

Legal Communication & Rhetoric: JALWD

Fall 2012 / Volume 9

.....

ARTICLES & ESSAYS

**A Picture is Worth a Thousand Words:
How Wordle™ Can Help Legal Writers**

Allison D. Martin

A Picture Is Worth a Thousand Words: How Wordle™ Can Help Legal Writers

Allison D. Martin*

I. Introduction

Wordle™¹ is a visualization tool that can help legal writers identify themes in their writing. The tool, which is available for free online, generates picture collages, called “word clouds,”² from text. The more frequently a word is used in that text, the larger that word will appear in the word cloud. To use it, a writer simply pastes any text in the program, reviews the resulting word cloud, which appears almost immediately, and confirms that the larger words match the writer’s theme. If a match exists, the writer likely achieved the desired theme. If a match does not exist, the writer can “see” that the draft may need revision; further, the word cloud provides the writer with visual clues about which words, and therefore ideas, to downplay when revising the draft. Thus, Wordle also can be used to supplement idea generation methods—all the while adding a bit of fun and levity to the sometimes dull stages of legal work.

.

* ©Clinical Professor of Law, Indiana University Robert H. McKinney School of Law. I thank Professor Ken Chestek, Clinical Professor of Law, Indiana University Robert H. McKinney School of Law, for creating such a great study to demonstrate Wordle’s usefulness and for his encouragement and helpful comments. I also am grateful to Susan Smith Bakhshian, Director of Bar Programs and Academic Success, and Clinical Professor of Law, Loyola Law School Los Angeles, for sharing her insights about Wordle’s usefulness generally and about this essay in particular. Finally, I am thankful for comments shared by other colleagues and for research conducted by librarians Miriam Murphy and Steven Miller.

¹ Jonathan Feinberg, *Wordle, Home*, www.wordle.net (2009) [hereinafter Feinberg, *Wordle*]. Jonathan Feinberg “created the Wordle word-cloud layout algorithms while working on a social bookmarking application at IBM Research, in 2005. [He] created the ‘Wordle’ web application in 2008.” Jonathan Feinberg, *mrfeinberg.com* (accessed Oct. 17, 2011). For a more scientific description of Wordle, see Jonathan Feinberg, *Wordle*, in *Beautiful Visualization: Looking at Data through the Eyes of Experts*, 37–58 (Julie Steele & Noah Iliinsky eds., O’Reilly Media 2010) [hereinafter Feinberg, *Wordle*, in *Beautiful Visualization*].

² Feinberg, *Wordle*, *supra* n. 1.

This sidebar essay will use Wordle to review three briefs and a judicial opinion filed in recent healthcare litigation, discussed by Professor Ken Chestek in his article, “Competing Stories: A Case Study of the Role of Narrative Reasoning in Judicial Decisions.”³ To demonstrate just how Wordle can help practitioners and judges identify themes in their writing, I will begin with a brief description of what Wordle is and how it works. I will then examine several word clouds generated by the program using those documents and consider whether the themes identified by Wordle are consistent with the “competing stories” Professor Chestek observed in those briefs. Finally, I will discuss Wordle’s limitations.

II. What is Wordle?

Text visualization is a powerful, emerging technique used in a wide variety of contexts to facilitate analysis.⁴ Wordle is one software program that creates text visualizations; it “generat[es] ‘word clouds’ from text that you provide.”⁵ It is one of several word cloud software services available online.⁶ It is free and easy to use. One simply copies text from any document and pastes it into the program. The program then quickly and automatically generates a word cloud based on the original text. The more frequently a word appears in the original text, the larger that word appears in the cloud.⁷

3 See generally Kenneth D. Chestek, *Competing Stories: A Case Study of the Role of Narrative Reasoning in Judicial Decisions*, 9 Leg. Commun. & Rhetoric: JALWD __ (2012).

4 Audrey A. Pureskiy, Gregory L. Shutt & Michael W. Berry, *Survey of Text Visualization Techniques in Text Mining: Applications and Theory*, 107 (Michael W. Berry & Jacob Kogan eds., John Wiley & Sons, Ltd. 2010).

