

## UNIGRAPHIC LIPOGRAMMATIC WINDOWS: A SURVEY

TRISTAN MILLER

Darmstadt, Germany

<http://www.nothingisreal.com/>

A *lipogram* is a span of text in which a particular letter or group of letters is avoided. Composition of very long lipograms is a popular pastime among logologists—indeed, some particularly dedicated constraint writers have produced entire novels which do not use the letter E [2, 3, 5]. However, lipograms can also occur entirely accidentally in running text. In this article, I present the results of my search for the longest unigraphic (single-letter) lipogrammatic windows in English literature.

Historically, literary lipogrammatic windows have not attracted the same level of interest as pangrammatic ones. This is probably down to two reasons. First, lipograms aren't as unique: whereas there is only one kind of pangram, there are as many different kinds of unigraphic lipograms as there are letters of the alphabet. Second, and perhaps more importantly, natural pangrammatic windows are much easier to spot with the naked eye. The object there is to find the *shortest* possible text containing all the letters of the alphabet, so the best examples will be at most a few dozen letters in length, and will be conspicuous by the close proximity of infrequent letters such as J, Q, X, and Z. Lipogrammatic windows, by contrast, can span tens or hundreds of thousands of words, and so any search for the very longest examples must be conducted with computer assistance.

The only previous lipogram search I am aware of is covered in Mike Keith's 1999 article "Literary lipogrammatic windows" [1]. Keith searched 15 000 books from Project Gutenberg [4] for passages missing either E or T, the two most common letters in English. This article extends his work to the other 24 letters of the alphabet. I can also improve upon his E and T results, thanks to a doubling in the number of Project Gutenberg editions in the intervening 15 years.

Before I present my results, I should like to give a short discussion of the sort of qualities embodied in a "natural" lipogrammatic window. First, any obviously intentional lipogrammatic writing should be excluded from consideration. Second, the text should be in English; we should discount any text which contains a significant amount of foreign words, such as would be found in a tourist's phrasebook or a textbook for second-language learners. The text should also be in a standard form of English with relatively modern spelling. Finally, the text should be coherent and cohesive, with a natural progression of thought and without excessive repetition in form or vocabulary. This last criterion excludes things like tables, indices, and mere lists of names, words, or phrases. Depending on one's point of view, it probably also excludes certain forms of esoteric or otherwise stilted poetry and prose.

In the results I present below, I aim to find the longest lipograms which unambiguously meet all the above criteria, though I also report some longer but still-interesting findings which violate these constraints. Where the lipogrammatic windows are quoted, I include the surrounding context, and mark the lipogram itself in italic text set off with square brackets. Those lipograms which I mention but which are too long to reprint here will be made available on my website at <http://www.nothingisreal.com/lipograms>.

**A** The longest A-lipogram in Project Gutenberg is Charles Bombaugh's poem "Incontrovertible Facts", written in 1875. A univocalic poem, it is a 549-letter lipogram for A, E, I, and U. Due to its artificial construction, however, it's not the sort of lipogram we're looking for. The next-longest contender is Lucy Sprague Mitchell's *Here and Now Story Book* (1921) for young children: it has one story, "How the Engine Learned the Knowing Song" which contains a chant consisting of nothing but the phrases "I'm coming," "I'm here", and "I'm going" repeated over and over again. Together with the surrounding text, this forms a 306-letter A-lipogram. However, the excessive repetition discounts this one as well.

The longest unequivocally natural A-lipogram in the corpus is the work of Charles Kinglsey, an Anglican priest and close friend of Charles Darwin. "The Unchangeable One", Sermon 33 in his collection *Town and Country Sermons* (1861), contains the following 275-letter specimen:

*Then, if the old Psa[<sup>l</sup>m]mist could trust God, how much more should we? If he could find comfort in the thought of God's order, how much more should we? If he could find comfort in the thought of his justice, how much more should we? If he could find comfort in the thought of his love, how much more should we? Yes; let us be full of troubles, doubts, sorrows; let times be uncert]ain, dark, and dangerous...*

**B, C** A 1902 presidential proclamation by Theodore Roosevelt establishing the Teton Forest Reserve contains a whopping 3817-letter B-lipogram, though it is little more than a lengthy prose boundary delimitation. I do not reproduce it in full here, though the following excerpt is enough to give one an idea of its nature:

