger showing from a mark that is in fact “widely recognized by the general consuming public.”¹³⁶

We finally consider how evidence from COCA and COHA can also be analyzed diachronically, providing us with a view of changing frequency over time.¹³⁷ Changes in frequency might offer clues to when a mark acquired fame. More provocatively, the evidence indicates some formerly famous marks may be falling below a minimum prominence benchmark. Thus, some marks that were once truly renowned might fall far enough out of regular usage that they are no longer sufficiently prominent to qualify as famous.¹³⁸

Relying on evidence of frequency or singularity alone would mislead the factfinder and fail to establish fame. A showing of singularity but low frequency may establish niche fame, but under the TDRA, niche fame is insufficient to qualify a mark for anti-dilution protection.¹³⁹ Corpus linguistics has similarly struggled with a frequency fallacy.¹⁴⁰ Critics have questioned whether frequency sufficiently correlates with ordinariness, and similar challenges could be raised to the correlation of frequency to fame.¹⁴¹ Moreover, looking only at raw data of the frequency of a word’s appearance in a corpus would not reveal whether those appearances all point back to the same singular source signifier, like Microsoft; multiple brands using the same mark for disparate goods, like Delta for faucets and airlines;¹⁴² or primarily non brand use, like Coach.¹⁴³

Before jumping to the conclusion that the *airplane* and *bicycle* uses of *vehicle* are entirely unattested in the corpora or the language at large, however, we should evaluate the use of *vehicle* in the concordance data.

136. 15 U.S.C. § 1125(c)(2)(A).

137. 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, 192 (2012) [hereinafter Mouritsen, *Hard Cases*]. But see Quentin Feltgen, *Diachronic Emergence of Zipf-like Patterns in Construction-Specific Frequency Distributions: A Quantitative Study of the Way Too Construction*, 16 DIACHRONIC LEXICAL SEMANTICS 1, 3 (2020), <https://perma.cc/HZ95-TFU3> (arguing COCA’s limited time horizon makes it unsuitable for diachronic investigation, but COHA is “one of the most reliable databases available for this purpose”).

138. See *infra* Part II.D.

139. See *supra* note 37 and accompanying text.

140. Stefan Th. Gries & Brian G. Slocum, *Ordinary Meaning and Corpus Linguistics*, 2017 B.Y.U. L. REV. 1417, 1459 n.181 (2017) (“As is well-known among corpus linguists, using only frequency as a measure of the commonness of a word, an expression, or a meaning is treacherous since words with very similar or even identical frequencies in a corpus can be very unevenly dispersed.”).

141. See, e.g., Shlomo Klapper, *(Mis)judging Ordinary Meaning?: Corpus Linguistics, the Frequency Fallacy, and the Extension-Abstraction Distinction in “Ordinary Meaning” Textualism*, 8 BRIT. J. AM. LEGAL STUD. 327, 352 (2019).

142. See *supra* Part II.C.2.

143. See *supra* Parts II.C.3, II.D.4.

A. COLLOCATION AND SINGULARITY

When using a search engine like COCA, one enters a search term called a node.¹⁴⁴ One can search for tokens or occurrences of the node within the database to estimate the frequency with which the node appears. COCA can compare tokens in different genres (academic, newspaper, spoken, fiction) and across years.

One can also search for collocates: words that appear near the node. Words that frequently appear near the node help one understand how the word is used in context and what it might mean.¹⁴⁵ Collocation is “the habitual juxtaposition” or association of a given word with other words at a frequency higher than chance.¹⁴⁶ As John Firth put it, “[c]ollocations are actual words in habitual company.”¹⁴⁷ Phillips and White refer to collocates as “word neighbors” that might reveal meaning consistent with the canon of *noscitur a sociis*: a word “is known by its associates.”¹⁴⁸ Collocation thus reveals evidence of ordinary usage of the target word.¹⁴⁹

COCA defaults to display the first one hundred collocates of a word in order of statistical frequency.¹⁵⁰ COCA uses a default minimum frequency to prevent the appearance of one-off collocates. COCA also weights collocates according to the statistical likelihood of the given word appearing at any point in the corpus using a metric known as Mutual Information (MI). The more frequently the collocate appears near the node compared to its expected appearance, the more comfortably we can conclude the collocation sheds light on the meaning of the node in context.¹⁵¹ In other words, a high MI score suggests that the node word and collocate are statistically more likely to appear together. A lower MI score indicates that the two words don’t associate as frequently. For example, the words *torrential* and *rain* have a high MI score because *torrential* is almost exclusively used to talk about rain and other weather patterns.

144. See, e.g., D. Carolina Núñez, *War of the Words: Aliens, Immigrants, Citizens, and the Language of Exclusion*, 2013 B.Y.U. L. REV. 1517, 1523 (2013).

145. See HUNSTON, *supra* note 16, at 68–69 (“Collocation is the tendency of words to be biased in the way they co-occur.”).

146. *Collocation*, OXFORD ENGLISH DICTIONARY ONLINE, <https://www.oed.com/view/Entry/36400> (last visited Dec. 10, 2021); Stefan Th. Gries, *Corpus-Based Methods and Cognitive Semantics: The Many Senses of To Run*, in CORPORA IN COGNITIVE LINGUISTICS: CORPUS-BASED APPROACHES TO SYNTAX AND LEXIS, 57, 60 (Anatol Stefanowitsch ed., 2006).

147. John R. Firth, *A Synopsis of Linguistic Theory, 1930–1955*, in STUDIES IN LINGUISTIC ANALYSIS 1, 14 (1957).

148. James Cleith Phillips & Sara White, *The Meaning of the Three Emoluments Clauses in the U.S. Constitution: A Corpus Linguistic Analysis of American English from 1760–1799*, 59 S. TEX. L. REV. 181, 200 (2017).

149. Stephen C. Mouritsen, *The Dictionary Is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning*, 2010 B.Y.U. L. REV. 1915, 1963 (2010) [hereinafter Mouritsen, *Fortress*]. Although we do not apply it to that purpose here, collocation data may also reveal change in meaning rather than change in frequency. See, e.g., Jake Linford, *Datamining the Meaning(s) of Progress*, 2017 B.Y.U. L. REV. 1531, 1552 (2017) [hereinafter Linford, *Datamining*].

150. Mouritsen, *Fortress*, *supra* note 149, at 1963.

151. See, e.g., Jennifer L. Mascott, *Who Are “Officers of the United States”?*, 70 STAN. L. REV. 443, 506 n.355 (2018).

We begin with collocates of claimed famous marks Microsoft, Blackberry, Delta, and Coach. Evidence of collocation may indicate whether a word claimed as a famous mark has a substantially singular use, multiple related polysemous meanings, disparate homonymous meanings, or no indications of source significance among its most frequent collocates.

If, as is the case with Microsoft, many or all of the top collocates of a trademark point to the owner, its goods and services, and its non-infringing competitors, the use of the mark is relatively singular. Collocates reflecting third party use or nonmark use would indicate the absence of singularity.

1. Coined Marks: Microsoft

Linguists have pointed to Microsoft as the prototypical monosemous or technical term.¹⁵² Collocation confirms that conclusion: Microsoft's most frequent collocates all refer back to the firm, its products, its non-infringing competitors, or its regulatory context. In particular, the four words statistically most likely to appear with Microsoft refer specifically to Microsoft products: Windows, Software, Office, and Word. Collocation provides a quick snapshot showing that use of Microsoft is substantially singular, if not entirely unique.

Rank	Collocate	Instances ¹⁵³	Mutual Information Score (MI) ¹⁵⁴
1	WINDOWS	1728	6.76
2	OFFICE	861	4.03
3	SOFTWARE	668	5.64
4	GOOGLE	630	5.80
5	WORD	625	3.86
6	APPLE	587	5.41
7	MICROSOFT	558	6.47
8	SURFACE	430	4.72
9	CORP	371	6.52
10	GATES	343	6.18
11	EXCEL	338	8.26

152. CRYSTAL, *supra* note 19, at 191.

153. "Instances" refers to the number of co-occurrences of collocates within four words of the target, controlled for the expected frequency of the collocate. COCA uses frequency in its tables, but we use that term for other purposes in this article.

154. Kenneth Ward Church & Patrick Hanks, *Word Association Norms, Mutual Information, and Lexicography*, 16 COMPUTATIONAL LINGUISTICS 22, 23 (1990) (explaining that a mutual information score "compares the probability of observing [word] *x* and [word] *y* together (the joint probability) with the probabilities of observing [word] *x* and [word] *y* independently (chance)"); see also Thomas R. Lee & James C. Phillips, *Data-Driven Originalism*, 167 U. PA. L. REV. 261, 302 (2019).

Rank	Collocate	Instances ¹⁵³	Mutual Information Score (MI) ¹⁵⁴
12	OPERATING	332	5.30
13	COMPANIES	303	3.19
14	INTERNET	284	3.77
15	PRODUCTS	272	3.85
16	OUTLOOK	256	6.46
17	SONY	253	6.90
18	XBOX	249	7.97
19	INTEL	246	6.86
20	SERVER	207	6.05

Table 1. Microsoft, Top 20 Collocates, COCA.

2. Arbitrary Marks: Blackberry and Delta

Blackberry and Delta are both well-known marks. Blackberry was once a ubiquitous pager and texting device. Delta is well-known for airlines. Both marks have been held famous in litigated disputes,¹⁵⁵ and both are marks derived from pre-existing terms. As courts have noted, fame is difficult to prove, particularly where “the mark is a common English word that has different meanings in different contexts.”¹⁵⁶ Collocation and concordance data can reveal whether an arbitrary mark has become sufficiently singular that fame is likely, or whether the claimed term has so many competing meanings that it is not likely famous. Collocation data suggests that between these marks, the use of Blackberry is more singular than the use of Delta. Of the top twenty collocates, the top two, three of the top five, and eight of the remaining fifteen all point to uses relating to Blackberry’s tech devices. The other nine collocates all point to the fruit, with terms like pick, bush, and similar fruits like raspberry and blueberry.

Rank	Collocate	Instances	MI
1	PHONE	248	4.55
2	DEVICE	196	5.61
3	RASPBERRY	157	9.77
4	SMARTPHONE	154	8.42
5	BLUEBERRY	120	9.34
6	ANDROID	113	7.35
7	RIM	109	7.64

155. *Rsch. in Motion Ltd. v. Defining Presence Mktg. Grp., Inc.*, 102 U.S.P.Q.2d 1187, 2012 WL 893481 (T.T.A.B. Feb. 27, 2012); *Delta Air Lines, Inc. v. Influence Direct, LLC*, No. 3-14-0926, No. 3-14-1112, 2016 WL 310068 (M.D. Tenn. Jan. 15, 2016).

156. *Coach Servs., Inc. v. Triumph Learning, LLC*, 668 F.3d 1356, 1369 (Fed. Cir. 2012).

Rank	Collocate	Instances	MI
8	IPHONE	100	6.70
9	DESKTOP	91	7.46
10	USER	85	4.28
11	E-MAIL	80	4.11
12	FRESH	78	4.31
13	PICK	75	2.83
14	BUSH	72	6.52
15	WILD	68	4.42
16	APP	66	3.77
17	STRAWBERRY	66	7.28
18	FARM	64	4.34
19	JAM	62	6.99
20	PLAYBOOK	62	9.09

Table 2. Blackberry, Top 20 Collocates, COCA.

Delta, on the other hand, is plausibly subject to diluting use by, among others, the use of the mark Delta for faucets by the Delta Faucet Corporation.¹⁵⁷ At least one commentator argues that in the face of widespread third-party use, a mark is not famous and “its strength has already been diluted.”¹⁵⁸

Our evidence suggests, however, that Delta is on the cusp of commercial singularity. For Delta, five of the top twenty collocates refer to the air travel brand: Air, Lines, Airlines, Flight, and Northwest (which merged with Delta in 2008).¹⁵⁹ The first collocate, Delta within four of Delta, refers at times to the airline,¹⁶⁰

157. DELTA, Registration No. 717,404.

158. Christopher L. Buongiorno, *Evidence of Fame and Dilution Before the Trademark Trial and Appeal Board*, 29 AIPLA Q. J. 1, 32 (2001); see also *Carnival Corp. v. SeaEscape Casino Cruises, Inc.*, 74 F. Supp. 2d 1261, 1271, (S.D. Fla. 1999) (arguing that multiple uses of “fun” for marks in various industries cut against plaintiff’s dilution claim); *TCPIP Holding Co. v. Haar Commc’ns, Inc.*, 244 F.3d 88, 96 (2d Cir. 2001) (“It seems unlikely that Congress could have intended that the holders of such non-distinctive marks [like American] would be entitled to claim exclusivity for them throughout all areas of commerce.”); *Michael Caruso & Co., Inc. v. Estefan Enters., Inc.*, 994 F. Supp. 1454, 1463 (S.D. Fla. 1998), *aff’d*, 166 F.3d 353 (11th Cir. 1998) (“[E]xtensive third party use of the word ‘bongo’ undermines the inherent distinctiveness” of the mark.); *Sports Auth., Inc. v. Abercrombie & Fitch, Inc.*, 965 F. Supp. 925, 941 (E.D. Mich. 1997) (holding that “third-party use of ‘authority,’” irrespective of competition, “diminishes any ‘distinctive or famous’ aspects of” the mark and renders it “not so famous as to deserve protection under the [FTDA]”); *Trs. of Columbia Univ. v. Columbia/HCA Healthcare Corp.*, 964 F. Supp. 733, 744–45, 750, (S.D.N.Y. 1997) (holding that fame of the mark “Columbia” for healthcare services “has been seriously undermined by third party use of the same or similar marks” both within the health care industry and in other industries).

159. John Crawley, *Delta Buys Northwest To Create Biggest Airline*, TIMES OF MALTA (Nov. 1, 2008), <https://perma.cc/W3VA-AARU>.