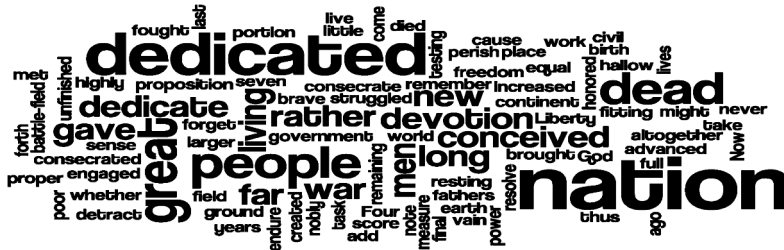
5 Feinberg, *Wordle*, *supra* n. 1.

6 One other is “TagCrowd.” Daniel Steinbock, *TagCrowd*, www.tagcrowd.com (accessed Oct. 17, 2011) and ToCloud, *ToCloud*, www.tocloud.com (accessed Oct. 17, 2011). I make no comment on these other software programs. I chose Wordle because I was introduced to it first, during an excellent presentation on an unrelated topic by Mark Urtel, EDD MS, Associate Professor, Department of Physical Education, Indiana University School of Physical Education & Tourism Management. I have since learned that Wordle has more graphical capabilities, such as the ability to manipulate text, background color, and orientation of words, than at least some of the other programs. See Pureskiy, Shutt & Berry, *supra* n. 4, at 109–10. I caution, though, that this technology is changing rapidly.

7 Wordle identifies and counts words in text. To count as the same word, words must be spelled and capitalized exactly the same way. The program treats “Commerce” and “commerce” as two different words, for example. Further, the program does not identify “stems” of words; thus, “state,” “states,” and “stating” would be considered different words. See Feinberg, *Wordle*, *FAQ*, *supra* n.1. Two-word phrases can be visualized as one word by using “the tilde character ~ between words that go together” *Id.* The default setting removes all numbers and common words, such as “and” and “the.” *Id.* Of course, the user also can omit any words from the original text. In addition, the program allows the user to change the font, color, and layout of the word cloud. See Feinberg, *Wordle*, *Home*, *supra* n. 1. The placement of words in a cloud is random, but Wordle offers some limited, more advanced options for manipulating placement. See Feinberg, *Wordle*, in *Beautiful Visualization*, *supra* n. 1, at 51.

Once a word cloud has been created, the program allows the user to take a screen shot of it, to print it, or to save it to the public gallery. According to the website, Wordle is safe to use on confidential text. See Feinberg, *Wordle*, *FAQ*, *supra* n. 1. A word cloud can become public only if the user saves it to the public gallery, and even then “only the word frequencies for the words that appear in Wordle are sent. There’s no way to reconstruct the source text from that information.” *Id.*

Applying Wordle to Lincoln's Gettysburg Address, a nonlegal text, demonstrates visually the prevalence of certain words that by their repetition Lincoln stressed.⁸ When creating the Address, Lincoln drew "on a classical rhetoric befitting the democratic burial of soldiers, on a romantic nature-imagery of birth and rebirth expected at the dedication of rural cemeteries, on biblical vocabulary for a chosen nation's consecration and suffering and resurrection, [and] on a 'culture of death' that made mourning serve life."⁹ The basic elements in the Address are "life and death."¹⁰ In the Address, "Lincoln tells us that the dead '*gave* their lives,' they did not simply lose them, and they did so for a single purpose, 'that that nation might live.'"¹¹ Lincoln "looks . . . to the birth of a *nation's* life . . . , its testing ordeal-by-death, and its new birth of freedom."¹² "When [Lincoln] spoke at the end of the Address, about government '*of the people, by the people, for the people*,' . . . he was saying that America is *a* people addressing its great assignment as that was accepted in the Declaration,"¹³ that "all men are created equal."¹⁴ Lincoln's Gettysburg Address formed the following word cloud:



The largest words in the cloud are nation, dedicated, great, people, and dead. These words closely track the themes of the Gettysburg Address: the dedication of the cemetery to remember those people who gave their lives to begin a rebirth of a nation that recognizes equality of people. It also is interesting to see what Lincoln decided not to emphasize.

8 President Abraham Lincoln delivered this speech on November 19, 1863, during the Civil War, at the dedication of a memorial cemetery in Gettysburg, Pennsylvania. I copied and pasted into Wordle the version of the speech known as the “Bliss Copy,” which is believed by many to be the final version of the speech, from *Abraham Lincoln Online, Speeches and Writings*, <http://showcase.netins.net/web/creative/lincoln/speeches/gettysburg.htm> (accessed on Oct. 17, 2011).