*... along the western boundary of said reservation to the point for its intersection with the township line b[<sup>e</sup>tween townships thirty-three (33) and thirty-four (34) north; thence westerly along said surveyed and unsurveyed township line to the southwest corner of section thirty-four (34), township thirty-four (34) north, range one hundred and nine (109) west; thence northerly to the northwest corner of section three (3), said township; thence westerly...*

Similarly formulaic border demarcations are the source of other long lipograms: a 1910 proclamation by William McKinley contains an L-lipogram of length 1540, an 1890 one by Benjamin Harrison has a monstrous 11 876-letter M-lipogram, and an 1899 one also by McKinley contains both a 11 402-letter V-lipogram and a 11 342-letter Y-lipogram. I mean to disregard all of these.

*Sacred Books of the East*, a 1900 selection of eastern scriptures by Epiphanius Wilson, contains a much more natural-sounding 3641-letter B-lipogram. This is found in James Darmestetter's translation of selections from the *Zend*, a Persian Zoroastrian text. The longest C-lipogram (2600 letters) and Q-lipogram (47 333 letters) are also found here.

**D** Novelist Gertrude Stein was renowned (some would say notorious) for her highly unconventional, repetitive style of writing. Her 1933 book *Matisse Picasso and Gertrude Stein with Two Shorter Stories* contains lengthy lipogrammatic windows for several letters, including D (length 1319), E (171), F (3809), I (337), K (8463), L (2938), M (1898), O (694), P (6048), R (1930), T (524), U (3217), W (2039), and Y (2261). But for the author's idiosyncratic literary style, each of

these could lay claim to being the longest lipogram for its respective letter. Except for her relatively short E-lipogram, I won't reproduce any of them in this survey, and shall exclude them from serious consideration.

Apart from Stein's experimental writings, I found a 787-letter D-lipogram in the exercises to Chapter XIX of *Deductive Logic* (1888) by St. George Stock. However, it consists of a list of isolated propositions rather than proper running text. If we admit poetic lipograms, then a 610-letter portion of John Martin Crawford's 1888 translation of the Finnish epic poem *The Kalevala* may count. However, this passage is nearly as repetitive as Stein's. Perhaps the most unassailable D-lipogram in running prose is this one of length 562 from *Sociology and Modern Social Problems* (1910) by Charles A. Ellwood:

Thus, in the state of New York 5.5 per cent are illiterate, but of the native whites only 1.2 per cent are illiterate, while 14 per cent of the foreign population can neither read [*nor write*. Again, in Massachusetts 5 per cent of the population are illiterate, but of the native whites only 0.8 per cent are illiterate, while 14.6 per cent of the foreign born are illiterate. Statistics of illiteracy for our cities show the same results. Thus, in the city of New York 6.8 per cent of the population are illiterate, but only 0.4 per cent of the native whites are illiterate, while 13.9 per cent of the foreign born are illiterate. Boston has 5.1 per cent of its total population illiterate, but only 0.2 per cent of its native white population are illiterate, while 11.3 per cent of its foreign-born population are illiterate. Of the total immigration in 1907, 30 per cent were illiterate. The number of illiterates from] different countries varies greatly.

**E** The longest E-lipogram found in Keith's search was a 133-letter passage from Charles Dickens's *Our Mutual Friend*. As mentioned above, Gertrude Stein's *Matisse Picasso and Gertrude Stein with Two Shorter Stories* bests this with a 171-letter window:

Any one [*continuing is continuing. Continuing is continuing. ¶ Continuing and assisting assisting continuing is assisting continuing. Continuing is continuing. Assisting continuing is assisting continuing. ¶ Some*]e one who is continuing is assisting continuing.

If such avant-garde texts are excluded, the next longest E-lipogram may be this 158-letter specimen from *The Harbor* by Ernest Poole:

"... [*Look at all you have done to your bosses—and laugh! To this town, to this nation—and laugh, laugh! Look—and think—of what you can do—all you—and you—and you—and you—by just folding your arms! Think of all you will do! And laugh—laugh! Laugh! Laugh!*]" ¶ *H*]e broke off with both arms raised, and there followed one moment without a sound.

