



do not  
resist  
artist  
eat  
buy  
& know  
soulle  
sness

*Subwords* is a series of post-Oulipian constrained textual pieces centred around one algorithmic idea: the decomposition of a (long) word into (shorter) ones, also known as the subsequence problem, where the letters of the ‘subwords’ must be present in the original word in order, but not necessarily contiguously. Example: “theory”, containing both “her” and “toy”. Our implementation also imposes that the non-contiguous parts of a subword must not be known words, and that all the letters in the original ‘superword’ be accounted for (exhaustive decomposition). Given a list of words, the algorithm finds all such decompositions, or ‘nexuses’, and from that database a few are then selected as the basis for writing – either as a seed for the imagination, fragments to be continued, enigma to be solved, or Rorschach shadows to be projected on. This paper also briefly exposes additional constraints adopted in the final phase of this process, and provides a short overview of (mostly Oulipian) letter-based techniques in order to contextualize the present set of constraints in the landscape of postwar constrained literature.

## J r mie Wenger

Goldsmiths College, University of London, London, UK  
j.wenger@gold.ac.uk  
jeremiewenger.com  
@jchwenger

**Keywords** Computational Literature, Generative Literature, Constrained Writing, Process-led Writing, Oulipo, Poetry, Literature

**DOI** [10.34626/2025\\_xcoax\\_031](https://doi.org/10.34626/2025_xcoax_031)

## Introduction

The present literary project is an ongoing series of post-Oulipian constrained pieces written from, or around, a particular generative process: the decomposition of a (super)word into (sub)words, where the superword and all subwords are a known words, with the subword letters remain in order but allowing for gaps, and where all letters of the superword are used, therefore a part of one (and only one) subword. Here is one example:

he ɿ  
t o y (theory)

The texts of this series all contain one such decomposition, attempting to find ways to write from, despite, or around it.

### A Short Primer on Letter-based Constraints

A short overview of past experiments with letter-based constraints can be beneficial to position this series in the landscape of experimental literature. Without a doubt the most important influence of this work is Georges Perec, perhaps the most famous member of the Oulipo (*Ouvroir de Littérature Potentielle*, Workshop of Potential Literature) who experimented with a wide array of constraints throughout his writing career. Another source of inspiration was Jean Starobinski's study of Saussure's late obsession with anagrams in *Les mots sous les mots* (1971), translated by Olivia Emmet as *Words Upon Words* in 1979. Figures such as John Cage and Jackson Mac Low are, of course, other well-known writers who experimented with such systems.

Studying the techniques of the Oulipo reveals that letter-based experimentation (as opposed to syntactic, or phonetic, constraints) was perhaps the most common approach to formulate systematic constraints. Here are a few notable examples:

- Acrostic: a poem where the initial letter of each line forms a text. Relaxing of the first-position constraint leads to the *mesostic*, explored by John Cage (Cage 1973; Cage 1979a; Cage 1979b; Cage 1983), and by Jackson Mac Low under the name *diastic* (Mac Low 1989).
- Lipogram: one or more letters are banned from the language available to the writer, most famously “e”, the most common letter in French, absent from Perec's novel *La Disparition* (1969), translated by Gilbert Adair as *A Void* in 1995, or all vowels except “e” in the univocalic piece *Les Revenentes* (1972) translated by Ian Monk as *The Exeter Text: Jewels, Secrets, Sex* in 1972.