9 Garry Wills, *Lincoln at Gettysburg: The Words That Remade America* 89 (Simon & Schuster 1992). Professor Wills won a Pulitzer Prize for General Non-Fiction in 1993 for this book. He is currently an emeritus professor in the history department at Northwestern University.

10 *Id.* at 62.

11 *Id.* at 59.

12 *Id.* at 77.

13 *Id.* at 145.

14 *Id.* at 120.

Noticeably absent in the word cloud is any direct reference to the Union, slavery, Gettysburg, or other particulars, showing that Lincoln chose not to directly address “the prickliest issues of its historic moment.”¹⁵ “The draining of particulars from the scene raises it to the ideality of a type.”¹⁶ “Lincoln was looking beyond the war to ‘the great task remaining before us’ as a nation trying to live up to the vision in which it was conceived.”¹⁷ My guess is that Lincoln would have been satisfied with the themes captured in his word cloud.

III. How Can Wordle Be Useful for Lawyers and Judges?

Wordle has been used as an analytic tool to examine writing styles, public speeches, and survey or focus-group results.¹⁸ I was interested in demonstrating how Wordle can be used as an analytic tool to examine the “big picture” or overall themes in briefs and judicial opinions. For my case study, I used briefs and an opinion from the healthcare litigation¹⁹ Professor Chestek discussed in his article. I then created word clouds from those documents and compared them to the themes identified by Professor Chestek. My goal was to determine whether those themes matched the resulting word cloud, demonstrating that Wordle actually captures or visualizes themes.

Professor Chestek categorized the plaintiffs in *Liberty University v. Geithner*²⁰ as “Private Individuals and Employers,” the protagonists in their

¹⁵ *Id.* at 90.

¹⁶ *Id.* at 54.

¹⁷ *Id.* at 145.

¹⁸ See Carmel McNaught & Paul Lam, *Using Wordle as a Supplementary Research Tool*, 15 Qualitative Rpt. 630, 631 (May 2010), <http://www.nova.edu/ssss/QR/QR15-3/mcnaught.pdf> (using Wordle to find common themes and compare differences between sets of responses from transcribed focus-group meetings and student blogs about educational experiences). In addition, the authors point to a literary study that was conducted using Wordle to compare Gertrude Stein’s writing style in *The Making of Americans* with the writing styles of novels written by Jane Austen, Charles Dickens, George Eliot, and George Meredith. *Id.* (citing Tanya Clement, Catherine Plaisant & Romain Vuillemot, *The Story of One: Humanity Scholarship with Visualization and Text Analysis*, (Tech Report HCIL-2008-33), College Park, MD: University of Maryland, Human-Computer Interaction Lab, <http://hcil.cs.umd.edu/trs/2008-33/2008-33.pdf> (accessed on January 18, 2012)). In addition, Jonathan Feinberg has examined presidential inaugural addresses. See *Beautiful Visualization*, *supra* n. 1 at 56. In the legal context, Wordle has been used by authors. See, e.g., David I. C. Thomson, *Law School 2.0: Legal Education for a Digital Age* (LexisNexis 2009) (using Wordle to summarize each chapter); Joseph S. Miller, *Hoisting Originality*, 31 Cardozo L. Rev. 451, 451 (2009) (using Wordle to summarize his working paper). For examples of how to use Wordle in teaching and learning, see Ramsden, *supra* n. 27, and Rhonda Huisman, Willie Miller & Jessica Trinoskey, *We’ve Wordled, Have You?: Digital Images in the Library Classroom*, 72 College & Research Libraries News 522 (October 2011).