160. “. . . 35 percent. Fitzgerald’s new ad strategy will debut in early 1992. **DELTA TRAFFIC RISES: Delta** Air Lines reported a 17.4 percent gain in traffic during September. . . .”

geography,¹⁶¹ mathematical equations,¹⁶² measurements,¹⁶³ and fraternities or sororities.¹⁶⁴ The rest of the list of top collocates is dominated by geographical uses. This use is not as singular as the brand-related collocates that dominate the Microsoft list, but all of the brand-related use appears here to point back to Delta. Competing use from Delta Faucet does not appear among the top collocates.

Rank	Collocate	Instances	MI
1	DELTA	562	8.73
2	AIR*	558	4.69
3	MISSISSIPPI	555	8.02
4	LINES*	497	5.78
5	RIVER	425	5.52
6	FORCE	362	4.53
7	AIRLINES*	339	7.73
8	PHI	231	10.20
9	BLUES	203	6.92
10	=	187	3.91
11	FLIGHT*	183	5.10
12	NIGER	170	9.73
13	GAMMA	170	8.81
14	QUADRANT	143	9.61
15	OKAVANGO	137	12.13
16	KAPPA	135	9.66
17	SIGMA	135	9.25
18	NORTHWEST*	132	6.37
19	REGION	128	3.95
20	MEKONG	116	11.16

Table 3. Delta Top 20 Collocates, COCA.

3. Coach: The Infamous Non-famous Mark

Coach is a leading brand for handbags.¹⁶⁵ But the Court of Appeals for the Federal Circuit held that Coach failed to present sufficient evidence that the mark was famous before the date the defendant began selling its non-confusing educational test-prep

161. “Arts and Education Program is to continue the great musical tradition born in the Mississippi **Delta**: the **Delta** blues. The Arts and Education program teaches students to play music. . .”

162. “. . . Beta 4 = (1- Lambda) # (19) ln S not = **Delta** 0 + **Delta** 1 - ln P t -**Delta** 2 ln U not - **Delta**. . .”

163. “. . . W/m2. The resulting **delta** temperature is, in application of Wien law, # **delta** T = 0.1885 **delta** F = 0.75 AK (for a black body). . .”

164. “. . . jet to fly Matt and Shartrina from Florida to Indy. A couple of his **Delta** Tau **Delta** fraternity brothers from Butler hopped into White’s specially equipped van and drove . . .”

165. Phil Wahba, *Coach Thinks Outside the Bag*, FORTUNE, May 24, 2017 (reporting Coach’s 2011 high-water mark, leading the U.S. handbag market with a 29.2% share).

materials.¹⁶⁶ In a different case, however, Coach prevailed on a motion for default judgment and secured a permanent injunction against a counterfeiter.¹⁶⁷ In its ruling in the counterfeiting case, a federal district court held that Coach had sufficiently alleged facts supporting that its mark was “famous,” and that the defendant began selling counterfeit marks after the Coach mark became famous.¹⁶⁸

The data on Coach bears out the holding of the Federal Circuit, in large part. In COCA, there are few references to the fashion brand. We found one fashion-related collocation to handbag in COCA, with sixty-nine occurrences (MI = 4.94).¹⁶⁹ Handbags can be collocated with *coach*, but that use is not common or frequent.

Rank	Collocates (COCA)	Frequency	MI
1	HEAD	8266	9.88
2	PLAYER	2997	4.37
3	ASSISTANT	2874	7.36
4	FOOTBALL	2848	5.75
5	TEAM	2129	2.96
6	COLLEGE	1517	3.13
7	MIKE	1237	8.71
8	SEASON	1118	2.76
9	ATHLETE	861	4.82
10	JOHN	780	8.53
11	BILL	686	2.79
12	HIRE	647	3.67
13	SPORT	583	2.72
14	OFFENSIVE	548	4.65
15	FIRE	521	3.07
16	MANAGER	494	2.81
17	JOE	485	9.66
18	TEAMMATE	483	5.42
19	DEFENSIVE	457	4.86
20	SOCCER	447	4.81

Table 4. Coach, Top 20 Collocates, COCA.

One might wonder whether the concordance measure is merely tracking the difference between fanciful marks and arbitrary marks, which the law treats as inherently source signifying, and descriptive marks, which are not treated that way. BlackBerry, Delta, and Coach are all arbitrary marks—there is no inherent connection between the marks and the respective goods or services—but as explained above, each of those marks show different results with regard to potential commercial singularity.

166. Coach Servs., Inc. v. Triumph Learning, LLC, 668 F.3d 1356, 1376 (Fed. Cir. 2012).

167. Coach, Inc. v. Bags Accessories, No. 10-2555, 2011 WL 1882403 (D.N.J. May 17, 2011).

168. *Id.* at *8–10. Notably, in making this determination, the court cited to FTDA precedent and the FTDA version of section 43(c).

169. We also found two fashion-related collocates in GloWbE, *handbags* (N=82, MI=7.71) & *bags* (N=65, MI=3.83). Neither were among the top fifty collocates for *coach*. For more on GloWbE, see *supra* note 127.

B. CONCORDANCE AND SINGULARITY

We also consider concordance, or key words in context (KWIC). Collocates provide a snapshot, but concordance lines present the word in fuller context.¹⁷⁰ COCA and COHA provided randomized samples from their respective databases. We analyzed KWIC displays as another check for monosemy and trademark usage for the terms *microsoft*, *panavision*, *apple*, *delta*, *trek* and *coach*.¹⁷¹ We coded KWIC results in one of the following four categories: target brand usage; non-brand-related use; third-party brand use; and generic use of the target brand.¹⁷² Interrater reliability was high at 99% and disagreements were resolved after further discussion and calibration.

This section offers test cases that demonstrate the process using concordance evidence to examine the target word in context. It's admittedly a rough first cut. If you were hoping to introduce this as evidence in litigation, you might reasonably take more time. But if you were trying to predict the likelihood that a court could or should conclude a mark is used more or less monosemously, analyzing KWIC would get you a good head start, and in some instances could lead a party to settle or even abandon a weak dilution claim.

1. Microsoft: Monolithic and Monosemous

We read 999 concordance results for *microsoft*. We coded 992 as brand usage, six as non-brand usage, and one as usage related to another brand. We coded none of the uses as generic. There is little doubt that the use of Microsoft, at least as it appears in the corpus, is effectively monosemous. Given the frequency per million of the mark in COCA, the mark easily clears both the prominence and singularity thresholds.

2. Panavision: A KWIC Look at a Short Concordance

Panavision International holds registered trademarks for Panavision for motion picture cameras, lenses, and photographic equipment.¹⁷³ Panavision claimed anti-dilution protection under the FTDA and prevailed in its case against a cybersquatter who initially registered the Panavision.com domain name.¹⁷⁴ As we will discuss

170. Mouritsen, *Fortress*, *supra* note 149, at 1958; Mouritsen, *Hard Cases*, *supra* note 137, at 201 (citing HUNSTON, *supra* note 16, at 79) ("It is tempting, when looking at a list of collocates, to draw conclusions about the overall frequency of compounds and phrases that may not be justified.")

171. We requested either 1000 or 500 concordance lines, where available, but have found that COCA frequently under-delivers by a line or two. Panavision appeared very infrequently in the corpus, so we analyzed all twenty-seven instances in COCA.

172. If consumers come to use a mark primarily as a generic designator for products or services, rather than as a source signifier, the mark will no longer qualify for trademark protection. *See, e.g.*, Linford, *Linguistic*, *supra* note 13, at 144. *See also infra* notes 298–299 and accompanying text.

173. *See, e.g.*, PANAVISION, Reg. No. 834,705.

174. *Panavision Int'l L.P. v. Toepfen*, 141 F.3d 1316, 1327 (9th Cir. 1998).

below, Panavision is a low frequency mark,¹⁷⁵ but it is a highly singular mark. All twenty-seven *panavision* tokens in COCA pointed back to the Panavision trademark. Panavision's concordance results look like those for Microsoft. If singularity were the sole requirement, Panavision would be entitled to protection as a famous mark. But as we discuss in Part C, the niche fame Panavision enjoys is not enough to qualify as famous.

3. Apple: Conflicted and Compelling

We read 999 KWIC results for *apple*. We coded 439 tokens as referring to Apple trademark use or referring to Apple Corp. (0.439). Another seven tokens were competing brand use (0.007), including uses of Big Apple, which New York City and other entities or firms might attempt to claim as part of a trademark.¹⁷⁶ The other 554 tokens were non-brand use (0.555), mostly referring to apples as fruit.

Based on these results, Apple is the type of arbitrary mark that many might suspect would achieve actionable trademark fame.¹⁷⁷ The overwhelming majority of the mark-related tokens (over 98%) point to Apple Corp. Generic fruit-related use could reasonably coexist with Apple's claim. But if strict monosemy is the test, Apple's use is not monosemous. In fact, brand use doesn't account for the majority of the tokens. That pattern should hold for nearly any mark drawn from a preexisting term, and it is similarly true of Delta, discussed next.

4. Delta: Not Quite Alone in Its Friendly Skies

As discussed above, Delta is potentially subject to diluting third-party use from the Delta Faucet Corporation, among others.¹⁷⁸ But faucets made no appearance among the top collocates, nor did they appear among the concordance lines we reviewed. Delta Airlines tokens were the dominant brand use (139/498, 0.27), but did not approach a majority of tokens. There was limited third-party use from mixed sources (17/498, 0.03). A court inclined to treat Apple's use as sufficiently singular to qualify as famous could reach the opposite conclusion about Delta, in light of heavier crowding by competitors and the lack of dominance of Delta Airlines in the corpus as a whole.

175. See *supra* Part II.A.

176. Brandon Marsh, *ICANN't Help Myself: Beneficial Adjustments to the New Generic Top-Level Domain Name Expansion Process*, 95 J. PAT. & TRADEMARK OFF. SOC'Y 195, 204 (2013).

177. As discussed above, for our frequency benchmarks, we estimate tokens and frequency per million by coding concordances and counting a proportional fraction of the total. See *supra* notes 133–135 and accompanying text. Apple is a high frequency token (33.39 fpm) and a ratio of 0.439 would still yield an estimated 14.65 fpm. That is sufficiently high frequency to meet our most stringent prominence benchmark, discussed in Part II.C.2, *infra*, if one reads the singular *trademark* use as sufficient to meet the substantial singularity standard.

178. See *supra* notes 157–158 and accompanying text.

5. Trek: Lagging Behind the Dominant Player

The manufacturer of Trek bicycles sought anti-dilution protection but was unable to persuade a court that its mark had acquired fame. Standing in Trek's way was the notable use of Trek in reference to the Star Trek television and film series. Our coding of *trek* supports that result. We read only twenty-four tokens related to Trek bicycles, a mere 0.048 of 500 concordance lines. Moreover, nearly every instance was from a bicycling magazine, suggesting that Trek had acquired, at best, niche fame. Conversely, Star Trek tokens dominated commercial uses across those same lines with 289 tokens (0.578). The majority of references pointed back to the Star Trek franchise. The use of Trek for bicycles was neither singular nor dominant.

6. Coach: Off-brand

Concordance analysis was unhelpful to a potential claim from the makers of the Coach handbag. We found no brand use in 1,000 concordance results. For a further refinement on how concordance and collocation might assist a factfinder in determining a likely ratio of trademark use for a high-frequency term adopted as a descriptive or arbitrary mark, see Part III.B.

7. A Spectrum of Singularity

The following spectrum suggests one of four broad categories of trademark singularity. The farther to the right, the more confidence a court or litigant might have that the mark in question is sufficiently singular to qualify as famous.

Singularity	Not Dominant	Commercially Dominant, Not Singular	Commercially Singular	Substantially Singular
<i>Exemplary Marks</i>	<i>Coach</i> <i>Trek</i>	<i>Delta</i>	<i>Apple</i> <i>Blackberry</i>	<i>Microsoft</i> <i>Panavision</i>

Figure 2. Spectrum of singularity.

C. FREQUENCY AND PROMINENCE

When attempting to establish actual recognition of a famous mark among the consuming public, litigants typically amass either survey evidence or the equivalent of press clippings to show unsolicited mentions by third parties. Corpus data may provide similarly probative information.¹⁷⁹ Indeed, survey evidence approximates

179. Jacoby proposed frequency as a proxy for fame, using a count of citations in press releases. Jacob Jacoby, *Considering the Who, What, When, Where and how of Measuring Dilution*, 24 SANTA CLARA COMP. & HIGH TECH. L.J. 600, 606 (2008).

what corpus data might plausibly show—the likelihood that the mark in question is prominent and renowned among the general consuming public.¹⁸⁰ In this part, we share data regarding ninety-seven litigated trademarks and propose benchmarks that courts and litigants could use to estimate whether a given mark is prominent and renowned among the general consuming public.

1. Frequency Analysis: Exemplary Famous Marks

We ran frequency analysis on ninety-seven unique marks listed in the well-regarded McCarthy treatise as exemplary marks held famous and not famous under both the FTDA and TDRA.¹⁸¹ Three listed marks—Audi, Nike, and Victoria’s Secret—were found famous in cases litigated under both statutes. We omitted from our analysis marks for which the fame was grounded in trade dress or non-word elements, like the adidas three-stripe mark,¹⁸² which the corpus is not suited to measure. We also omitted the Mc- family of marks, as the corpus doesn’t provide a ready way to search for a set of terms sharing, in this case, a common prefix.¹⁸³ As discussed above, for those marks derived from common, high-frequency words like Top for tobacco, we analyzed approximately 500 concordance lines, coding each line for target mark use or other use. We then took the percentage of target mark usage from those concordance lines to estimate a percentage of total tokens pointing to the target mark, our estimated frequency per million.¹⁸⁴

Our list includes thirty-nine marks courts held not famous: twenty-three analyzed under the FTDA and sixteen analyzed under the TDRA. Our list also includes sixty-one marks held famous (including the repeated Nike, Audi, and Victoria’s Secret marks): thirty under the TDRA and thirty-one under the FTDA. This frequency data provides several interesting insights. As shown below, the frequency per million of the selected marks that were found famous (*Table 5*) was significantly higher than the frequency per million of the selected marks found not famous (*Table 6*).