- Anagram: given an initial string of text, only permutations of the letters are thenceforth allowed. Examples include Michelle Grangaud's *Memento-fragments* (1987), using titles of books, paintings, music pieces, or famous quotes, and *Stations* (1990), using Paris metro station names. Another extreme example is Perec's *Ulcérations* (1974), where the title word also sets the allowed set of letters, and each of the subsequent 400 lines is a permutation of this set.
- Heterogram: a poem where each line is written using all letters of the alphabet (exemptions for very rare letters are often applied to relax the stringency of the constraint). An accomplished example is Perec's collection *Alphabets* (1976), where the allowed letters are reduced to the 10 most frequent plus 1 (the 'tonality' of the poem, that can be 'in H', or 'in L...'), yielding 11x11-lettered square poems.
- '*Beau présent*' (beautiful present, with a *double entendre* on gift and presence): a love poem, or some other form address, written using only the letters present in the addressee's name (Perec 1994).
- '*Belle absente*' (beautiful absentee, formed as the 'antonym' of the beautiful present): the same as above, where each line is a heterogram, containing all letters of the alphabet once, *except* the letter being emphasized. Like the acrostic, each line reveals/hides a new letter, and the poem, like the acrostic, contains as many lines as the addressee's name (Perec 1994).
- Palindrome: a string of text that is identical when read forward or backward (punctuation usually excluded). Perec's "Great Palindrome" contains 5'556 letters, 1'247 words (Perec 1980).
- One should also mention an interesting, idiosyncratic setup occurring in chapter LI – the so-called *Compendium* – of Perec's *La Vie mode d'emploi* (1978), translated as *Life: A User's Manual* by David Bellos in 1987: three blocks of 60 lines of 60 characters each (including spaces, and with the last block containing only 59 lines – a wilful 'error', or 'clinamen', often introduced by Oulipians in their own systems (Andrews 2022, 127-162). Each line summarises events around 179 characters from the book, without naming them. Each block contains a 'letter diagonal', the same letter occurring – last in the first line, then next-to-last in the second line, next-to-next-to-last in the third, etc. (The letters for the three blocks are A, M and E, in French *âme*: soul).

In all these instances, a specific, letter-based rule is adopted *a priori*, and combined with either additional processes (such as chance operations, somewhat favoured by the Anglo-American avant-garde) or a subjective approach (“Given constraint X, how can ‘natural’ writing proceed?” – the predominant orientation of the Paris-centric Oulipo). The present series, even when it combines multiple constraints, firmly sits in the second camp, aspiring to a renewed and personal literary subjectivity, rather than an ‘emancipation from the subject’ or to ‘asubjective writing’.

## Subwords

### Subsequence as Constraint and Generative Principle

In this context, the proposed constraint is grounded in a foundational algorithmic idea: the decomposition of longer words into shorter ones, known in computer science as the subsequence problem. In this context, we seek to find whether a certain string is present in another, respecting the order of the letters, but not necessarily in their contiguity, allowing for ‘gaps’ between parts of the string (Cormen et al. 2022). Given a list of words or ‘dictionary’, our script (Python, 2017-19, unpublished) loops through all sufficiently long words (six letters or more), and attempts to decompose those using only subwords from that same dictionary.

Our implementation adds the following constraints:

- The non-consecutive parts of subwords (decomposing the longer superword) are not allowed to be words themselves: ‘anti + establishment’ would be not a valid subword of ‘antidisestablishmentarianism’, thus removing ‘trivial’ results.
- All letters of the superword must be accounted for (exhaustive decomposition).
- To limit the amount of results, subwords of two letters or less are excluded.

The peculiarity of such an approach is the creation, from the dictionary, of a database of constrained material. When run on a small list – say, the 10,000 most frequent words of English from the Google’s Trillion Word Corpus (Michel et al. 2010), available in the GitHub repository *Wordhoard* (Wenger 2019) under English/google/google-10000.txt – the algorithm already finds more than 5,000 decompositions. This presents the prospective writer with a first question: what to do with such quantities? Here, a first important subjective act is required: the writer could decide to present the plethora as is, in

one form or another; one could devise a specific procedure to select only a small subset from that data (a procedure which could be formal, based on letters again, or thematic, searching for the presence of particular words in super- and/or subwords); or one could select, as it were, 'fully' subjectively, according to personal preferences. Our approach so far has oscillated between the last two options, developing relatively crude scripts to search through the data in either a formal or a thematic way.