¹⁹ I used the plaintiffs’ brief, the defendant’s brief, and the district court opinion from *Liberty U. v. Geithner*, 753 F. Supp. 2d 611 (W.D. Va. 2010). I also used the plaintiffs’ brief from *Florida ex rel. Bondi v. Dept. of Health and Human Servs.*, 780 F. Supp. 2d 1256 (N.D. Fla. 2011). For the briefs, I omitted from text all cover pages, tables of authority, and tables of content. For the briefs and the court opinion, I omitted these common legal terms and cites: Plaintiffs, Defendants, U.S., S. Ct., and *Id.*

²⁰ 753 F. Supp. 2d 611 (W.D. Va. 2010).

story. Their goal was to express their right not to buy health insurance, so their brief focused on “Congress’ assertion of broad powers under the Commerce Clause” while spending “[l]ittle space” on “why individual freedom is important.”²¹ The brief cast Congress as the obstacle or “villain” in their view of the case.²² In addition, two individual plaintiffs advanced arguments in the same brief that the healthcare-reform law would somehow facilitate abortions and that they should not be forced to participate in such a program because abortion was against their religious beliefs.²³ The collective plaintiffs’ brief formed the following word cloud:



The largest words in the word cloud are Act, Congress, religious, health, Second, and Clause. These words closely track the two main themes in this brief: the challenge based on the Commerce Clause and the challenge based on freedom of religion.²⁴ It also confirms that the plaintiffs spent a great deal of energy discussing the role of Congress and less time on individual freedom and states’ rights, which are much smaller words in the word cloud; the word “individuals” appears on the far right side, the word “state” appears above “Congress,” and the word “States” appears below “Congress.”

The defendant in the *Liberty University* case was the “United States Government.” As the protagonist in its story, the defendant placed “the people it protects (the Everyperson hero, all American citizens) into the center of the story,” making them co-protagonists.²⁵ Their goal was to make healthcare “more universally available, and at a lower cost.”²⁶ The obstacle is “a health care system that is badly broken.”²⁷ “The antagonists include the greedy insurance companies who seek to ‘exclude from coverage those they deem most likely to incur expenses.’”²⁸ “The solution

21 Chestek, *supra* n. 3 at 114 (p. 28.)

25 *Id.* at 118.

22 *See id.*

26 *Id.*

23 *Id.* at 113, n. 70.

27 *Id.* at 119.

24 *Id.* at 107.

28 *Id.*

to the problem . . . is to require insurance companies to cover everybody”²⁹ This brief formed the following word cloud:



The largest words in this cloud are Health, Insurance, Coverage, U.S., provision, minimum, Congress, and clause. This word cloud is consistent with the theme Professor Chestek identified, especially the terms “Insurance” and “Coverage.”

Given the themes from both sides, it is now interesting to test the trial court’s opinion. Does the word cloud depict the winning side’s theme more prominently than the losing side’s theme? In this case, the United States Government won at the trial court level. Here’s the word cloud formed from the judicial opinion:



The largest words in the word cloud are Act, health, coverage, religious, Congress, and insurance. Further, the word “religious” is secondary in size to “health” and “coverage.” Thus, the theme of the court’s opinion, as visualized here, is consistent with the United States

Government's—the winning side's—theme. In other words, the word cloud shows that the court was persuaded more by the Government's arguments than by those of the Private Individuals and Employers.

It also is interesting to examine what the court decided not to emphasize. The words “state” and “States” in the right corner of the cloud appear relatively smaller, which supports Professor Chestek’s point that the Private Individuals and Employers did not succeed at persuading the court to focus on states’ rights, perhaps because these plaintiffs lacked the credibility to tell that story.³⁰

These plaintiffs were successful, however, in capturing the court's attention on the religious theme, as demonstrated by the size of "religious" in the word cloud. As Professor Chestek pointed out, though, because the Act provides an exception for religious objectors and requires that at least one health plan not provide coverage for abortion services, this argument was not likely to prevail.³¹ Thus, the fact that the court focused more of its attention on this theme was likely not advantageous to the plaintiffs. Again, the plaintiffs may have fared better if they had devoted more space to "why individual freedom is important"³²

Finally, I thought it would be interesting to compare themes raised by different plaintiffs in a brief filed in a similar lawsuit. The plaintiffs—protagonists in the “State Government” category told a story different from the plaintiffs—protagonists in the “Private Individuals and Employers” category: The State Government plaintiffs—protagonists told a story about “federalism and states’ rights.”³³ Their goal was to preserve “state power.”³⁴ The obstacles were the PPACA itself and Congress, although they took “a more nuanced, and reasonable, view of Congress as antagonist.”³⁵ The following is the word cloud formed by the brief filed by the plaintiffs in *Florida ex rel. Bondi v. Dep’t of Health and Human Servs.*:



30 *Id.* at 129.

31 *Id.* at 113, n. 69.

32 *Id.* at 114.

33 *Id.* at 125.