A number of other texts have even longer lipograms, though in all cases they are intentional word-play, onomatopoeic poetry, or long lists of proper names. For example, Richard Francis Burton's *Personal Narrative of a Pilgrimage to Al-Madinah and Meccah* (1855) includes a translation of a "specimen of a murshid's diploma" which contains an E-lipogram of length 434. However, all but the first eight letters of it are part of a lengthy genealogy.

**F** Children's writer Gertrude Smith provides what is probably the longest natural-sounding F-lipogram. "The Sleepy-time Story", a story of hers which appears in Volume I of *Boys and Girls*

*Bookshelf: A Practical Plan of Character Building* (1920), contains a stretch of 1732 F-less letters. The text in question describes the antics of two young girls, Arabella and Araminta, and their playful kittens.

**G** Project Gutenberg distributes an e-book containing three parallel translations of the Qur'an into English, from Abdullah Yusuf Ali, Marmaduke Pickthall, and Mohammad Habib Shakir. The 29th sura contains a G-lipogram of length 1805. Of course, we should probably discount this on account of it being the same material reproduced three times with only slightly different wording. More acceptable perhaps is a 1464-letter G-lipogram from Thomas Aquinas's *Summa Theologica, Part III*, as translated by the Fathers of the English Dominican Province.

**H** Amazingly, the five texts with the best claim to containing the longest H-lipogram are all on military subjects. *Kelly Miller's History of the World War for Human Rights* (1919) defines some basic units of military organization, abstaining from H for 683 letters. However, this can hardly be said to be running text. An 1839 military report by Captain C. Hagart of the Army of the Indus includes a 676-letter H-lipogram, though it too is not proper prose but rather a list of troops. Vol. 1, Issue 4 of *The New York Times Current History of the European War* (1915) includes a timeline of World War I, wherein can be found a 602-letter H-lipogram, and the United States' *Homeland Security Act of 2002* contains a 533-letter H-lipogram; however, in both cases the text is mostly in point form rather than fluent paragraphs.

The longest H-lipogram of indisputably "normal" prose may be this 396-letter specimen from *Civilization and Beyond: Learning from History* (1975) by Scott Nearing:

The logical outcome of such a situation is preparation for a war of independence by the vanquished, countered by military occupation, rigid suppression, and exploitation by the victors in th[e previous struggle. ¶ War is taken for granted as an instrument of policy. It is employed by civilized nations and empires as a means of expansion. Wars of independence and restitution follow conquest, dismemberment and annexation. Civilized nations and empires prepare for war and wage war as a normal aspect of civilized life. ¶ Civilization, and in particular western civilization, is a time-bomb, built to detonate and scatter its fragments far and wide. It is a type of booby trap in w]hich humanity has been caught periodically and horribly mangled.

**I** A library at Oxford University holds the untitled notebook of a 16th-century London grocer, Richard Hill, which is said to be filled with "the mental furniture of an average-educated man of the time". Included in the manuscript is a lengthy list of proverbs; my discovery therein of a 1422-letter I-lipogram is perhaps not so remarkable given Hill's archaic spelling (e.g., "A byrde yn honde ys better than three yn the wode"). A more running text, though still in obsolete spelling, is *The Principal Navigations, Voyages, Traffiques and Discoveries of the English Nation* (1589–1600) by Richard Hakluyt. Volume VIII of this work contains a 936-letter I-lipogram.

The longest I-lipogram in modern spelling which does not resort to excessive repetition may be this 296-letter example from *Flag and Fleet: How the British Navy Won the Freedom of the Seas* by William Wood (1919):

But this itself depends on the second reason, which, in its turn, depends upon the thi[rd]. *For we never could have won the greatest sea-power unless we had bred the greatest race of seamen. And we never could have bred the greatest race of seamen unless we ourselves had been mostly bred from those hardy Norsemen who were both the terror and the glory of the sea.* ¶ *Many thousands of years ago, when the brown and yellow peoples of the Far South-East were still groping their way about their steamy Asian rivers and hot shores, a race of great, strong, fair-haired seamen was growing in the North.*

**J** The letter J is a relatively recent addition to the English alphabet, originating as a consonantal variant of I. Older texts use it rarely or not at all. For example, the 1623 collection of Shakespeare's plays which is known today as the *First Folio* has long stretches with nary a J, the longest of which is 739 210 letters in length.