We conducted an independent t-test to compare frequency per million words in COCA for marks deemed famous and marks deemed not famous. To the extent our results deviated from the outcomes in previous litigated cases, frequency analysis might fail to predict likely outcomes in future cases. A significant difference was found between famous marks ($M=1.54$, $SD=2.49$) and non-famous marks ($M=0.27$, $SD=1.22$) where $t(97)=3.01627$, $p=0.001634$, significant at $p < 0.05$. Effect size was calculated using Cohen’s d where $d=0.657716$, which can be interpreted as a

180. *Id.* at 617–18.

181. See MCCARTHY, *supra* note 51, at §§ 24:107–110 (listing in order marks found famous under the TDRA and FTDA, and then marks found not famous under the TDRA and FTDA).

182. *adidas Am., Inc. v. Skechers USA, Inc.*, 890 F.3d 747 (9th Cir. 2018) (affirming finding that the three-stripe sports shoe trade dress was “famous,” but reversing preliminary injunction against infringement of the three-stripe trade dress for failure to prove irreparable injury).

183. See, e.g., *McDonald’s Corp. v. McSweet LLC*, 112 U.S.P.Q.2d 1268, 2014 WL 5282256, at *1 (T.T.A.B. Sept. 29, 2014).

184. See *supra* notes 132–136 and accompanying text.

medium effect size.¹⁸⁵ This suggests that there is a meaningful difference between the frequency of use of famous and non-famous marks in COCA. The following box plot helps visualize our findings. The minimum, first quartile, median, and third quartile of the non-famous marks all cluster at 0 fpm.

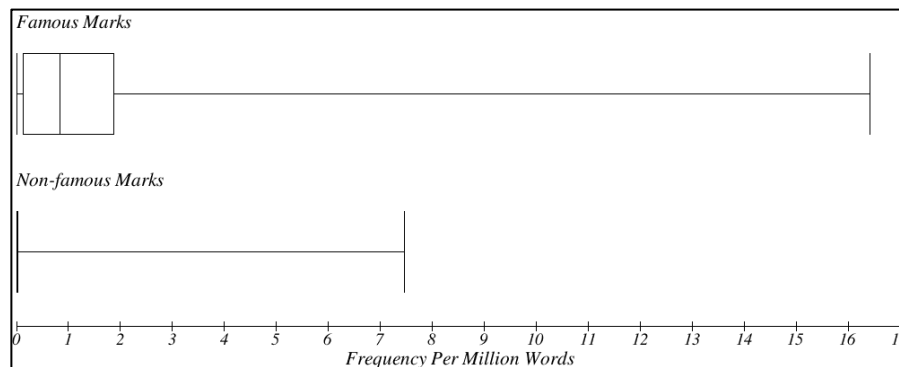


Figure 3. Box plot of frequency of famous and non-famous marks.

Perhaps surprisingly, given the criticisms of judicial application of the FTDA fame factors, we found no significant difference between the results in cases when courts found a mark famous under the FTDA ($M=1.62$, $SD=3.27$) compared to cases where the court found a mark famous under the TDRA ($M=1.35$, $SD=1.44$).¹⁸⁶ Ostensibly different standards under the FTDA and TDRA do not appear to drive differences in frequency rates of marks found famous.

As shown below, the highest per million frequency for any of the non-famous marks was Columbia, claimed by the New York City university as a mark used in connection with the provision of medical and healthcare services,¹⁸⁷ with an estimated frequency per million of 7.46. App Store, which was found to be a generic designation for stores selling software applications,¹⁸⁸ follows with a frequency per million of 1.91. The Trek mark for bicycles follows with an estimated 0.49 fpm, and Sports Authority's use of Authority at an estimated 0.26 fpm. The remaining non-famous marks all appeared in COCA with a frequency of 0.1 per million or less. Several marks held famous also appeared less frequently than 0.1 per million, including Newport (0.097 estimated fpm), Big Gulp (0.09 efpm), and Panavision (0.04 fpm). Many marks held famous appear less frequently than Trek's estimated 0.49 per million, including Bentley (0.48 efpm), Velveeta (0.20 fpm), Miss USA (0.17 fpm), and Candyland (0.15 fpm).

185. See generally Jacob Cohen, *A Power Primer*, 112 PSYCH. BULL. 155 (1992).

186. The t -value is 0.39466. The p -value is 0.347339. The result is *not* significant at $p < 0.05$.

187. *Trs. of Columbia Univ. v. Columbia/HCA Healthcare Corp.*, 964 F. Supp. 733, 736 (S.D.N.Y. 1997).

188. *Apple, Inc. v. Amazon.com, Inc.*, No. C 11-1327, 2011 WL 2638191 (N.D. Cal. July 6, 2011).

Microsoft was not listed among McCarthy's exemplary marks, but compared to the other marks held famous, Microsoft is a significant outlier (28.16 fpm). That is the same approximate frequency as the noun *survival*, reported in one summary as the 2,836th most frequently-appearing word in the corpus.¹⁸⁹ A slightly different list includes Microsoft as the 3,192nd most frequently-used word in the corpus.¹⁹⁰

Table 5. Marks held Famous	Year ¹⁹¹	Tokens	fpm	Concordance	Estimated Tokens ¹⁹²	Estimated fpm
Ford ¹⁹³	2002	34006	34.24	240/500	16322	16.43
Nike ¹⁹⁴	2007 /2003	5639	5.68			5.68
NASDAQ ¹⁹⁵	2003	5421	5.46			5.46
Starbucks ¹⁹⁶	2009	5257	5.29			5.29
Blackberry ¹⁹⁷	2012	4619	4.65			4.65
AOL ¹⁹⁸	1998	3901	3.93			3.93
Delta ¹⁹⁹	2016	12801	12.89	144/498	3701.5	3.73
Barbie ²⁰⁰	1998	3314	3.34			3.34
Chanel ²⁰¹	2014	2897	2.92			2.92
The Sporting News ²⁰²	2000	285	2.9			2.9

189. WORD FREQUENCY DATA, <https://perma.cc/2VVA-D2M5> (last visited Nov. 8, 2021) (listing the top 60,000 lemmas—the root of a word form).

190. *Id.* (listing approximately the top 220,000 word forms that occur at least twenty times in the corpus and in at least five different texts).

191. Year of the first case making a fame determination. Where two years are listed, those are marks for which courts found fame under both the TDRA and the FTDA.

192. For marks derived from a high-frequency word, we examined approximately 500 concordance lines to estimate the percentage of the total tokens that likely reflected trademark use, and then estimated a frequency based on that proportion. *See also supra* notes 132–136 and accompanying text.

193. Ford Motor Co. v. Lloyd Design Corp., 184 F. Supp. 2d 665 (E.D. Mich. 2002).

194. Nike, Inc. v. Nikepal Int'l, Inc., No. 2:05-cv-1468, 2007 WL 2782030 (E.D. Cal. Sept. 8, 2007).

195. NASDAQ Stock Mkt., Inc. v. Antarctica, S.R.L., 69 U.S.P.Q.2d 1718, 2003 WL 22021943 (T.T.A.B. June 30, 2003).

196. Starbucks Corp. v. Wolfe's Borough Coffee, Inc., 588 F.3d 97, 105 (2d Cir. 2009).

197. *Rsch. in Motion Ltd. v. Defining Presence Mktg. Grp., Inc.*, 102 U.S.P.Q.2d 1187, 2012 WL 893481 (T.T.A.B. Feb. 27, 2012).

198. America Online, Inc. v. IMS, 24 F. Supp. 2d 548 (E.D. Va. 1998).

199. Delta Air Lines, Inc. v. Influence Direct, LLC, No. 3-14-0926, No. 3-14-1112, 2016 WL 310068 (M.D. Tenn. Jan. 15, 2016).

200. Mattel Inc. v. Jcom, Inc., No. 97 Civ. 7191, 1998 WL 766711 (S.D.N.Y. Sept. 11, 1998).

201. Chanel, Inc. v. Makarczyk, 110 U.S.P.Q.2d 2013, 2014 WL 2531211 (T.T.A.B. May 27, 2014).

202. Times Mirror Mags., Inc. v. Las Vegas Sports News, L.L.C., 212 F.3d 157 (3d Cir. 2000), *cert. denied*, 531 U.S. 1071 (2001).

Table 5. Marks held Famous	Year ¹⁹¹	Tokens	fpm	Concordance	Estimated Tokens ¹⁹²	Estimated fpm
James Bond ²⁰³	1998	2794	2.81			2.81
Pepsi ²⁰⁴	2007	2692	2.71			2.71
Porsche ²⁰⁵	1997	2601	2.62			2.62
Viagra ²⁰⁶	2009	2280	2.3			2.3
Episcopal Church ²⁰⁷	2019	1946	1.96			1.96
Visa ²⁰⁸	2008	7488	7.54	119/500	1782.14	1.78
Prozac ²⁰⁹	2000	1548	1.56			1.56
Audi ²¹⁰	2008 /2006	1520	1.53			1.53
Motown ²¹¹	2011	1473	1.48			1.48
Versace ²¹²	2018	1329	1.34			1.34
Louis Vuitton ²¹³	2008	1263	1.27			1.27
Calvin Klein ²¹⁴	2017	1140	1.15			1.15
TiVo ²¹⁵	2018	1047	1.05			1.05
Rolex ²¹⁶	2011	1013	1.02			1.02
Budweiser ²¹⁷	1996	964	0.97			0.97
Victoria's Secret ²¹⁸	2008 /2003	966	0.96			0.96

203. Danjaq LLC v. Sony Corp., No. CV97-841, 1998 WL 957053 (C.D. Cal. July 29, 1998), *aff'd without op.*, 165 F.3d 915, (9th Cir. 1998).

204. PepsiCo, Inc. v. #1 Wholesale, LLC, No. 07-CV-367, 2007 WL 2142294 (N.D. Ga. July 20, 2007).

205. Porsche Cars N. Am., Inc. v. Manny's Porshop, Inc., 972 F. Supp. 1128 (N.D. Ill. 1997).

206. Pfizer Inc. v. Sachs, 652 F. Supp. 2d 512 (S.D.N.Y. 2009).

207. VonRosenberg v. Lawrence, 412 F. Supp. 3d 612 (D.S.C. 2019).

208. Visa Int'l Serv. Ass'n v. JSL Corp., 590 F. Supp. 2d 1306 (D. Nev. 2008), *aff'd*, 610 F.3d 1088, (9th Cir. 2010).

209. Eli Lilly & Co. v. Nat. Answers, Inc., 86 F. Supp. 2d 834, 847 (S.D. Ind. 2000), *aff'd*, 233 F.3d 456 (7th Cir. 2000).

210. Audi AG v. D'Amato, 469 F.3d 534, 547 (6th Cir. 2006).

211. UMG Recordings, Inc. v. Mattel, Inc., 2011 WL 5014005, at *18 (T.T.A.B. Sept. 30, 2011).

212. Gianni Versace, S.p.A., v. Versace 19.69 Abbigliamento Sportivo SRL, 328 F. Supp. 3d 1007, 1022 (N.D. Cal. 2018), *appeal dismissed*, 2019 WL 2005744 (9th Cir. 2019).

213. Louis Vuitton Malletier S.A. v. Haute Diggity Dog, LLC, 507 F.3d 252, 257, 265 (4th Cir. 2007).

214. Coty Inc. v. Excell Brands, LLC, 277 F. Supp. 3d 425, 460 (S.D.N.Y. 2017).

215. TiVo Brands LLC v. Tivoli, LLC, 2018 WL 6921323 (T.T.A.B. Dec. 31, 2018).

216. Rolex Watch U.S.A., Inc. v. AFP Imaging Corp., 2011 WL 6780738 (T.T.A.B. Dec. 5, 2011), *appeal dismissed*, 2012 WL 3306996 (Fed. Cir. Aug. 14, 2012); Rolex Watch U.S.A., Inc. v. Rolex Deli Corp., 2012 WL 5177517 (S.D.N.Y. Oct. 18, 2012).

217. Anheuser-Busch, Inc. v. Andy's Sportswear, Inc., 1996 WL 657219 (N.D. Cal. Aug. 28, 1996).

218. Moseley v. V Secret Catalogue, Inc., 537 U.S. 413 (2003).

Table 5. Marks held Famous	Year ¹⁹¹	Tokens	fpm	Concordance	Estimated Tokens ¹⁹²	Estimated fpm
Winston ²¹⁹	2004	6843	6.89	64/499	877.66	0.88
Cartier ²²⁰	2000	861	0.87			0.87
Tylenol ²²¹	2000	839	0.84			0.84
Citibank ²²²	2010	829	0.83			0.83
Burberry ²²³	2009	670	0.67			0.67
Rolls-Royce ²²⁴	2016	649	0.65			0.65
Harley-Davidson ²²⁵	2018	615	0.62			0.62
Hotmail ²²⁶	1998	587	0.59			0.59
Bentley ²²⁷	2013	2505	2.52	95/500	475.95	0.48
Camel ²²⁸	2004	3670	3.7	62/500	455	0.458
Polo ²²⁹	1998	4062	4.09	49/500	398.08	0.4
Toys "R" Us ²³⁰	1996	351	0.35			0.35
Etch-A-Sketch ²³¹	1998	213	0.21			0.21
Velveeta ²³²	2002	203	0.2			0.2
Miss USA ²³³	2009	166	0.17			0.17
Candyland ²³⁴	1996	146	0.15			0.15

219. R.J. Reynolds Tobacco Co. v. Premium Tobacco Stores, Inc., 2004 WL 1613563 (N.D. Ill. July 19, 2004).

220. Cartier, Inc. v. Deziner Wholesale, L.L.C., 2000 WL 347171 (S.D.N.Y. Apr. 3, 2000).

221. McNeil Consumer Brands, Inc. v. U.S. Dentek Corp., 116 F. Supp. 2d 604 (E.D. Pa. 2000).

222. Citigroup Inc. v. Cap. City Bank Grp., Inc., 2010 WL 595586 (T.T.A.B. Feb. 17, 2010), *aff'd*, 637 F.3d 1344 (Fed. Cir. 2011).

