LIPOGRAMS AND OTHER CONSTRAINTS

A. ROSS ECKLER
Morristown, New Jersey

Many kinds of literary constraint are available to the logologist. On the one hand, E-less text is so easy to create that the reader may be unaware that it is happening. (Some reviewers of Perec's *La Disparation* never caught on.) On the other hand, palindromic text--probably the most popular form of literary constraint--is so difficult to create that, at best, it makes only local sense; coherent development of a topic is essentially impossible. Can one characterize in an objective way the relative difficulty of different sorts of constraint? Which ones should be cultivated more widely?

The literary challenge of constrained writing can be expressed in two forms. The stricter one, alluded to above, is the creation of text that sounds so natural that the reader is unaware that a constraint is present. A more modest goal is the creation of grammatically-acceptable text that clearly conveys a story or topic selected in advance (that is, not dependent upon the exigencies of the constraint). One simple way to demonstrate this is to paraphrase a well-known (short) text such as Mary Had a Little Lamb or the Gettysburg Address. The reader knows that <u>something</u> is going on, and his task is to deduce what sort of constraint has been employed.

The measure of difficulty that I propose is a simple one:what fraction of words are available when selecting the next word of the text? High-usage words are weighed more heavily than low-usage ones in this calculation; what is the probability that a randomly-selected word in running English text will be available? For the standard E-less lipogram, this fraction is always the same: the fraction of words in running text that do not contain E (about 0.5). In other constraints, this fraction may vary from word to word, and one must calculate its average value. The smaller this value, the more difficult the writer's task. Writing under restriction can, I think, meet the strict challenge if the fraction is 0.4 or more. The lesser challenge can, perhaps, be met with fractions as low as 0.2. It is hard to specify a rigid cutoff--sense and grammar gradually fade, like the grin of Lewis Carroll's Cheshire Cat.

It's time to move beyond E-less text. Obviously, one can write text omitting any other letter, although such a task becomes ridiculously easy once one leaves the high-frequency letters. It is better to move in the opposite direction, and require that each word contain a specified letter. For E, the task is just as easy as it was without E; both restrictions use about half the available words in running text. One can exhibit a pair of constrained writings on the same topic. Constructing T-

text, A-text, etc. becomes harder. As a rule of thumb, the fraction of words containing a given letter in running text is about four times the fraction of that letter in running text. Text containing (say) only C-words, M-words, P-words or the like will seriously challenge most writers.

One can omit more than one letter. Starting with the rarest letters, how many can depart before the task becomes formidable? A goal to shoot for is a half-alphabet--13 letters of one's choice. If these are ETAOINSHRDLCM, a short sample of running text revealed that 0.4 of all words were usable (62 out of 145, 11 being THE). I strongly doubt that anyone can write a paragraph or more of half-alphabet text that would not attract the reader's suspicion, but perhaps topics can be intelligibly paraphrased.

Univocalic prose omits all but one vowel (Y counted as well). Most leave the writer with too few alternatives to create smoothly-reading paraphrases. This E-univocalic for Mary Had a Little Lamb is about as good as they get:

> Meg kept the wee sheep.
> The sheep's fleece resembled sleet.
> Then, whenever Meg went,
> The sheep went there next...

In a sample of 145 words in running text, 31 (a fraction of 0.2) were E-vocalic, but of these 11 were THE. Other univocalics probably generate only labored paraphrases.

A natural way to deal with multiple letter-omissions or -insertions is homoliteral or heteroliteral text. In the former, each word has at least one letter in common with its immediate predecessor; in the latter, no letters in common. For typical running text a rough sort of parity is achieved: 0.4 of all words share no letters with their predecessor, and 0.6 do. Therefore, it should be possible to write fool-the-eye prose with either rule; texts such as the Gettysburg Address can be paraphrased both ways.

Doubly-homoliteral text (each word has at least two letters in common with its predecessor) can probably be used for paraphrases, but the task is much more difficult. Single-letter words (a, I) are unusable, and about 0.2 of words in running text are doubly homoliteral with their predecessor.