The King James Version (KJV) of the Bible, at least in modern editions, does use J consistently, but its first appearance in *Leviticus* isn't until after 58 036 letters. (The last word of the preceding book, *Deuteronomy*, is "journeys"; if we consider the KJV to be a single text, the length of the J-lipogram is actually 58 043.)

**K** An 1867 letter from US president Andrew Johnson to the House of Representatives, as reproduced in Volume VI of *A Compilation of the Messages and Papers of the Presidents* by James D. Richardson, contains a K-lipogram of length 12 712. The letter is the signing statement for an appropriations bill, and contains a number of largely templated paragraphs objecting to various claims made in the bill. There are K-lipograms of similar length, and employing similarly formulaic structures, to be found in other legal, governmental, and technical documents. The longest K-lipogram which does not employ overly repetitive or formulaic language is probably an 8212-character passage from an 1855 veto from US president Franklin Pierce, as reproduced in Volume V of the aforementioned *Compilation*.

**L** Most of the longest L-lipograms are extremely repetitive texts in the form of avant-garde writing, boundary demarcations, parish rolls, cumulative rhymes, etc. The longest window of what I would consider to be ordinary, narrative text is found in Volume III of *Tales from the Arabic*, a 1901 edition of the Arabian Nights. One of its stories, "Shehrzad and Shehriyar", has a 725-letter L-lipogram.

**M** Religious texts provide two M-lipograms of impressive length, though in both cases the passages should probably be excluded due to excessive repetition: The aforementioned *Zend* contains a 1302-letter M-lipogram in its instructions for "Cleansing the Unclean", and a 1289-letter specimen can be found in that part of the KJV's *Exodus* which describes the construction of the tabernacle.

The longest M-lipogram in text which is not overly repetitive is probably a 1186-letter example from *Bunny Brown and His Sister Sue at Aunt Lu's City Home*, a 1916 children's book penned by the pseudonymous Laura Lee Hope.

**N** In 1917 Australian poet Frank Wilmot (a.k.a. Furnley Maurice) published the poem “Invalid”, written in the voice of a bedridden cold sufferer: “Raid, raid, go away, / Dote cub back uttil I say, / That wote be for beddy a day. . .” This near-total abstention from nasal consonants results in a charming, if artificial, 718-character N-lipogram. I found several other similar texts, including “To Bary Jade”, a 690-letter poem by Charles Follen Adams which appears in Volume X of *The Wit and Humor of America* (1908), and “Bood”, a 474-letter poem found in the 14 November 1917 issue of *Punch*.

The longest prose example which does not employ reiteration for its own sake may be these 255 letters, which appear in an 1867 veto from Andrew Johnson (printed in Volume VI of *A Compilation of the Messages and Papers of the Presidents*):

The officer of the Army must, if “detailed,” go upon the supreme bench of the State with the same prompt obedience as if he were detailed to go upon [a court-martial. *The soldier, if detailed to act as a justice of the peace, must obey as quickly as if he were detailed for picket duty.* ¶ *What is the character of such a military civil officer? This bill declares that he shall perform the duties of the civil office to which he is detailed. It is clear, however, that he does*] not lose his position in the military service.

**O** Many of the longer O-lipograms are of dubious quality. For instance, there is a 509-letter specimen in Kisari Mohan Ganguli’s 19th-century translation of Book VI of *The Mahabharata of Krishna-Dwaipayana Vyasa*, though it’s part of a colossal list of places and peoples. Another example can be found in “The Cataract of Lodore”, an 1820 ode to the Lodore Falls by English poet laureate Robert Southey. Hailed as a “masterpiece of onomatopoeia”, it contains a 396-letter O-lipogram.

