

# The hidden sound-structures of poetry

Stephanie Smolinsky

New York City Technical College, CUNY

E-mail: smolins@earthlink.net

## Introduction

Can phonetics help us understand how poetry creates its emotional and aesthetic effects? We talk of the “poet’s ear” as especially sensitive to speech sounds, or of the “music” of a poem, which must be read aloud for its full effect. We notice that the sounds of a poem may have a remarkable fittingness to its subject, even though that fittingness cannot be traced to obvious poetic techniques, like alliteration or assonance. We recall a line of poetry without having made a conscious effort to memorize it: a mysterious ‘something’ has fixed it in our mind.

I would like to propose that much—maybe all—of poetry is rather more structured phonetically we generally recognize, and that some of its mysteries may be explained via phonetic observation and analysis. We know that one of the primary tools of the poet is *repetition* occurring at every level of language. At the phonetic level, this is seen in alliteration (repetition of a consonant), in assonance (repetition of a vowel), in rhyme (repetition of a syllable’s nucleus plus its coda), in half-rhyme (generally of a different nucleus with the same coda) and in meter (repetition of feet, or of numbers of strong stresses per line). A second, equally fundamental poetic tool is *iconicity*, in which the form of the language, at all levels, is shaped so as to embody the poet’s meaning. The most obvious phonetic form of iconicity is onomatopoeia, in which a sound referred to in the poem is imitated in the type of segment(s) used. Thus, phonetic repetitions can feed iconicity, as in the famous lines of Tennyson’s (from *The Princess*, Part VII) evoking a dreamy summer euphoria:

“The moan of doves in immemorial elms...the murmuring of innumerable bees”

My first hypothesis is that in poetry, below the conscious level of alliteration, assonance and rhyme, certain types of sub-segmental repetition are also present: repetitions of phonetic features, such as bilabial, plosive, high, rounded, and so on. I am by no means the first person to stumble on this idea: Kenneth Burke (1973), in his essay *On Musicality In Verse*, first published in 1938, describes the technique as ‘concealed alliteration.’ Thus, in the Tennyson example, we see not just eight cases of *m*-alliteration, but of three other nasals: two *n*’s and one *N* (the velar nasal), in other words, the nasal resonance of the *m*-effects is supported by other nasal segments. In addition, the labiality of the *m*’s is echoed in the *v* of

*doves* and the *b* of *bees*. Moreover, every segment in the lines is voiced, evoking their referent natural noises. Finally, observing the vertical and horizontal sequence of movement of the tonic vowels: *ow*, *V*, *O*, *E*, *3r*, *uw*, *iy*, we find the lower, darker note of the doves is caught by the mid and mid-low back vowels while the higher note of the bees is captured by the high vowels.

The Tennyson example illustrates my second hypothesis: that repetitions of certain phonetic features at certain points in a poem (clusterings) and absences of other phonetic features (gaps) may serve iconicity by embodying the message of the poem in ways arrived at unconsciously or semi-consciously by the poet and which affect the audience in a similar way. The reader that jibs at my attribution of ‘darkness’ to the low back vowels should be assured that this connection is not new. From Plato onwards, qualities have been attributed to individual speech sounds, and to phonetic or phonological features. There exists a sizeable literature on the phenomenon known as ‘sound symbolism’ or ‘phonetic symbolism’ (see Sapir (1929) and Newman (1933) on brightness and size in vowels, as linked to tongue height and anteriority, or Marchand (1959) on the aptness of consonant manner features to mimic sounds and processes in the real world). Therefore if a poem subjected to a feature analysis is indeed found to display clusterings and gaps, two kinds of matches are possible. The first, we can look for an arguable iconic connection (i.e., one which is not far-fetched, and which truly explains the poem’s effect) between feature distribution and meaning. If such a connection is established, then we can check feature-to-meaning links found in the poem against those linkings of features and qualities described in the literature of phonetic symbolism.

## 2. INSTRUMENTATION AND PROCEDURES

A computational approach seemed obvious, given the quantity of speech sounds and the large number of features and combinations of features in even a short poem. This avoids the fatigue of doing such analysis by hand and minimizes the chances of error. Accordingly, the Pattern-Finder program was built by Mr. Constantine Sokoloff, according to my requirements but with his own creative suggestions, especially regarding the visual displays. It has a web-based interface, and is built using the Java programming language and the Java Server Pages technology. The Pattern-Finder permits a very large number of types of search (single-feature or combined-feature). Combinations may be of two features shared by a

segment (e.g., coronal + fricative) of two features, which may overlap on some segments only (fricative + posterior, with h the overlap segment), or of two features on separate segments (fricative vs. liquid). The program can generate frequencies, alerting its operator to abnormally high or low numbers of a specific feature in a text, before a display is viewed. But its most important function is in distribution display, since an apparently normal frequency might still mask an abnormal distribution (e.g., a feature-clustering at one location only, and no scatter of this particular feature over the rest of the poem).