At this stage, the writer has 'bare' decompositions, or nexuses, at hand, that can in some cases stand out as individual finished pieces, offering already interesting, or beautiful, clusters of meaning:

w a r  
e ve (weaver)

\*

ea t  
b u y (beauty)

\*

rat s  
a te  
l ion (alterations)

\*

s ms  
e y e  
cos t (ecosystems)

However, it seemed that if left there, the series would in effect have reduced the act of writing to something akin to curation, an idea which, despite its popularity in certain artistic or literary circles, felt to us resolutely insufficient. In order to make room for more subjectivity into the process, the most obvious option was to consider those as fragments, or starting points, in need of completion – echoing, surprisingly, the work facing translators of ancient texts (see, for instance, Carson 2002). Viewed as a fragment, the nexus can now be viewed as surrounded by empty (erased? yet uncreated?) space, where writing can occur.

u n e q u a l t r e a t i e s o f a g i n g : b a r b a r i a n s  
a t t h e g a t e s o f m y b r a i n , c a n n o n a d i n g

Our literary cake is almost ready, missing only the archaic cherry of writing.

### Nexus-based Writing: Beyond the Database

Given a particular nexus, two ‘modes’ are immediately available to the writer: the horizontal one, where each subword sits on one line, with available blank space on the left and the right, see Fig. 1); and the vertical one, where instead blank lines exist above and below the decomposition (in this setup, an elegant constraint can be to restrict each line to contain exactly as many characters as the superword, see Fig. 2). It goes without saying that the two can be combined, and that other possibilities exist (for fiction, for instance, the superword could be the title, each subword governing a chapter).

The evolution of the series so far also led to a final layer of constraints:

- In the horizontal mode, the boundaries between the nexus and the rest of the text can be clear (no writing ‘encroaching’ on the nexal space, with boundaries corresponding with the length of the superword, which can help preserve the legibility of said superword) or blurred (regular spacing before and after each subword).
- In this mode, called ‘internal’, one stringent exercise is to allow writing only in the interstices left around the subwords: filling the blanks without going beyond the superword width.
- A ‘uniform’ (square, rectangular) effect with the same length for all lines. (Other possibilities include: ‘staircase’, incrementing/decrementing the line length at each line, with the in-/decrement either static or itself in-/decreasing; ‘corner/saw tooth’, the same, but repeated, with alternating in-/decrements).

With the reframing of nexuses as fragments, something peculiar occurs. This ‘new’ blank space is nearly the same as what we started with: undetermined – except that it is now the space around that particular nexus. The nexus can then act like the abstract shapes on a Rorschach test – a test of our own choosing –, calling for subjec-

Fig. 1. “bargaining” (horizontal mode, uniform line length).

do not  
resist  
artist  
e a t  
b u y  
& know  
soull e  
ssness

Fig. 2. “beauty” (vertical mode, 6 letters per line, incl. spaces).

tive completion, and intimate associations. Meanwhile, the superword hovers above (or below) like a ‘negative’ title or theme, recalling the *‘belle absente’*.

Thinking about how to characterize this last phase – perhaps the least determined, even though it is the most constrained – various metaphors emerged:

- Writing ‘from’: the constraint as a seed, from which a plant (the text) grows. Computer-scientifically infected people may be triggered by the use of the word ‘seed’ – the name for the nugget of true randomness injected into pseudo-random number generators. We would be glad, even if this analogy is less leafy.
- Writing ‘despite’: the constraint now as an obstacle, to be overcome. This one is very much in line with the Oulipo itself, resonating with a well-known quote by Queneau, one of the founding members of the movement: “Oulipians are rats who construct the labyrinth from which they plan to escape” (Matthews 2005, 205; Bens 2005, 49; Andrews 2022, 21, 162; Higgs 2009; Long 2015). Trees have the power to grow roots through rocks or concrete.
- Writing ‘around’: the nexus as support, fostering growth, with the auctorial subject as creeper (or fig tree, if the host ends up stunted). This also recalls the animal kingdom: the creation of pearls (the text) by oysters, if triggered by small rocks or specks of dust (the constraint).