Homoliteral (and heteroliteral) text has one great advantage for the writer of restricted prose. Although on the average a certain fraction of words are not available, <u>any</u> word in English is theoretically usable by suitably selecting its neighbors. In contrast, E-less text and its relatives permanently ban a certain set of words, making it somewhat more difficult to express a given line of thought.

In acrostic-style restricted writing, one insists that the first letters of successive words in running text spell out a message of their own. Consider first the less-restrictive condition that <u>any</u> letter in the word (not just the first) can make the match. For instance, if the acrostic template is ACROSTIC, then one can write fAt Children requiRe mOre activitieS Than thIn Children. More interestingly, one can match the template to the text by requiring that the nth word of the text contain the nth letter of the text: Many cAts caN easilY disCern A tasTy morSel conCealed beneAth iNedible garbagE... These can be termed self-replicating texts. On the average, a fraction 0.2 of words in running text is available to choose from for the next word in the message. Paraphrases ought to be possible.

True acrostic writing is, of course, harder; the fraction of available words is probably only 0.05 or so. Howard Bergerson introduced Word Ways readers to self-replicating acrostic text in August 1975, dubbing it an automynorcagram. He also introduced the idea of enchainment: message A can automynorcagrammatize message B, message B can automynorcagrammatize message C, and so on, back eventually to message A.

One can insist that the first letter of a word be the same as the first letter of the preceding word, or alternatively the same as the last letter of the preceding word. Both are too difficult for paraphrasing, even though the latter can theoretically reach any word in the language. (But it's tough to get to J or Q!)

Shifting from prose to poetry, much looser acrostic restrictions become possible: typically, the acrostic template is read off from the initial letters of the successive lines of the poem.One can without much difficulty weave additional messages (all of the same length) into the poem: line endings (a double acrostic) or even intermediate letters (the nth letter of the nth line, running a diagonal). When the number of letters in the acrostic template does not match the number of lines, one can use variable-length acrostic indents:

ACorns grow to mighty trees;
ROSes never fail to please;
TICks now carry Lyme disease.

This generalization affords the opportunity to conceal messages that the reader is then challenged to find, as illustrated by the following doggerel, written in an hour or so by the author's sister on the day following his seventieth birthday:

Fiddles play, anthems raise!
Vestal virgins chant his praise.
Corn sheaves hanging on the door;
Eminent person has the floor.
Usual toasts and celebration:
The man's a genuine inspiration.
Irrational? No, still in top form

(Typical Eckler, a bit off the norm,
Luminary of the logological game).
Song and story should tell his fame.
Edification is what he's about --
Aye, birthday boy without a doubt.

The foregoing prose constraints have all been applied on a word-by-word basis--text can be built up by hand, one word at a time. In contrast, some constraints must be administered in larger doses. Take, for example, anagrammatic text in which one selects a chunk of n letters (some may be repeated) and repeatedly rearranges them to form intelligible text. For small values of n--5 to 20--one cannot hope to paraphrase intelligibly, let alone fool the reader, but either goal may well be achievable for n equal to 100 or so. The trouble is that the bookkeeping labor of selecting and arranging suitable words is immense, best left to a computer. (Note that a word may bridge two chunks in the text.) A similar, but less onerous, task confronts the writer who wishes to ensure that the letters of the alphabet, taken one at a time but in any order, are present (one letter per word) in 26-word chunks of text. The strategy is to select words containing the rarest letters first, placing them at suitable positions in the chunk, then fill in the narrative with words using the remaining letters. It is likely that almost any choice of these will result in most of the alphabet being included; only a little tweaking will be needed to capture an errant letter or two. 26 words averaging 5 letters apiece totals 130 letters, considerably exceeding the minimum pangrammatic window in literature (67 letters). Harder variations of this literary challenge: (1) place the letters at the start of each word, (2) keep the alphabet in order. If both restrictions are heeded, it is unlikely that the Gettysburg Address can be paraphrased.