The most natural prose example I could come up with was this 323-letter one from Arnold Bennett’s novel *A Great Man* (1904):

A lady entered, chattered at an incredible rate in Italian, and disappeared behind the screen, where she knocked a chair over and rang fo[r the waiter. *Then the waiter entered and disappeared behind the screen, chattering at an incredible rate in Italian. The waiter reappeared and made his exit, and then a gentleman appeared, and disappeared behind the screen, chattering at an incredible rate in Italian. Kissing was heard behind the screen. Instantly the waiter served a dinner, chattering always behind the screen with his cust]omers at an incredible rate in Italian.*

**P** The fairy tale “Chicken-Licken”, as printed in the 1927 anthology *Childhood’s Favorites and Fairy Stories*, does not use the letter P until the very last sentence. Including the title of the story itself, this amounts to 2519 consecutive non-P letters.

**Q** As mentioned above, the longest Q-lipogram contains 47 333 letters, and forms part of James Darmestetter’s translation of the *Zend*.

**R** Many long R-lipograms are non-rhotic speech written in eye dialect. Examples in excess of 400 letters are to be found in *American Adventures: A Second Trip "Abroad at Home"* (1917) by Julian Street, *Dr. Sevier* (serialized 1883–1884) by George W. Cable, *The Boss of Little Arcady* (1905) by Harry Leon Wilson, the poem "The Capture" (before 1914) by Paul Laurence Dunbar, and the short story "Ching, Ching, Chinaman" (1917) by Wilbur Daniel Steele. However, these should all be discounted on the grounds that the writer has made a conscious decision to avoid using R.

A 438-letter R-lipogram occurs in William Tyndale's 1526 translation of *Gospel of Luke*. While it is indisputably accidental, it occurs in a lengthy genealogy as opposed to fluent narrative or expository prose. The longest R-lipogram in unequivocally natural text is probably this 388-letter specimen from George W. Peck's *Peck's Bad Boy at the Circus* (1905):

Another [fellow and I got a system on slot machines, and one day we beat the machines out of a shotbag full of nickels, and when we showed up at the tent all the fellows wanted to know how we did it, and pa said it was gambling, and we ought not to do it, but he also wanted to know how we managed to win, and when we told pa about it pa said it was no sin to beat a slot machine, 'cause it was an inanimate thing, just a machine, and anybody who could beat a nickel in the slot machine at his own game was equal to a] Rockefeller.

**S** As with the letter R, many writers deliberately refrain from using the letter S to affect a particular regional accent or speech impediment. The longest such example I discovered is in Charles Dickens's *Hard Times* (1854), where 1339 S-less letters are used to transcribe the lisping speech of circus master Mr. Sleary: "Thethilia, it doth me good to thee you. You wath alwayth a favourite with uth. . ."

In *The Dodge Club or, Italy in 1859* (1872) by James De Mille, a character tells a story which includes some very repetitive and noncommittal haggling between his grandfather and a ship captain. The dialogue contains an S-lipogram of length 379, though it should probably be discounted on account of its excessive repetition.

A more natural example is the following 359-character S-lipogram, from Marmaduke Pickthall's aforementioned translation of the Qur'an:

And unto you belongeth a half of that which your wives [leave, if they have no child; but if they have a child then unto you the fourth of that which they leave, after any legacy they may have bequeathed, or debt (they may have contracted, hath been paid). And unto them belongeth the fourth of that which ye leave if ye have no child, but if ye have a child then the eighth of that which ye leave, after any legacy ye may have bequeathed, or debt (ye may have contracted, hath been paid). And if a man or a woman have a di]stant heir (having left neither parent nor child), and he (or she) have a brother or a sister (only on the mother's side) then to each of them twain (the brother and the sister) the sixth. . .

**T** Keith's survey did not turn up any prose T-lipograms meeting my criteria; the longest passage he found (213 letters) was a selection of definitions from the 1913 edition of *Webster's Unabridged Dictionary*.

The longest T-lipogram I found was in *Rolling Stones* (1919) by O. Henry. In this book the author reproduces one of his letters to his editor, which contains a 231-letter T lipogram. However, the intentionally excessive repetition, combined with the fact that the letter was never originally intended for publication, raises doubt over whether this lipogram should be accepted.