223. Burberry Ltd. v. Euro Moda, Inc., No. 08 CIV. 5781 (CMAJP), 2009 WL 4432678 (S.D.N.Y. Dec. 4, 2009).

224. Rolls-Royce Motor Cards Ltd. v. Davis, 2016 WL 3913640 (D.N.J. Mar. 11, 2016).

225. H-D U.S.A., LLC v. SunFrog, LLC, 311 F. Supp. 3d 1000, 1041 (E.D. Wis. 2018).

226. Hotmail Corp. v. Van\$ Money Pie Inc., 1998 WL 388389 (N.D. Cal. Apr. 16, 1998).

227. Bentley Motors Ltd. Corp. v. McEntegart, 976 F. Supp. 2d 1297, 1313 (M.D. Fla. 2013).

228. R.J. Reynolds Tobacco Co. v. Premium Tobacco Stores, Inc., No. 99 C 1174, 2004 WL 1613563 (N.D. Ill. July 19, 2004).

229. Polo Ralph Lauren L.P. v. Schuman, No. Civ.A. H97-1855, 1998 WL 110059 (S.D. Tex. Feb. 9, 1998).

230. Toys "R" Us, Inc. v. Akkaoui, No. C 96-3381, 1996 WL 772709 (N.D. Cal. Oct. 29, 1996).

231. Ohio Arts Co. v. Watts, No. 98 CV 7338, 1998 WL 34072046 (N.D. Ohio June 23, 1998).

232. Kraft Foods Holdings, Inc. v. Helm, 205 F. Supp. 2d 942 (N.D. Ill. 2002).

233. Miss Universe, L.P., LLLP v. Villegas, 672 F. Supp. 2d 575 (S.D.N.Y. 2009).

234. Hasbro, Inc. v. Internet Ent. Grp., Ltd., No. C96-130WD, 1996 WL 84853 (W.D. Wash. Feb. 9, 1996).

Table 5. Marks held Famous	Year ¹⁹¹	Tokens	fpm	Concordance	Estimated Tokens ¹⁹²	Estimated fpm
WaWa ²³⁵	1996	140	0.14			0.14
The Greatest Show on Earth ²³⁶	1996	122	0.12			0.12
Just Do It ²³⁷	2011	3622	3.65	15/500	108.66	0.11
Newport ²³⁸	2012	3998	4.03	10/500	97.96	0.097
Big Gulp ²³⁹	2007	168	0.17	90/167	90	0.09
The Other White Meat ²⁴⁰	2010	52	0.05			0.05
Panavision ²⁴¹	1996	44	0.04			0.04
The House That Ruth Built ²⁴²	2015	43	0.04			0.04
Jews for Jesus ²⁴³	1998	40	0.04			0.04
Lexington ²⁴⁴	1998	3434	3.46	1/498	6.9	0.01
Arthur the Aardvark ²⁴⁵	1999	3	0			0
Intermatic ²⁴⁶	1996	2	0			0
America's Team ²⁴⁷	2009	1	0			0
Don't Leave Home Without Us ²⁴⁸	1996	1	0			0

235. Wawa Dairy Farms v. Haaf, No. 98 CV 7338, 1998 WL 34072046 (E.D. Pa. 1996), *aff'd*, 939 F.3d 1032 (3d Cir. 1997).

236. Ringling Bros.-Barnum & Bailey Combined Shows, Inc. v. B.E. Windows Corp., 937 F. Supp. 204 (S.D.N.Y. 1996); Ringling Bros.-Barnum & Bailey Combined Shows, Inc. v. Utah Div. of Travel Dev., 955 F. Supp. 605 (E.D. Va. 1997), *aff'd*, 170 F.3d 449 (4th Cir. 1999).

237. Nike, Inc. v. Maher, 49 U.S.P.Q.2d 1957, 2011 WL 3828723 (T.T.A.B. Aug. 9, 2011).

238. Lorillard Tobacco Co. v. Cal. Imports, LLC, 886 F. Supp. 2d 529, 535 (E.D. Va. 2012).

239. 7-Eleven Inc. v. Wechsler, 83 U.S.P.Q.2d 1715, 2007 WL 1431084 (T.T.A.B. May 15, 2007).

240. Nat'l Pork Board v. Supreme Lobster & Seafood Co., 96 U.S.P.Q.2d 1479, 2010 WL 2513872 (T.T.A.B. June 11, 2010).

241. Panavision Int'l, L.P. v. Toeppen, 945 F. Supp. 1296, 1302-03 (C.D. Cal. 1996), *aff'd*, 141 F.3d 1316 (9th Cir. 1998).

242. N.Y. Yankees P'ship v. IET Prods. & Servs., Inc., 114 U.S.P.Q.2d 1497, 2015 WL 2455162 (T.T.A.B. May 8, 2015).

243. Jews for Jesus v. Brodsky, 993 F. Supp. 282 (D.N.J. 1998), *aff'd without opinion*, 159 F.3d 1351 (3d Cir. 1998).

244. Lexington Mgmt. Corp. v. Lexington Cap. Partners, 10 F. Supp. 2d 271 (S.D.N.Y. 1998).

245. Brown v. It's Ent., Inc., 34 F. Supp. 2d 854 (E.D.N.Y. 1999).

246. Intermatic, Inc. v. Toeppen, 947 F. Supp. 1227 (N.D. Ill. 1996).

247. Dall. Cowboys Football Club, Ltd. v. America's Team Properties, Inc., 616 F. Supp. 2d 622 (N.D. Tex. 2009).

248. Am. Express Co. v. CFK, Inc., 947 F. Supp. 310 (E.D. Mich. 1996).

Table 5. Marks held Famous	Year ¹⁹¹	Tokens	fpm	Concordance	Estimated Tokens ¹⁹²	Estimated fpm
Nailtiques ²⁴⁹	1997	1	0			0
NYC Triathlon ²⁵⁰	2010	0	0			0

Table 6. Marks held Not Famous	Year	Tokens	fpm	Concordance	Estimated Tokens	Estimated fpm
Columbia ²⁵¹	1997	24798	24.97	148/499	7354	7.46
App Store ²⁵²	2011	1895	1.91			1.91
Trek ²⁵³	2002	10229	10.3	24/500	490	0.49
Authority ²⁵⁴	1997	66367	66.83	2/499	266	0.26
Clue ²⁵⁵	1999	15771	15.88	3/499	94.81	0.1
Washington Speakers Bureau ²⁵⁶	1999	6	0.01			0.1
Shaklee ²⁵⁷	2019	69	0.07			0.07
Blue Man Group ²⁵⁸	2008	62	0.06			0.06
The Children's Place ²⁵⁹	2001	48	0.05			0.05
Rocky Top ²⁶⁰	2014	39	0.04			0.04
It's a 10 ²⁶¹	2013	26	0.03			0.03
Petro ²⁶²	1997	339	0.34	26/337	26	0.026

249. Nailtiques Cosmetic Corp. v. Salon Scis., Corp., No. 96-2709-CIV-NESBITT, 1997 WL 244746 (S.D. Fla. Jan. 10, 1997).

250. N.Y.C. Triathlon, LLC v. NYC Triathlon Club, Inc., 704 F. Supp. 2d 305 (S.D.N.Y. 2010).

251. Trs. of Columbia Univ. v. Columbia/HCA Healthcare Corp., 964 F. Supp. 733, 750 (S.D.N.Y. 1997).

252. Apple, Inc. v. Amazon.com Inc., No. C 11-1327 PJH, 2011 WL 2638191 (N.D. Cal. July 6, 2011).

253. Thane Int'l, Inc. v. Trek Bicycle Corp., 305 F.3d 894, 912 (9th Cir. 2002).

254. Sports Auth., Inc. v. Abercrombie & Fitch, Inc., 965 F. Supp. 925 (E.D. Mich. 1997).

255. Hasbro, Inc. v. Clue Computing, Inc., 66 F. Supp. 2d 117 (D. Mass. 1999), *aff'd*, 232 F.3d 1 (1st Cir. 2000).

256. Wash. Speakers Bureau Inc. v. Leading Auths. Inc., 33 F. Supp. 2d 488 (E.D. Va. 1999), *aff'd*, 217 F.3d 843 (4th Cir. 2000).

257. Superior Consulting Servs., Inc. v. Shaklee Corp., No. 616CV2001ORL31GJK, 2019 WL 913374 (M.D. Fla. Feb. 25, 2019), *aff'd*, No. 19-10771, 2021 WL 4438518 (11th Cir. Sept. 28, 2021).

258. Blue Man Prods., Inc. v. Tarmann, 75 U.S.P.Q.2d 1811, 2005 WL 2034544 (T.T.A.B. Aug. 18, 2005), *rev'd*, No. 05-2037, 2008 WL 6862402 (D.D.C. Apr. 3, 2008).

259. TCPIP Holding Co., Inc. v. Haar Communications, Inc., 244 F.3d 88 (2d Cir. 2001).

260. House of Bryant Publ'ns, LLC v. City of Lake City, 30 F. Supp. 3d 711 (E.D. Tenn. 2014).

261. It's a 10, Inc. v. Beauty Elite Grp., Inc., No. 13-60154-CIV, 2013 WL 6834804 (S.D. Fla. Dec. 23, 2013).

262. Petro Shopping Ctr. L.P. v. James River Petroleum, No. CIV.A. 3:96CV530, 1997 WL 187335 (E.D. Va. Jan. 17, 1997), *aff'd on other grounds*, 130 F.3d 88 (4th Cir. 1997).

Table 6. Marks held Not Famous	Year	Tokens	fpm	Concordance	Estimated Tokens	Estimated fpm
Dentsply ²⁶³	2020	24	0.02			0.02
Star Market ²⁶⁴	1996	23	0.02			0.02
Fun Ship ²⁶⁵	1999	17	0.02			0.02
Avery Dennison ²⁶⁶	1999	12	0.01			0.01
Tycos ²⁶⁷	2002	5	0.01			0.01
Stealth ²⁶⁸	1998	3467	3.49	1/500	6.934	0.00698
Appleseed ²⁶⁹	1997	297	0.3	2/296	2	0.00203
We ²⁷⁰	2004	518157 2	5217. 98	0/499	0	0
Top ²⁷¹	2007	260566	262.4	0/498	0	0
Coach ²⁷²	2012	87013	87.62	0/1000	0	0
Charlotte ²⁷³	2008	18146	18.27	0/498	0	0
Tornado ²⁷⁴	1998	5791	5.83	0/498	0	0
Majestic ²⁷⁵	2012	2760	2.78	0/498	0	0
Dearest ²⁷⁶	1999	2352	2.37	0/499	0	0
Lingo ²⁷⁷	2011	1188	1.2	0/495	0	0

263. Dentsply Sirona, Inc. v. Dental Brands for Less LLC, No. 15 CIV. 8775 (LGS), 2020 WL 1643891 (S.D.N.Y. Apr. 2, 2020).

264. Star Mkts., Ltd. v. Texaco, Inc., 950 F. Supp. 1030 (D. Haw. 1996).

265. Carnival Corp. v. SeaEscape Casino Cruises, Inc., 74 F. Supp. 2d 1261 (S.D. Fla. 1999).

266. Avery Dennison Corp. v. Sumpton, 189 F.3d 868 (9th Cir. 1999).

267. Welch Allyn Inc. v. Tyco Int'l Servs. AG, 200 F. Supp. 2d 130 (N.D.N.Y. 2002).

268. S Indus., Inc. v. Diamond Multimedia Sys., Inc., 991 F. Supp. 1012 (N.D. Ill. 1998).

269. Appleseed Found. Inc. v. Appleseed Inst., Inc., 981 F. Supp. 672 (D.D.C. 1997).

270. WE Media, Inc. v. Gen. Elec. Co., 218 F. Supp. 2d 463 (S.D. N.Y. 2002), *aff'd*, 94 Fed. Appx. 29 (2d Cir. 2004). The corpus couldn't produce a collocation for a high frequency word like *we*, so we instead analyzed collocations of *magazine* within four words of *we*, and found no results pointing to plaintiff's trademark for magazines.

271. Top Tobacco, L.P. v. N. Atl. Operating Co., 509 F.3d 380, 384 (7th Cir. 2007).

272. Coach Servs., Inc. v. Triumph Learning LLC, 668 F.3d 1356, 1376 (Fed. Cir. 2012).

273. GMA Accessories, Inc. v. Croscill, Inc., No. 06 CIV. 6236 GEL, 2008 WL 591803 (S.D.N.Y. Mar. 3, 2008).

274. Breuer Elec. Mfg. Co. v. Hoover Co., No. 97 C 7443, 1998 WL 427595, 48 U.S.P.Q.2d 1705 (N.D. Ill. July 23, 1998).

275. Perfect Pearl Co., Inc. v. Majestic Pearl & Stone, Inc., 887 F. Supp. 2d 519, 540–41 (S.D.N.Y. 2012).

276. Something Old, Something New, Inc. v. QVC, Inc., No. 98 CIV. 7450 SAS, 1999 WL 1125063 (S.D.N.Y. Dec. 8, 1999).

277. Lingo v. Lingo, 785 F. Supp. 2d 443 (D. Del. 2011).

Table 6. Marks held Not Famous	Year	Tokens	fpm	Concordance	Estimated Tokens	Estimated fpm
Bongo ²⁷⁸	1998	513	0.51	1/494	1.03	0
King of the Mountain ²⁷⁹	1997	77	0.08	0/76	0	0
We'll Pick You Up ²⁸⁰	2001	47	0.05	0/47	0	0
Bio-Safe ²⁸¹	2007	3	0			0
Marco's Pizza ²⁸²	2019	2	0			0
Movie Mania ²⁸³	2014	2	0			0
Teton Glacier ²⁸⁴	2002	1	0			0
Weather Guard ²⁸⁵	1997	1	0			0
Millionaire's Reading Club ²⁸⁶	2019	0	0			0
Steak-Umm ²⁸⁷	2011	0	0			0
Timberstone ²⁸⁸	2017	0	0			0
Lane Capital Management ²⁸⁹	1998	0	0			0

2. Proposing Prominence Benchmarks

One challenge facing a factfinder is how to determine whether a mark appears with sufficient frequency in a corpus to indicate that mark is “widely known” among consumers. Psycholinguistics literature measures frequency effects on language

278. Michael Caruso & Co. v. Estefan Enters., Inc., 994 F. Supp. 1454, 1463 (S.D. Fla. 1998), *aff'd without opinion*, 166 F.3d 353 (11th Cir. 1998).