Thanks to this last stage, and despite their Dickinsonian concision (the haiku also comes to mind), we hope that these pieces can become more than the mere exposition of the constraints underpinning them, and start enjoying a fleeting existence in the thick, wild and generous undergrowth of literature.

### **Post Scriptum: Towards a Series of Series**

This series has been growing alongside other nexus-oriented projects such as *Squares* (using square of letters, where rows, columns, and/or diagonals, form valid words), *Cubes* (writing on cubic surfaces, two of which are squares), *Werewords* (close to *Subwords*: here two words are connected by a common, not necessarily semic letter pattern), and *Laces* (circular paths of Levenshtein-adjacent words).

Website:

[jeremiewenger.com/subwords](http://jeremiewenger.com/subwords)

ethical pe t project  
#1: r es c a le all  
supe r b ness

wizardry and w it be  
kings in this r ing

how many b r in g-it-downs  
for every a ga in-it-rises?

lest th e conspiracies of our ravens,  
or our r ats' mischiefs, be unleashed

thought c-beams glitter in the dark after  
the fear, when he, former slave, discerns  
the unending ep ic struggle between  
the becoming-l ion of ambitions and  
the becoming-r at of satisfactions  
near the tannhäuser gate – events lost in  
the time rain – off the shoulder of orion

mu s i c ?

p o e s y ?

a l l

p o o

( b e s t ,

l a c a n )

m a z e w / o

o v e r l a p ,

p e r i o d s ,

b u t w i t h

f i v e f o l d

r o t a t i o n

s y m m e t r y

– w h a t a

t i l e

m e s s

( k i n d e s t

r e g a r d s ,

p e n r o s e )

## References

- Andrews, Chris.** 2022. "Meaningful Forms and Clinamen." *How to Do Things with Forms: The Oulipo and Its Inventions*. Montreal & Kingston, London, Chicago: McGill-Queen's University Press.
- Bens, Jacques.** 2005. *Genèse de l'Oulipo 1960–1963*. Bordeaux: Le Castor Astral.
- Cage, John.** 1973. M: *Writings '67–'72*. Middletown, CT: Wesleyan University Press.
- . 1979a. *Roaratorio: An Irish Circus on Finnegans Wake*. Leipzig, London, New York: Edition Peters.
- . 1979b. *Empty Words: Writings '73–'78*. Middletown, CT: Wesleyan University Press.
- . 1983. *X: Writings '79–'82*. Middletown, CT: Wesleyan University Press.
- Carson, Anne.** 2002. *If Not, Winter: Fragments of Sappho*. London: Virago.
- Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest and Clifford Stein.** 2022. "Longest common subsequence", *Introduction to Algorithms, fourth edition*. Cambridge, MA: MIT Press, 393-399.
- Higgs, Christopher.** 2009. "The Annoying Lacuna: One Unofficial History of the Oulipo." *AGNI Online*, July 1. <https://agnionline.bu.edu/essay/the-annoying-lacuna-one-unofficial-history-of-the-ouliipo>.
- Long, Michael.** 2015. "Rats Build Their Labyrinth: Oulipo in the 21st Century." *Hyperallergic*, May 17. <https://hyperallergic.com/206802/rats-build-their-labyrinth-ouliipo-in-the-21st-century>.
- Mac Low, Jackson.** 1989. *Words and Ends from Ez*. New York, NY: Avenue B.
- Mathews, Harry, and Alastair Brotchie, eds.** 2005. *Oulipo Compendium*. 2nd ed. London: Atlas Press.
- Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, Joseph P. Pickett, et al.** 2010. "Quantitative Analysis of Culture Using Millions of Digitized Books." *Science* 331 (6014): 176–82. <https://doi.org/10.1126/science.1199644>.
- Perec, Georges.** 1980. *La Clôture et autres poèmes*. Paris: Hachette.
- . 1994. *Beaux présents, belles absentes*. Paris: Seuil.
- Wenger, Jérémie.** 2019. *Wordhoard*. Accessed May 8, 2025. <https://github.com/jchwenger/wordhoard>.