34 *Id.* at 116.

35 *Id.* at 117.

The largest words in this word cloud are States, Medicaid, federal, Congress, State, Power, Clause, and Commerce. “States” is the most prominent word in the visualization of the “State Government” plaintiffs’ brief. By contrast, Act, Congress, and religious were the prominent words in the “Private Individual and Employers” plaintiffs’ brief. This difference is consistent with the varying themes identified by Professor Chestek.³⁶

IV. What are Wordle’s Limitations?

Given its rather simplistic treatment of words, Wordle has limitations. One limitation is that Wordle does not recognize word stems. In the Gettysburg Address word cloud, above, for example, if “dedicated” and “dedicate” had been treated as the same word, it would have appeared once, but larger, giving it more prominence.³⁷ Similarly, in the word cloud formed by the plaintiffs’ brief in the *Florida* case, above, if “State” and “States” had been combined into one word, its prominence would have been even stronger.

In addition, context can be lost by simply counting words.³⁸ For example, if a word were consistently used with a “not” preceding it in the original document, a viewer may draw the wrong conclusion from the word cloud.³⁹ Similarly, in the judicial opinion in the *Liberty University* case, for example, “religious” is prominent, which could lead the casual viewer to draw a conclusion that the court favored that claim when, in fact, it did not.

Further, even if one were to “combine meaningful phrases into joint words” (e.g., change “not happy” to “nothappy”) in the original text before creating the Wordle, “ambiguity could not be completely avoided.”⁴⁰ More-subtle connotations can be lost.⁴¹ Looking at the word cloud formed by the plaintiffs’ brief in the *Florida* case again, for example, the word “state” could mean a government body or it could mean to express.

Finally, “merely counting words does not permit meaningful comparisons of like text,” especially on the same topic.⁴² The three briefs analyzed above all had a different story to tell, allowing for meaningful

36 Because the state-government plaintiffs in this case prevailed, it appears that the states’ rights story, as visualized in this word cloud, was the winning story to tell.

37 See Andy Ramsden & Andrew Bate, *Using Word Clouds in Teaching and Learning* 4, <http://opus.bath.ac.uk> (August 2008).

38 Ramsden & Bate, *supra* n. 37 at 4; McNaught & Lam, *supra* n. 19 at 641.

39 Ramsden & Bate, *supra* n. 37 at 3; McNaught & Lam, *supra* n. 19 at 641.

40 McNaught & Lam, *supra* n. 19 at 641.

41 *Id.*

42 Feinberg, *Wordle*, in *Beautiful Visualization*, *supra* n. 1, at 55.

comparison; had the briefs told the same story, however, the word clouds would likely have been mostly indistinguishable.

Given its limitations, Wordle should be used only to supplement other traditional text-analysis methods.⁴³

V. Conclusion

Wordle can be useful for legal writers.⁴⁴ It is probably most useful to check or validate a theme after a document has been created, basically in the final stages of editing. A legal writer could, however, use it at the beginning of the writing process, too, by pasting notes into Wordle to help identify a theme from the get-go. Thus, Wordle can help lawyers and judges check and identify themes. It also offers a visually stimulating way to generate ideas, generally, and creates a fun escape from the sometimes monotonous legal tasks of the day.

Of course, it is not without limitation; words are “retrieved out of context,” and word forms are treated rather simplistically.⁴⁵ However, one can compensate for these limitations to a certain extent, and the program should only be used to supplement other analytic and idea generation tools. Further, my guess is that the next generation of visualization tools, which are likely already being created, will address these limitations and create even more useful applications in the legal context. I encourage legal writers to give it a try. A Wordle is a visualization as pleasing as it is revealing of what is essential about the text it illustrates.⁴⁶

43 *Id.*

44 Although beyond the intended scope of this essay, lawyers could use Wordle in many ways relevant to the practice of law. Wordle has been used to “illustrate news articles . . . and to distill and abstract personal and painful memories for victims of abuse.” Feinberg, *Wordle*, in *Beautiful Visualization*, *supra* n. 1, at 57. Similarly, it may be helpful to distill and to abstract depositions, interrogatory responses, admissions, and even witness testimony, to name a few.

45 McNaught & Lam, *supra* n. 19 at 641.

46 See Feinberg, *Wordle*, in *Beautiful Visualization*, *supra* 1, at 58.