279. King of the Mountain Sports, Inc. v. Chrysler Corp., 968 F. Supp. 568, 577–78 (D. Colo. 1997), *aff'd on other grounds*, 185 F.3d 1084 (10th Cir. 1999).

280. Advantage Rent-A-Car, Inc. v. Enter. Rent-A-Car, Co., 238 F.3d 378 (5th Cir. 2001).

281. Biosafe-One, Inc. v. Hawks, 524 F. Supp. 2d 452, 466 (S.D.N.Y. 2007).

282. Marco's Franchising LLC v. Marco's Coal Fired Pizza Inc., No. 17-CV-2550-MSK-NYW, 2019 WL 4645431 (D. Colo. Sept. 23, 2019).

283. Movie Mania Metro, Inc. v. GZ DVD's Inc., 857 N.W.2d 677, 687 (2014).

284. Nat'l Distillers Prod. Co., LLC v. Refreshment Brands, Inc., 198 F. Supp. 2d 474, 486 (S.D.N.Y. 2002).

285. Knaack Mfg. Co. v. Rally Accessories, Inc., 955 F. Supp. 991 (N.D. Ill. 1997).

286. Springboards To Educ., Inc. v. Hous. Indep. Sch. Dist., 912 F.3d 805, 818 (5th Cir. 2019), *as rev'd*, (Jan. 29, 2019), *as rev'd*, (Feb. 14, 2019).

287. Steak Umm Co., LLC v. Steak 'Em Up, Inc., No. CIV.A. 09-2857, 2011 WL 3679155 (E.D. Pa. Aug. 23, 2011).

288. Idaho Golf Partners, Inc. v. TimberStone Mgmt., LLC., No. 1:14-CV-00233-BLW, 2017 WL 3531481 (D. Idaho Aug. 17, 2017).

289. Lane Cap. Mgmt., Inc. v. Lane Cap. Mgmt., Inc., 15 F. Supp. 2d 389 (S.D.N.Y. 1998), *aff'd*, 192 F.3d 337 (2d Cir. 1999).

acquisition and recall,²⁹⁰ and to that end, researchers group words into high and low frequencies.²⁹¹ Frequency counts depend in part on the size of the corpus, with counts normalized in frequency per million words.²⁹² In the 1940s, Thorndike & Lorge proposed benchmarks of above 37.0 fpm and below 1.0 fpm for high and low frequency words, respectively.²⁹³ Recent decades have embraced standardized measures with different benchmarks. Low frequency words are typically defined as having less than 5.0 fpm (e.g., gloom, frenzy, objection) and high frequency words as having more than 100.0 fpm (e.g., energy, market, area).²⁹⁴

As shown in the chart of frequency measures above, nearly every investigated mark falls below the 5.0 fpm benchmark. The exceptions are famous marks Ford, Nike, NASDAQ, and Starbucks. Columbia is the only mark held not famous with an estimated frequency above 5.0 fpm. None of the examined marks would qualify as a high frequency word under any traditional frequency count measure. That doesn't necessarily mean none are prominent trademarks.

In the last decade, these higher benchmarks have been challenged for the simple reason that most of the words in large corpora appear with a frequency below 1.0 fpm.²⁹⁵ Recent work proposes a slightly more forgiving threshold, with low frequency words below 1.0 fpm and high frequency words above 10.0 fpm.

Just under half of the listed marks found famous (24/59) have an estimated prominence above 1.0 fpm. Thus, using a 1.0 benchmark for a prominence threshold would not entirely undercut the last twenty-five years of fame holdings. In addition, two of the marks found not famous (Columbia, App Store) have an estimated prominence above 1.0 fpm. However, strictly applying that 1.0 fpm benchmark would leave several potentially surprising candidates on the outside looking in: Budweiser (0.98 fpm); Victoria's Secret (0.97 fpm); Winston (0.88 fpm); Cartier (0.87 fpm); Tylenol (0.84 fpm); and Citibank (0.83 fpm).

290. See, e.g., Ben Ambridge, Evan Kidd, Caroline F. Rowland & Anna L. Theaktson, *The Ubiquity of Frequency Effects in First Language Acquisition*, 42 J. CHILD. LANG. 239 (2015); HERMANN EBBINGHAUS, *MEMORY: A CONTRIBUTION TO EXPERIMENTAL PSYCHOLOGY* (1913); Marc Brysbaert, Matthias Buchmeier, Markus Conrad, Arthur M. Jacobs, Jens Bölte & Andrea Böhl, *The Word Frequency Effect: A Review of Recent Developments and Implications for the Choice of Frequency Estimates in German*, 58 EXPERIMENTAL PSYCH. 412 (2011); Marc Brysbaert & Michael J. Cortese, *Do the Effects of Subjective Frequency and Age of Acquisition Survive Better Word Frequency Norms?*, 64 Q.J. EXPERIMENTAL PSYCH. 545 (2011). But see Tom Roeper, *What Frequency Can Do and What It Can't*, in FREQUENCY EFFECTS IN LANGUAGE ACQUISITION: DEFINING THE LIMITS OF FREQUENCY AS AN EXPLANATORY CONCEPT 23 (Insa Güllow & Natalia Gagarina eds., 2007).

291. See, e.g., Marc Brysbaert, Pawel Mandera & Emmanuel Keuleer, *The Word Frequency Effect in Word Processing: An Updated Review*, 27 CURRENT DIRECTIONS IN PSYCH. SCI. 45 (2018).

292. *Id.* at 46.

293. Richard W. Howell & Harold J. Vetter, *High and Low Frequency Nouns as Sources of Hesitation in the Production of Speech*, 12 PSYCHONOMIC SCI. 157, 157 (1968) (citing EDWARD LEE THORNDIKE & IRVING LORGE, *THE TEACHER'S WORD BOOK OF 30,000 WORDS* (1944)).

294. Brysbaert, Mandera & Keuleer, *Word Frequency Effect*, *supra* note 291, at 46.

295. *Id.*; cf. Walter J.B. van Heuven, Pawel Mandera, Emmanuel Keuleers & Marc Brysbaert, *SUBTLEX-UK: A New and Improved Word Frequency Database for British English*, 67 Q.J. EXPERIMENTAL PSYCH. 1176, 1179–80 (2014).

Turning to our data, we find that for marks held famous, the median frequency is 0.87 per million. In other words, half of the marks determined to be famous had frequencies above this benchmark. We suggest that a frequency over 0.87 provides strong evidence that the mark is famous. Marks below this benchmark may also be deemed famous but may require additional evidence to establish their fame.

We can also draw insights from the frequency findings for the listed non-famous marks. Here, we found a median frequency of 0.0, meaning that at least half of the non-famous marks did not appear in the corpus. While there were also a handful of famous marks that had 0.0 frequencies, the majority of marks found famous had at least some presence in the corpus. We therefore suggest that if a mark occurs at a rate of 0.0 hits per million words, it is quite unlikely to qualify as famous.

As illustrated in the box plot above, there is some slight overlap between the most frequent quartile of non-famous marks and the lowest quartile of famous marks. This indicates a grey area in determining fame and is possibly an indication that previous courts have not maintained consistent standards for determining fame. Thus, some deserving trademarks may not have received protection while some non-qualifying marks may have been extended protection. However, on the whole, the overlap in frequencies of a handful of famous and non-famous marks is fairly small. As illustrated by the box-and-whisker plots, there appears to be a notable difference between the bulk of the famous marks and the bulk of the non-famous marks. We consider this promising evidence that frequency information can be useful in determining if a mark is famous or not.

Thus, our three potential fame benchmarks are the externally suggested 5.0 fpm and 1.0 fpm, and the internally derived 0.87 fpm. But these corpus tools are not anchored to any particular benchmark. A court might reasonably conclude our preferred benchmark goes beyond or falls short of standards established by existing precedent. Whatever standard courts apply, parties can use corpus analysis to argue that a mark meets or fails to meet the articulated standard, or to propose why one standard surpasses another.

Benchmark (fpm)	Prominence	Exemplars
$x \geq 5.0$	Highly likely prominent	Microsoft Ford Nike
$5.0 < x \leq 1.00$	Likely prominent	Blackberry Porsche Louis Vuitton
$1.00 < x \leq 0.87$	Strong evidence of prominence	Budweiser Victoria's Secret Cartier
$0.87 < x \leq 0.00$	Ambiguous to weak evidence of prominence	Tylenol Bentley Panavision
~ 0.00	Possible but unattested prominence	Lexington Arthur the Aardvark NYC Triathlon

Figure 4. Prominence Benchmarks.

As discussed above, frequency analysis helps correct for the niche fame that might be mistaken for wide recognition if one relied only on contextual evidence. For example, the Panavision mark looks like Microsoft through the lenses of collocation and concordance. If evidence of singularity were the only requirement, Panavision would easily qualify. But Panavision is a term that shows up in the corpus as a whole at a much lower frequency than Microsoft. Its relatively low frequency in COCA—0.04 per million—suggests the mark does not currently have prominence necessary for a court to conclude it is widely recognized.²⁹⁶ The collocation profile points to singularity, but the frequency does not suggest fame. Panavision’s frequency is so low that it probably should not have been held famous. Assuming Panavision was famous in its niche, that niche fame no longer qualifies a mark for anti-dilution protection.²⁹⁷

D. PROMINENCE OVER TIME

In this section, we provide some data for how the frequency with which a term appears in a corpus can change over time. In particular, we provide some evidence of frequency changes over time for four marks: AOL, Kodak, Xerox, and Starbucks. The first three marks have shown frequencies at some point in COHA and/or COCA that provide solid evidence of their prominence. Nonetheless, token frequency for each of those marks has dipped in recent years, raising the possibility that they might still be source signifying but no longer famous. In comparison to the first three marks, the Starbucks mark continues to rocket heavenward, likely acquiring sufficient prominence to qualify as famous around the turn of the twenty-first century.

The question is what courts should do with such evidence. The Lanham Act provides a mechanism for canceling a registration when the mark becomes a generic designator for the goods or services offered.²⁹⁸ Courts have also recognized a loss of enforceable common law rights if the mark is shown to have slid into genericness.²⁹⁹ Conversely, there is no statutory mechanism for the loss of fame, although courts have denied anti-dilution protection to marks that have not acquired fame prior to dilutive use.³⁰⁰ Nonetheless, if courts intend to enforce the act as written, evidence that a mark is no longer widely recognized by the general consuming public should call into question the suitability of anti-dilution protection. That evidence would be important for courts to note and perhaps difficult for them

296. One might wonder whether Panavision would appear more frequently in COHA, which has data extending back to the launch of the brand. But Panavision does not appear once among COHA’s 405 million words.

297. See *supra* note 37 and accompanying text.

298. 15 U.S.C. § 1064(3); see also Linford, *Linguistic*, *supra* note 13, at 125.

299. Jake Linford, *Valuing Residual Goodwill After Trademark Forfeiture*, 93 NOTRE DAME L. REV. 811, 822–25 (2017).

300. See, e.g., *Coach Servs., Inc. v. Triumph Learning LLC*, 668 F.3d 1356, 1376 (Fed. Cir. 2012).

to credit: The gravity of previous eras of fame might well draw judicial notice and lead courts or juries to err in favor of a mark owner whose fame is dwindling.

Schechter himself provides an excellent example of the phenomenon. When his article was published in 1927, he offered what he must have concluded was an obvious example of a prominent mark with a singular meaning: Blue Goose for fruit.³⁰¹ Schechter posited that:

when the public hears or sees the phrase “Blue Goose” it thinks, not of “a North American wild goose having a grayish plumage resembling that of the young snow goose found chiefly during its migrations in the Mississippi Valley,” but of oranges or grapefruit with a certain trademark and certain meritorious qualities.³⁰²

The venerable Blue Goose makes barely a ripple on the waters of COCA, with seventeen *blue goose* tokens in the last thirty years, none of which refer to the citrus sellers. None of fifteen tokens in COHA refer to the fruit vendors, either. A court might have extended anti-dilution to Blue Goose in the late 1920s but would likely refuse to do so today.

1. AOL: What Goes Up Must Come Down

Corpus evidence indicates the AOL trademark was particularly prominent at its early-twenty-first century peak. Internet users of the era will remember the ubiquitous AOL CD-ROM mailer that would help computer users connect to the World Wide Web. AOL tokens crested at 11.62 fpm from 2000 to 2004, approaching—but not quite reaching—Microsoft levels of fame.