Perhaps the longest indisputable prose example is the following (223 letters) from Joseph A. Altsheler's *The Texan Star: The Story of a Great Fight for Liberty* (1912):

He had a look of health despite the dead whit[eness of his face, which Ned now knew was caused by prison pallor. Ned liked him. He liked him for many reasons. He liked him because his eyes were kindly. He liked him because he was one of his own race. He liked him because he was a fellow prisoner, and he liked him above all because] this was the first human companionship that he had had in a time that seemed ages.

If we admit poetic lipograms, then at 296 letters, the entirety of Victor Hugo's "Guitar Song" (1838), as translated by Evelyn Jerrold, qualifies.

**U** A very formulaic, 2392-letter U-lipogram is included in the instructions for making a doiley in *Golden Stars in Tatting and Crochet* (1861) by Eleonore Riego de la Branchardière. A less formulaic U-lipogram, albeit in archaic spelling, is a 1935-letter example from the ballad "Robin Hood and the Potter", known from a Cambridge manuscript dating from the reign of Henry VII.

Perhaps the most natural-sounding U-lipogram in modern spelling is one of 1023 letters (not counting the Greek letters used as variable names) from *Geometrical Solutions Derived from Mechanics* by Archimedes, as translated by J. L. Heiberg and Lydia G. Robinson (1909).

**V** "Sis' Becky's Pickaninny", a short story by Charles W. Chesnutt published as part of his famous *The Conjure Woman* collection, contains a V-lipogram of 7281 letters. However, as with many of the R- and S-lipograms, it's written in eye dialect. For standard English, we can turn to *A Kindergarten Story Book* (1916) by Jane L. Hoxie, a collection of short stories for children. A passage spanning three of the stories ("Billy Bobtail", "Kid Would Not Go", and "Fox Lox") forms a lipogram of length 5829.

**W** Like J, the letter W is a relatively recent addition to the English alphabet, originating as a consonantal variant of the double U; U in turn was originally a variant form of the letter V. It's not surprising, then, that many early texts often or always use UU or VV where we would normally write W. One such example is the 1582 Douay–Rheims translation of the *Epistle of Jude*, which has a W-lipogram of length 2125.

It's not hard to find other examples spanning many thousands of letters, but in most cases these are contained in mere lists, or are written in eye dialect. For example, the short story "Little Jakey" by S. H. Dekroyft, published in *Stories of Childhood* (1875) contains a 1826-letter W-less narration by a German-accented character. The longest truly unassailable example of a W-lipogram is probably a 1362-letter passage in *Evolution Of The Japanese, Social And Psychic* (1905) by missionary Sidney Lewis Gulick.



**X** A notable X-lipogram in poetry can be found in George Henry Needler's 1904 translation of the German epic *Nibelungenlied*. One 38 742-letter passage, beginning in the Sixteenth Adventure and ending in the Twentieth, contains nary an X. The longest prose example I found is a 35 028-letter stretch in the World English Bible revision of the *Book of Jeremiah*, beginning in the 34th chapter and ending in the 46th. A close runner-up is a 34 819-letter specimen in *Gods and Fighting Men: The Story of the Tuatha de Danaan and of the Fianna of Ireland*, arranged and translated by I. A. Gregory in 1905.

**Y** Finding an unambiguously fluent Y-lipogram in fully modern English spelling turned out to be quite a challenge. There are dozens of Y-lipograms from dance manuals, cookbooks, Biblical genealogies, medieval chronicles, and other texts which employ highly repetitive and/or archaic language. Perhaps the least contentious example is a 1688-letter specimen from the 1913 *Popular Science* tome *The Boy Mechanic, Volume 1: 700 Things For Boys To Do*.

**Z** Eugene Mason's prose translation of the Medieval historiographic poem *Brut* includes a Z-lipogram of 141 151 letters—nearly half the length of the poem.

## References

- [1] Michael Keith. Literary lipogrammatic windows. *Word Ways*, 32(4):294–296, November 1999.
- [2] Georges Perec. *La Disparition*. Éditions Gallimard, Paris, 1969.
- [3] Georges Perec. *A Void*. Trans. Gilbert Adair. The Harvill Press, London, 1995.
- [4] *Project Gutenberg*. <http://www.gutenberg.org/>.
- [5] Ernest Vincent Wright. *Gadsby*. Wetzel Publishing Co., Los Angeles, 1939.