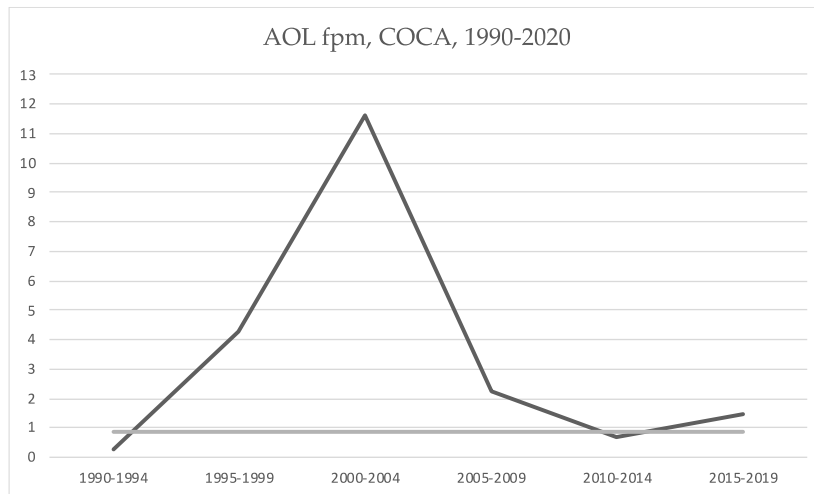


Figure 5. AOL, COCA, frequency over time, 1990–2019.

301. Schechter, *supra* note 41, at 829.

302. *Id.* at 830.

As we can see, however, AOL has recently dipped into the ambiguous end of the frequency pool, with an average of 0.68 fpm between 2010 and 2014. COCA allows us to take a closer look at AOL’s dip between 2010 and 2014, as it provides data by year. As shown below, AOL tokens appear at a rate above 1.0 from 2009 to 2011 but drop below the 0.87 benchmark (the solid gray line) from 2013 to 2015. AOL tokens spike in 2016 at 4.6 fpm. That spike may have something to do with Verizon’s acquisition of Yahoo and its proposed integration of AOL and Yahoo.³⁰³ AOL tokens creep back down to 0.8 fpm in 2019. Over the past decade, AOL appears to be flirting with frequencies that fall below our most forgiving proposed benchmark.

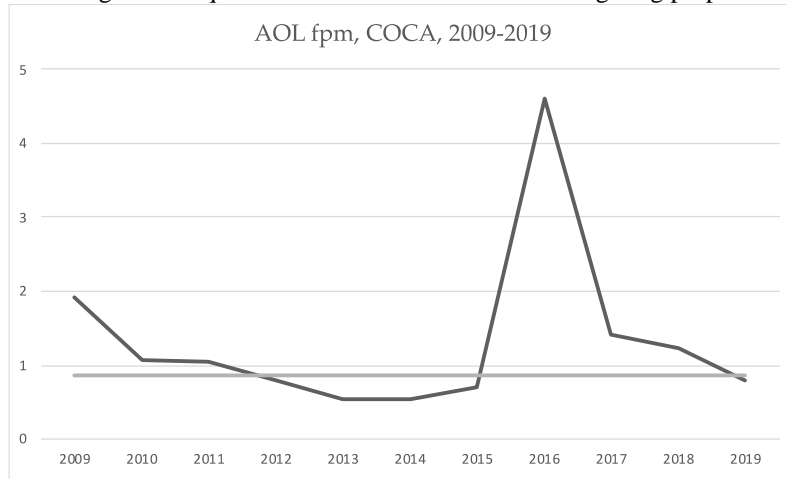


Figure 6. AOL, COCA, frequency over time, 2009–2019.

AOL, COCA	Avg	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
fpm	1.33	1.91	1.08	1.05	0.79	0.54	0.54	0.7	4.6	1.41	1.24	0.80

Table 7. AOL, COCA, frequency, 2009–2019.

2. Mama, Don’t Take My Kodak Fame Away!

Depending on how one views the evidence, Kodak’s fame may also have disappeared in an epochal flash. Kodak was once identified as the prototypical famous mark.³⁰⁴ Data from COHA supports that intuition, indicating that over its

303. Lara O’Reilly, *This Woman Has Been Given the Job of Merging Yahoo With AOL*, BUSINESS INSIDER, July 25, 2016, <https://perma.cc/AT52-PREM>.

304. See, e.g., *It’s a 10, Inc. v. Beauty Elite Grp., Inc.*, No. 13-60154-CIV, 2013 WL 6834804 (S.D. Fla. Dec. 23, 2013) (identifying Kodak as “a household name” and a “giant[] of branding”); *Fruit of the Loom, Inc., v. Girouard*, 994 F.2d 1359, 1362 (9th Cir. 1993) (identifying Kodak as among those marks that are sufficiently mature and well-known to qualify for anti-dilution protection); Monica Hof Wallace, *Using the Past To Predict the Future: Refocusing the Analysis of a Federal Dilution Claim*, 73 U.

life, Kodak tokens appeared with a frequency of 1.30 per million, strong evidence of prominence to accompany Kodak's singular use. Similarly, COCA reports an average of 2.45 fpm over the last thirty years.

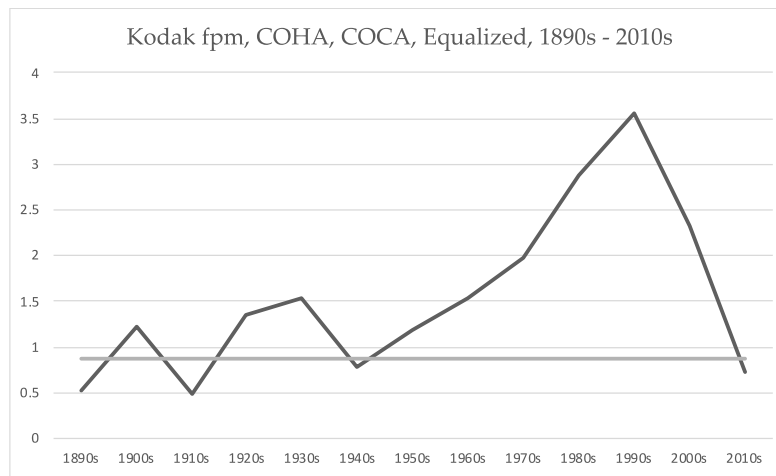


Figure 7. Kodak Frequency Report, COCA, COHA, 1890s–2010s.

CINCINNATI L. REV. 945, 947–48 (2005); H.R. REP. NO. 104-374 (1995) (giving as an example of dilution of famous marks the unauthorized use of Kodak on pianos). *See also supra* note 44 and accompanying text.

In the last decade, however, we see a significant drop in Kodak's frequency. From 2010 to 2014, the average frequency was 0.89, and only 0.55 from 2015 to 2019. Perhaps this should not surprise. In the past decade, Kodak was delisted from the S&P 500. The Eastman Kodak Company sought bankruptcy protection and shifted its business away from cameras and film toward chemicals.³⁰⁵ While Kodak hasn't lost its trademark significance, data from the past half-decade might suggest statutory fame is slipping through the corporation's fingers.

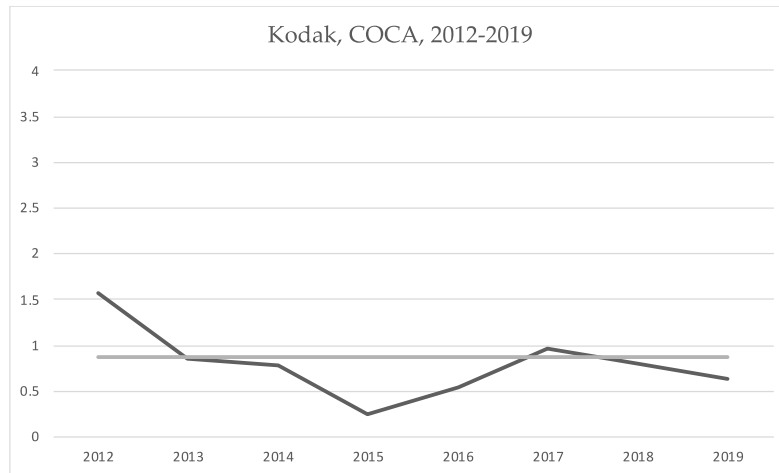


Figure 8. Kodak Frequency Report, COCA 2012-2019.

3. Fading Xerox?

The Xerox trademark is frequently invoked as another obviously famous trademark.³⁰⁶ As with other coined terms, the fanciful nature of the mark increases the perception.³⁰⁷ At a first glance, the frequency numbers bear out the mark's undeniable fame. Xerox first appears in COHA in the 1950s, and over the ensuing decades, averages a respectable 0.78 fpm, peaking in the 1990s at 3.33 fpm. COCA tells a similar tale of the '90s, with Xerox averaging 2.75 fpm across the decade.

305. Clare Duffy, *How Kodak Went From Photography Pioneer To Pharmaceutical Producer*, CNN BUS. (Aug. 4, 2020), <https://perma.cc/X5TJ-6H8U>.

306. See, e.g., *Savin Corp. v. Rayne*, No. 00-CV-11728 PBS, 2001 WL 34815751, at *8 (D. Mass. Mar. 26, 2001) (assessing a cybersquatting claim by noting the accused cybersquatter acquired domain names that included "indisputably famous trademarks such as Xerox"); Butters, *supra* note 132, at 514 (categorizing Xerox as "extraordinarily famous"); Miles J. Alexander & Michael K. Heilbronner, *Dilution Under Section 43(c) of the Lanham Act*, L. & CONTEMP. PROBS. 93, 106 (Spring 1996) (including Xerox among a list of "undeniably famous marks").

307. See, e.g., David M. Klein & Daniel C. Glazer, *Reconsidering Initial Interest Confusion on the Internet*, 93 TRADEMARK REP. 1035, 1047 (2003).

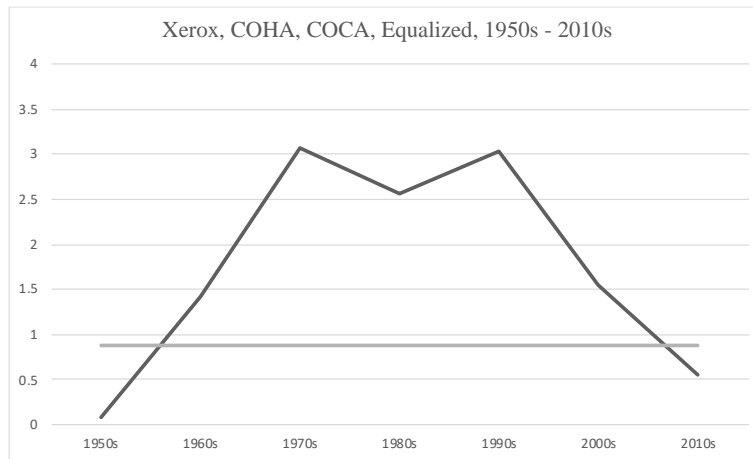


Figure 9. Xerox Frequency Report, COCA, COHA, 1950s–2010s.

But in the past fifteen years, Xerox’s frequency on COCA falls off, appearing at 0.63 fpm in the last half of the 2010s. Xerox tokens in the past eight years have fallen below our proposed lenient prominence benchmark of 0.87 fpm, suggesting its fame may be dissipating, although no one would argue the mark has lost its distinctiveness or source significance.

Even that data has some complications, however. Nestled between lower frequency years in 2011 (0.82 fpm) and 2013 (0.82 fpm), COCA’s 2012 sources were jammed with mentions of Xerox product (11.84 fpm). One question litigants should raise is whether the recent drops in frequency are aberrations, whether the spike in frequency is the deceptive outlier, or whether courts observing the noise in the data should shy away from the question of lost fame entirely.

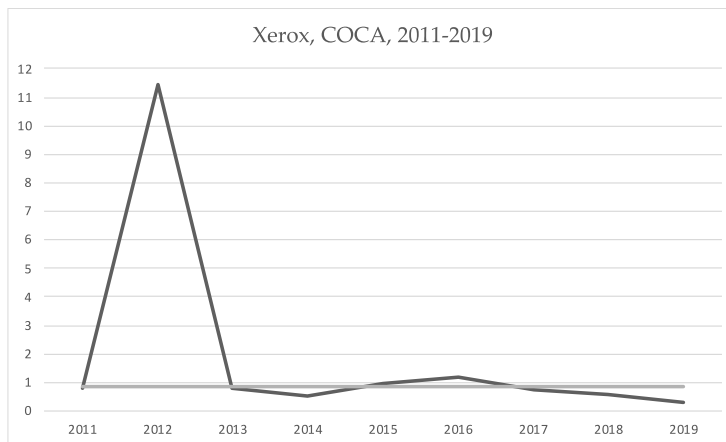


Figure 10. Xerox Frequency Report, COCA, 2011–2019.

4. Starbucks: Reaching for the Stars

Diachronic evidence may also shed light on when a mark acquires fame—if it does at all. Consider Starbucks, for which there was no evidence of use in COCA prior to 1993. The corpus reveals a steady growth of Starbucks tokens throughout the end of the '90s. Starbucks tokens appeared in 1999 with a frequency of 2.57 per million. That is well into the band of frequencies that provide likely evidence of prominence.

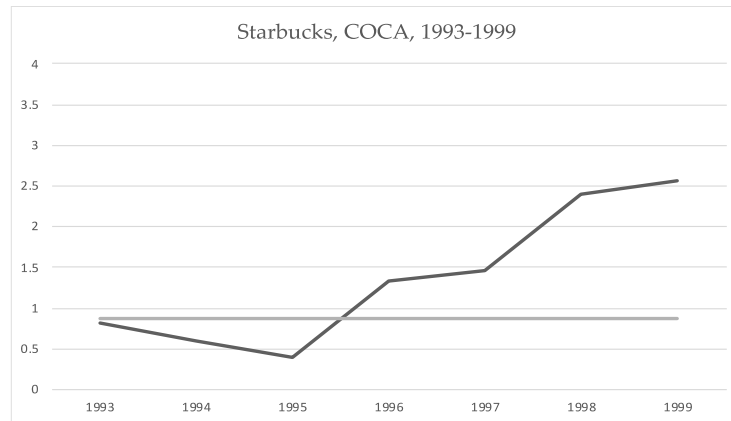
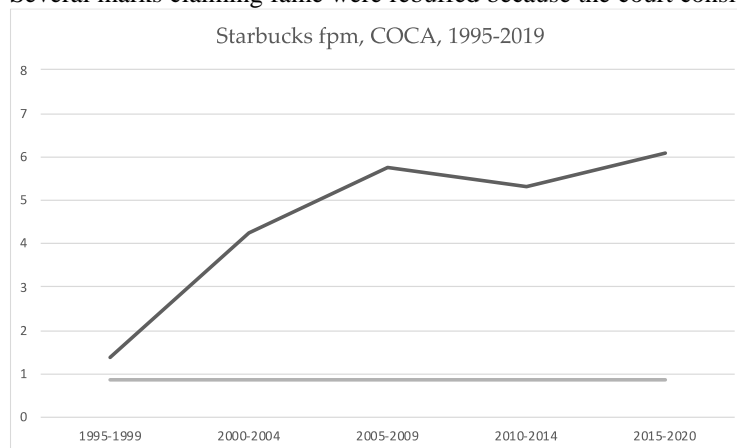


Figure 11. Starbucks, COCA 1993–1999.

Courts applying the FTDA frequently state that fame is acquired over time.³⁰⁸ Several marks claiming fame were rebuffed because the court considered their time



308. See, e.g., *WAWA, Inc. v. Haaf*, Civ. A. No. 96-4313, 1996 WL 460083 (E.D. Pa. Aug. 7, 1996) (*Wawa*, used for almost ninety years, is famous.), *aff'd*, 116 F.3d 471 (3d Cir. 1997); *Pirelli Armstrong*

in the commercial sphere (three years, nine years, fifteen years) too short.³⁰⁹ Perhaps a court should ask a mark like Starbucks to percolate a bit longer before concluding the mark has acquired fame, but one now sees a strong trendline for Starbucks.

Figure 12. Starbucks Frequency Over Time, COCA, 1995–2020.

III. APPLICATIONS, IMPLICATIONS, AND COMPLICATIONS

This final Part discusses how a court might use corpus analysis to dispose of some fame inquiries at an early stage in dilution litigation. We then discuss some of the advantages of corpus analysis over other tools for assessing fame. We also address some concerns about the use of corpus evidence in dilution cases.

A. APPLICATIONS

Meritorious famous marks should not be denied anti-dilution protection, but neither should a mark holder threaten unrelated markets by invoking a weak case for fame. Sometimes, a court declines to decide the fame issue on a motion to dismiss or a motion for summary judgment in a case where corpus evidence would support the defendant's motion to set aside a dilution claim for failure to establish fame, as in the Choose Your Own Adventure and Prima examples discussed below.³¹⁰

We do not argue that the corpus evidence is strong enough to dispose of the fame issue in every case. Nonetheless, in close cases, corpus evidence may provide additional useful insight and fortify an uncertain court in disposing of the fame question on a relatively early motion. Corpus evidence of the marks at issue in *Chooseco* and *Gerwan Farming* favor a finding that neither mark was famous.

1. Choose Your Own Fame Inquiry

Chooseco sued Netflix for, inter alia, dilution by tarnishment of its ostensibly famous mark, Choose Your Own Adventure (CYOA) for adventure books for young

Tire Corp. v. Titan Tire Corp., 4 F. Supp.2d 794, 802 (C.D. Ill. 1998) (Pirelli and Armstrong, registered trademarks for more than 80 years, are famous); Toys “R” Us, Inc. v. Akkaoui, No. C 96-3381 CW, 1996 WL 772709 (N.D. Cal. Oct. 29, 1996) (Toys “R” Us, used for thirty-six years, is famous). See also Edward E. Vassallo & Maryanne Dickey, *Protection in the United States for Famous Marks: The Federal Trademark Dilution Act Revisited*, 9 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 503, 511 (1999).

309. See, e.g., Michael Caruso & Co., Inc., v. Estefan Enters., Inc., 994 F. Supp. 1454, 1463 (S.D. Fla. 1998), *aff'd*, 166 F.3d 353 (11th Cir. 1998) (finding that use of mark for only fifteen years contradicts the contention that the mark is uniquely famous); Genovese Drug Stores, Inc. v. TGC Stores, Inc., 939 F. Supp. 340, 350 (D.N.J. 1996) (“We’ll Take Good Care of You,” used for nine years, is not famous.); Appleseed Found., Inc. v. Appleseed Inst., Inc., 981 F. Supp. 672, 677 (D.D.C. 1997) (Appleseed, used for three years, is not famous.).

310. See, e.g., William McGeveran, *Life in the Fast Lane: Of Presumptions, Defenses, and Burdens*, 1 IP THEORY 25, 28 (2010), for a discussion of when we might prefer courts to make presumptions in favor of defendants or weaken presumptions in favor of plaintiffs in trademark cases.

adults.³¹¹ A CYOA book is written from a second-person point of view, allowing the reader to make choices that lead through the story in a nonlinear fashion to multiple potential endings.

Netflix allegedly tarnished the mark by using it in a commercial context in Netflix's *Black Mirror: Bandersnatch* film. Netflix's use allegedly associated the mark with *Bandersnatch*'s "dark and disturbing content."³¹² The court rejected Netflix's motion to dismiss the tarnishment claim. The court did not opine on whether CYOA was famous, apparently assuming that Chooseco sufficiently pled fame. Although the parties recently settled the dispute,³¹³ the case provides an opportunity to test whether corpus analysis confirms or calls into question the court's presumption of fame.

An examination of the corpora suggests that CYOA is not famous. The first CYOA book, *The Cave of Time*, was published in 1979.³¹⁴ A COHA of the phrase "Choose Your Own Adventure" delivered no results. COCA results were a bit better but still in the Panavision range. Overall, CYOA appeared eighty-three times in COCA, a per-mil rate of 0.05. It spiked with fourteen hits from 2015 to 2019. That evidence suggests the court's presumption of fame was misguided.

Moreover, examining concordance lines suggests a relatively even mix of trademark and generic use. For example, the following string from *Salon* magazine in 2019 suggests the use of the phrase descriptively:

... a plot twist to that narrative. It's a **choose your own adventure** situation with paternity. On the one hand, there is with the advent of . . .

Other uses are expressive uses that appear to simultaneously reinforce the brand as the prototype for a category but also potentially engage in generic use:

... It's like a '**Choose Your Own Adventure** of an album,' says Ryan Holladay, citing the series . . .

All told, we concluded that thirty-nine instances of CYOA were clearly brand use. Another forty-three were at the border of expressive and generic uses, explaining how an item with choice architecture was like a CYOA story. This sort of evidence presents coding challenges. In some cases, expressive uses of a clearly famous mark to identify another seller's goods might be neither diluting nor generic when they refer to the other brand as "the Cadillac of" refrigerators or "the Choose Your Own

311. *Chooseco LLC v. Netflix, Inc.*, 439 F. Supp. 3d 308 (D. Vt. Feb. 11, 2020) (denying motion to dismiss, granting motion for leave to file sur-reply).

312. *Id.* at 27.

313. *Chooseco LLC v. Netflix, Inc.*, 2:19-cv-00008-wks (D. Vt. Nov. 30, 2020) (stipulation and order of dismissal with prejudice); Adi Robertson, *Netflix Settles Bandersnatch "Choose Your Own Adventure" Lawsuit*, VERGE, Nov. 25, 2020, <https://perma.cc/AP6F-6NRY>.

314. EDWARD PACKARD, *THE CAVE OF TIME, CHOOSE YOUR OWN ADVENTURE #1* (1979).

Adventure of” video games.³¹⁵ Thus, the evidence of singularity is somewhat mixed. The court might reasonably have concluded it should not presume CYOA is a famous mark, even on a motion to dismiss.

2. Prima Peaches

In one unreported case,³¹⁶ plaintiff Gerawan Farming, Inc., argued that its mark Prima for fresh produce (primarily stone fruits and grapes) was famous, and brought a dilution claim against defendant Prima Bella Produce, Inc. for its use of the Prima Bella mark for sweet corn. Gerawan offered minimal evidence that the Prima mark was truly famous, having spent “a relatively modest amount of money on product advertising.”³¹⁷ Prima lacks the prominence of well-known fruit brands like Chiquita for bananas, and Gerawan’s evidence was not overwhelming.³¹⁸ The court nonetheless denied Prima Bella’s motion for summary judgment. The court noted that the Prima mark was used “aggressively” on its packaging, paperwork and peripherals, and that some consumers asked for “Prima produce grown by the Gerawan name.”³¹⁹ Thus, “[a]lthough the Prima trademarks may lack the fame of Dole, Del Monte or Sunkist, Gerawan has raised factual issues as to the famous and distinctive element of trademark dilution.”³²⁰

Corpus evidence suggests the court overestimated the contestability of the factual issues raised. Instead, the court should have granted the Prima Bella’s motion for summary judgment on the dilution claim because Gerawan’s mark Prima is not famous. In a frequency search, the word *prima* appears in COCA at 1.69 fpm, cresting at 1.95 fpm in 2000–2004. To constitute strong evidence of prominence at that frequency, Gerawan’s use would need to account for roughly half of the tokens of *prima*. But a collocation search found no incidents of *prima* collocated with fruit, peaches, or grapes.

We also pulled concordance lines from the seven years leading up to defendant’s Prima Bella launch. We found 360 tokens of *prima* over those five years (mean = 51.43, median = 40), but none plausibly referred to Prima for fruits or produce of any

315. See Ann Bartow, *Likelihood of Confusion*, 41 SAN DIEGO L. REV. 721, 803–04 (2004) (discussing this phenomenon and its analysis by the Ninth Circuit in the “Barbie Girl” case, *Mattel, Inc. v. MCA Records, Inc.*, 296 F.3d 894 (9th Cir. 2002)); Ullrich, *supra* note 124, at 1010 (characterizing this as metaphoric and not genericizing use); Lee Goldman, *Proving Dilution*, 58 U. MIAMI L. REV. 569, 602 (2004) (classifying *x* is the *y* of *z* use as expressive and non-dilutive). But see Steven Y. Reeves, *Speech-Zilla Meets Trademark Kong?: How the Hollywood Circuit Got It Wrong in the Barbie Battle*, *Mattel, Inc. v. MCA Records, Inc.*, 4 MINN. INTELL. PROP. REV. 285, 337 (2003) (classifying the structure as “[d]ilutive blurring [that] involves relatively little expressive commentary”).

316. *Gerawan Farming, Inc. v. Prima Bella Produce, Inc.*, No. CV F 10-0148 LJO JLT, 2011 WL 3348056, at *1 (E.D. Cal. Aug. 2, 2011).

317. *Id.* at *27.

318. *Id.*

317. *Id.*

320. *Id.* (listing Gerawan’s uses of the Prima trademarks in connection with its products, such as on its “boxes, buildings, website, labels, tags, recipe cards, pallets, clothing, headwear, pens, beverage containers, computer bags, paper tablet holder” and so on as a basis of factual dispute).

kind. This evidence suggests that Gerawan's Prima mark was no longer truly prominent and renowned in the early 2000s, if it ever was. Moreover, the concordance analysis revealed third-party use for brands like Prima Energy, Prima Kase cheese, Prima Publishing, and Rolf Prima bicycle tires. That suggests *prima*, an Italian word invoking concepts such as primacy or prominence, evidences the crowding we might expect around a laudatory mark, the type of crowding that typically undermines a claim to fame.³²¹

3. Visualizing Singularity and Prominence

The following figure places Choose Your Own Adventure, Prima, and other aforementioned marks into a rough framework. The closer a mark lands to the upper right corner, the more confident a court or litigant can be in concluding that the mark has sufficient singularity and prominence to qualify as famous. A shift to the left suggests insufficient singularity, while a shift downward indicates lack of prominence.

Prominence / Singularity	Not Dominant	Commercially Dominant, but Crowded	Commercially Singular	Substantially Singular
<i>Prominent</i> $x \geq 5.0$	Columbia		Apple Ford	Microsoft Starbucks
<i>Likely Prominent</i> $5.0 < x \geq 1.00$	App Store	Delta	Blackberry	Porsche
<i>Strong Evidence</i> $1.00 < x \geq 0.87$			Winston ³²²	Budweiser
<i>Ambiguous Evidence</i> $0.87 < x \geq 0.00$	Trek	Camel	Big Gulp	Hotmail Panavision
<i>Unattested</i> ~ 0.00	Coach Prima	Choose Your Own Adventure		Arthur the Aardvark

Figure 13. Singularity and Prominence.

B. IMPLICATIONS

The corpus linguistic method provides an alternative to survey evidence. Surveys are often touted as exemplary evidence of consumer confusion. But surveys are also

321. See *supra* note 158 and accompanying text.

322. Winston, the cigarette brand, appears with high frequency from its former sponsorship of NASCAR's Winston Cup Series from 1971 to 2003.

expensive.³²³ While they are frequently relied on, litigants regularly challenge their reliability and probity.³²⁴ Moreover, a survey is essentially a snapshot of how consumers perceive the trademark at issue at the time the survey is conducted, not during a prior era.

Collocation, concordance, and frequency evidence may be significantly cheaper to generate than survey evidence if litigants can use a publicly accessible corpus like COCA, and equally probative as survey evidence for the reasons discussed above. Therefore, prior to discovery, it will be easier for a litigant to estimate the usage of an allegedly famous mark via corpus analysis than to uncover evidence of market penetration that an opposing party might keep close to the vest.

Corpus data may also reduce judicial overreliance on dictionaries. In their collective role of parsing constitutional and statutory text,³²⁵ courts have become enamored with dictionaries as the best tool to help establish the ordinary meaning of the words to be understood.³²⁶ Dictionaries can show multiple meanings of a given word, but dictionaries are generally not organized in a manner to clearly designate the ordinary or prototypical meaning that should dominate textualist analysis.³²⁷ Courts make a similar error relying too heavily on dictionaries to assess *ex ante* genericness when investigating trademark validity.³²⁸ Corpus evidence provides a different view of the term in question as used in various contexts. In the fame context, dictionaries generally afford space to only a fraction of potentially famous

323. See Linford, *Democratizing Access*, *supra* note 62, at 6–8 (“A survey is useful, but gathering survey evidence of distinctiveness can be prohibitively expensive.”).

324. See *generally id.* at 9–10 (explaining courts’ role as a gatekeeper of surveys as admissible evidence in trademark cases).

325. Lee & Mouritsen, *supra* note 14, at 792 (“[T]he threshold question for the ‘standard picture’ of legal interpretation . . . starts with a search for the ‘ordinary communicative content’ of the words of the law.” (quoting William Baude & Stephen E. Sachs, *The Law of Interpretation*, 130 HARV. L. REV. 1079, 1106 (2017))).

326. Mark A. Lemley, *Chief Justice Webster*, 106 IOWA L. REV. 299, 304–5 (2020).

327. Craig Hoffman, *Parse the Sentence First: Curbing the Urge to Resort to the Dictionary When Interpreting Legal Texts*, 6 N.Y.U. J. LEGIS. & PUB. POL’Y 401, 401 (2003) (“[J]ust as medical science has progressed since the time of leech treatments, the science of linguistics has progressed since the time that scholars believed that dictionaries held the key to sentence meaning.”); Brief of Trademark and Internet Law Professors in Support of Respondent, at 5–6, 19–20, 28, U.S. Patent and Trademark Office v. Booking.com B.V., No. 19-46 (4th Cir. Feb. 19, 2020) (explaining that dictionaries over-rely on a term’s history rather than focusing on how it is actually used by sellers and consumers in commerce); Clark D. Cunningham, Judith N. Levi, Georgia M. Green & Jeffrey P. Kaplan, *Plain Meaning and Hard Cases*, 103 YALE L.J. 1561, 1614–16 (1994) (“[T]here is no single reference book which is *The Dictionary*, but rather a number of competing publications which themselves may differ significantly from edition to edition.”); Mouritsen, *Fortress*, *supra* note 149, at 1935–37, 1945–46 (arguing that at best, a dictionary can provide “the most frequently *encountered* meaning,” not “the most frequently *occurring* meaning”); cf. Lawrence B. Solum, *Triangulating Public Meaning: Corpus Linguistics, Immersion, and the Constitutional Record*, 2017 BYU L. REV. 1621, 1631 (“Lexicographers report these conventional semantic meanings in dictionary definitions, but such definition are secondary evidence of patterns of usage.”); Linford, *Datamining*, *supra* note 149, at 1553 (“Corpus lexicography done well should provide better evidence of original meaning than an appeal to a dictionary alone.”).

328. See *generally* Linford, *Linguistic*, *supra* note 13.

brands.³²⁹ The inclusion of a trademark in a dictionary certainly hints at prominence—as the number of entries for the word used as a mark can fortify or undermine singularity—but a dictionary is not optimized for investigating those questions.

Third, as highlighted above in Part II.D, corpus analysis can be particularly well-suited to show language change over time if the corpus collects from sources across eras. As Bedi and Schuster note, “it seems appropriate that a court would prefer more than just a single snapshot of [a] fame measurement to establish a continued pattern of fame, as well as to establish exactly when a mark became famous.”³³⁰ Corpus evidence of changing meaning may establish a point at which a mark has become famous.

For example, Coach Services, Inc.—maker of the well-known Coach handbags—sought to establish that the mark was famous and that the use of the Coach mark on education services would likely dilute its mark.³³¹ Unfortunately for Coach Services, the evidence did not establish to the court’s satisfaction that the mark was famous prior to Triumph Learning’s applications on December 21, 2004, to register Coach-related marks for educational services.³³²

A collocation search in COCA for *coach* within four words of *bag* indicates twelve potential matches between 1993 and early 2004, nine of which refer back to the leather goods firm. A similar search of *coach* within four of *leather* yields another six hits from 1999 to early 2004. A look generally for *coach* in the same eras shows 8,441 tokens between 1990 and 1994, 11,319 tokens between 1995 and 1999, and 12,260 tokens between 2000 and 2004. This evidence supports the court’s conclusion that Coach Services failed to establish fame. Fifteen trademark-related tokens over a fifteen-year window is a negligible fraction (0.0005) of the total tokens of *coach* in COCA over the same period (n=32,020; 73.86 fpm). Multiplying that fraction by the fpm of all instances of *coach* would yield 0.03 efp, far below any reasonable threshold for fame.

Compare uses of *coach* with evidence of Microsoft’s prominence. *Microsoft* appears over twenty-four times per million words. Collocation results indicate that its use singularly points to the mark in its commercial context and concordance lines overwhelming refer to Microsoft as a brand. This difference suggests the Coach mark for handbags does not approach Microsoft’s singularity, and the Coach mark’s estimated prominence barely clears the level of unattested use, which indicates Coach is not widely known among the general consuming public.

329. Butters, *supra* note 132, at 510–11 (explaining the lack of correlation between terms listed (and not listed) in the dictionary and the level of fame for such terms).

330. Cf. Bedi & Schuster, *supra* note 15, at 479–80 (discussing the temporal snapshot nature of survey evidence and recommending that companies measure consumer perception at regular intervals to preserve evidence for dilution causes of action); Bedi & Reibstein, *supra* note 15, at 723–24 (same).

331. *Coach Servs., Inc. v. Triumph Learning LLC*, 668 F.3d 1356, 1371–72 (Fed. Cir. 2012).

332. *Id.*; see *COACH AMERICA’S BEST FOR STUDENT SUCCESS*, Reg. No. 4,219,848 (registered Oct. 9, 2012) (canceled May 10, 2019) (word and design *COACH AMERICA’S BEST FOR STUDENT SUCCESS* computer programs and printed materials for use in child and adult education).

C. COMPLICATIONS

Recent applications of corpus linguistics to legal analysis promise to increase the rigor of methodologies like originalism and textualism.³³³ However, other scholars caution that judicial reliance on corpus-based research “risks producing misguided judicial outcomes that will prove resistant to review.”³³⁴ Some of the concerns about judicial reliance on corpora when engaging in constitutional or statutory interpretation are likely less problematic, or at least less acute, in the context of trademark interpretation because the use of data-driven analysis does not deviate from historical precedent. Scholars have used corpus analysis to analyze whether sellers are running out of good trademarks,³³⁵ whether and how trademark dilution occurs,³³⁶ and how to supplement the standard primary significance test with corpus evidence in genericness disputes.³³⁷ Moreover, trademark litigants and courts have made use of data-driven interpretive tools, including consumer surveys, for decades.³³⁸ Considering corpus evidence does not constitute a significant departure from that practice.³³⁹ Indeed, in certain circumstances, corpus and survey evidence can be used to effectively triangulate better results than using either alone.³⁴⁰

From our perspective, corpus evidence cannot singlehandedly dispose of the fame question, but provides a relatively quick look at prominence and singularity at negligible cost. Moreover, corpus evidence is limited because while it can

333. 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, 1181 (2017) (arguing that “originalists who embrace a Big Data transformation will be able to reliably and accurately reveal original language conventions”).

334. John S. Ehrett, *Against Corpus Linguistics*, 108 GEO. L.J. ONLINE 50, 61 (2019) (“[A]t best corpus-based research may constitute a time-consuming diversion; at worst, it risks producing misguided judicial outcomes that will prove resistant to review.”); see generally Carissa Byrne Hessick, *Corpus Linguistics and the Criminal Law*, 2017 B.Y.U. L. REV. 1503, 1505 (2017) (arguing that corpus linguistics represents a radical break from current interpretive theories and ought not be adopted as an interpretive theory for criminal laws, because corpus linguistics is unlikely to deliver on its promise to increase predictability and decrease the use of the judge’s personal preferences, while it may sacrifice important values such as notice and accountability); Tammy Gales & Lawrence Solan, *Revisiting a Classic Problem in Statutory Interpretation: Is a Minister a Laborer?*, 36 GA. ST. U. L. REV. 491, 500 (2020) (discussing Justice Breyer’s use of corpus analysis in *Muscarello v. United States*, 524 U.S. 125 (1998)); see also Anya Bernstein, *Legal Corpus Linguistics and the Half-Empirical Attitude*, 106 CORNELL L. REV. 1397, 1406–07 (2021) (explaining the ambiguous role of corpus linguistics in statutory interpretation).

335. See generally Beebe & Fromer, *supra* note 124, at 947 (analyzing corpus data and concluding that trademark applicants are increasingly being forced to resort to second-best, less competitively effective marks).

336. Heald & Brauneis, *supra* note 124, at 2574–75.

337. Hoopes, *supra* note 124, at 410–11 (evaluating the utility of corpus linguistics in genericide disputes); Ullrich, *supra* note 124, at 992–93 (proposing using corpora to measure primary significance in trademark disputes).

338. Cf. Linford, *Democratizing Access*, *supra* note 62, at 1 (“Consumer surveys provide evidence of distinctiveness on which courts frequently rely.”).

339. See generally *id.*

340. See, e.g., Kyra Nelson, *Informing Lexicographic Choices through Corpus and Perceptual Data*, 33 INT. J. LEXICOGRAPHY 251, 251 (2020) (integrating corpus data and perceptual survey data to examine “trends between native speaker perceptions of lexical items and their actual use”).

demonstrate the absence of a word used in certain contexts in a corpus, it cannot disclose the use of the word in other contexts. The absence of a word in COCA or COHA is not dispositive of trademark use in every context. But with corpora of similar size, infrequent appearance or no appearance in trademark context may indicate that the mark has not acquired sufficient prominence or singularity to qualify as famous.³⁴¹

As we see with attempts of brand owners to direct trademark usage in dictionaries and other media, it is fair to ask whether corpora may be subject to attempted manipulation. Publishers get angry letters from mark owners when a mark is used in generic fashion.³⁴² Lexicographers must also fend off pressure from unhappy mark owners when a dictionary entry indicates a mark is used generically.³⁴³ We have not encountered anecdotal evidence of such attempted corpus manipulation. Nonetheless, as with any other system that courts or litigants rely on to establish questions of law or fact, mark owners could seek to exert pressure on corpus compilers or game corpus analysis. On the other hand, the nature of corpus gathering may limit the effect of such strategies. A mark owner might arguably bring pressure to bear on the host of a corpus, but it will prove more difficult to bring pressure to bear equally on all the various sources whose uses are included in a wide-ranging corpus.

Finally, focusing on singularity and uniqueness will make it comparatively easier for the owner of a coined or fanciful mark like Xerox or Kodak—or perhaps a new portmanteau like Microsoft or Chik-Fil-A—to assert fame because of the monosemous origin of the term. Coca-Cola is clearly a famous mark by any measure, but a test that looks for high levels of monosemy might disfavor other marks built from relatively common or descriptive elements. This concern, however, is not unique to corpus analysis, but raises the question of how we should conceive of fame. We have argued the prominence and singularity benchmarks are consistent with the TDRA and current case law, and why they are appropriate given the scope of protection offered to marks that qualify for antidilution protection.

Finally, we note again that many scholars are skeptical of the anti-dilution enterprise. We take no position on that debate but expect that anti-dilution law is

341. See Lawrence M. Solan, *Corpus Linguistics as a Method of Legal Interpretation: Some Progress, Some Questions*, 33 INT. J. SEMIOTICS L. 283, 290 (2020) (arguing that the absence of a word in a corpus can be meaningful “from the fact that the missing concept is expressed in other [words]”). But see Anaol Stefanowitsch, *Negative Evidence and the Raw Frequency Fallacy*, 2 CORPUS LINGUISTICS & LINGUISTIC THEORY 61, 62 (2006) (noting that corpus analysis can identify the absence of a word but not why it is absent).

342. Frank Ahrens, *So Google Is No Brand X, but What Is “Genericide”?*, WASH. POST (Aug. 5, 2006), <https://perma.cc/RS7S-4K3Z>.

343. SIDNEY I. LANDAU, *DICTIONARIES: THE ART AND CRAFT OF LEXICOGRAPHY* 280 (1st ed. 1984) (“The dictionary editor must do battle to include any trademarks, and he is under great pressure to distort the facts of usage by entering all such terms only in capitalized form, even though the record clearly shows they are often written in lower-case letters.”); Butters, *supra* note 132, at 509–10 (describing how dictionaries may skew for utility reasons toward more obscure usage).

here to stay. Refining the fame inquiry may serve to cabin the costs that some ascribe to rights against dilution.

IV. CONCLUSION

The Trademark Dilution Revision Act sets a high standard for mark owners hoping to cleanse the commercial lexicon of blurring and tarnishing uses. Only famous marks need apply. If one takes the fame requirement seriously, it seems unquestionable that some marks have successfully enforced anti-dilution rights without truly becoming “widely recognized by the general consuming public.”

Our approach asks courts to inquire both whether the mark in question is substantially singular in its use and whether it is truly prominent and renowned. Applying our dual fame benchmarks—singularity and prominence—can help a court focus its inquiry on the demanding fame requirement and ensure the powerful rights against blurring and tarnishment are not extended to marks that do not qualify. Courts that accept the invitation will likely grant anti-dilution rights to fewer marks than have historically qualified, as our discussion in Part II highlights, but this outcome is consistent with the text and history of the TDRA.

More importantly, corpus analysis tools can provide potential litigants with relatively affordable and readily accessible evidence of whether a purportedly famous mark has reached necessary levels of prominence and singular use. Corpus evidence can likewise provide courts and fact finders with better insight into mark use over time and consumer use in context, compared to existing practices.

We propose singularity standards and prominence benchmarks sufficiently stringent to restrict dilution suits to a handful of litigants. The majority of mark owners should instead pursue remedies against likely confusion, unfair competition, and false advertising, if at all. But the use of corpus tools doesn’t require embracing a particular benchmark. Indeed, we suspect courts and litigants will find reasons to resist any benchmark we offer. Nonetheless, these tools provide an opportunity to shore up or undermine evidence of fame that until now has been drawn from unsolicited third party mentions, consumer surveys, and the other bits and bobs that litigants lay before the bench and jury box, hoping to sway the court about a trademark’s fame. Better evidence from corpus analysis will help courts and litigants reach earlier and more accurate resolution of dilution claims.