The piloted poems are short; each display shows the whole (transcribed) text. *Background segments* (those lacking the feature) are in 10-point italic letters; while all *highlit segments* are in non-italic, enlarged (approximately 20-point) type. The highlit segments indicate a single pattern of distribution with characters which are of normal thickness, underlined and square-bracketed; if needed, a second pattern of distribution is contrastingly indicated by adding characters in boldface with no underline or brackets. Our present transcription system is based on the SAMPA alphabet for American English, somewhat modified in order to render transcriptions as transparent as possible. The affricates are represented by single letters, C and j, the low front vowel by A, the low unrounded back vowel by a, the schwa by ^ and the diphthongs of *cow*, *coy* and *kite*, respectively by aw, Oy, and ay. At present, displays are in HTML, and therefore cannot be shown in this paper formatted as described above, but one example of a two-feature ‘versus’ distribution demonstrates full display.

**Figure 1: Pound’s *In a Station of the Metro*: Labial distribution and frequency**

total phonemes 44

total consonants 27

total vowels 17

total labials 7 (26% of consonants)

Diy %A [p]^“rIS=n — ^ [v] %Diyz “ [f]eysIz — In  
D^ “krawd ...

“ [p]Et=lz — an ^ “ [w]Et — “ [b]lAk —  
“ [b]aw

It was decided that for the pilot, input transcriptions for the Pattern-Finder should be of the broad phonemic type. Phonological rules which might contribute to poetic effects (such as tonic vowel lengthening before voiced consonants) could be built in to the program later. The features that were chosen were mostly the traditional articulatory ones, e.g., “fricative” for consonants or “high” and “front” for vowels. The idea here was to try to

represent the ear (i.e., the aural-oral consciousness) of the poet; for this, the groupings produced by the acoustics-based Jacobsen Fant and Halle (1969) distinctive features, such as strident/mellow, seemed counter-intuitive. (For example, their ‘stridency’ feature contrasts the labiodentals with the non-strident linguadentals, whereas all four dental fricatives can be heard as the ‘breathy’ group characterized by lower-frequency frication, contrasting with the six-member ‘sibilant’ group of the coronal fricatives and the affricates, characterized by higher-frequency frication.) So far, “sonorant” has been the only feature included that groups vowels with consonants; this was in order not to bias answers to the question of whether connections could be found in comparing distributions of consonant and vowel features.

Besides breathy/sibilant, some other consonantal features were added. These included fleeting/durative, the criterion being whether a consonant is capable of being prolonged (as are, e.g., the fricatives and nasals) or not (as are, e.g., the stops and glides). The simplifying place features anterior/coronal/posterior were added, and similarly, with the vowels, the simplified vertical place features of upper/middle/lower (making the mid-high and mid-low vowels members both of the middle group and respectively, of the upper and lower groups). Tongue-movement in the free vowels was described in the features limited rise/expanded rise, which distinguished the group iy, ey, uw and ow from that of ay, aw and Oy. However, it should be understood that the Pattern-Finder is a work in progress: if these extra features are not found to capture any interesting distributions, or if other features appear relevant and need to be added, then the feature list will be changed accordingly.

In the transcriptions, a very simplified two-level stress system has been used, to enable four kinds of search: for all stressed vowels regardless of type, for primary-stressed, for secondary-stressed, or for both types, with each type being marked. Likewise, a simplified two-length pause system has been used, to enable four kinds of search: for all pre-pause segments, for pre-major pause, for pre-minor pause, or for both types, with each type being marked.

In this pilot run, data for the Pattern-Finder has been taken from two sources: the control sample is a transcription of three paragraphs from a piece of fairly formal journalistic prose by Nat Hentoff, published in the *Village Voice*. The four experimental samples have been taken from an anthology of modern verse, *Poetry Speaks*, in which many of the written texts are accompanied by recordings of the poets performing their own work. Three of the poems chosen are Langston Hughes’, *The Negro Speaks of Rivers*, Carl Sandburg’s *Cool Tombs* and Wallace Stevens *Bantams in Pine-Woods*. All are short and well-known. For these, transcriptions made from the written texts have been checked against the poets’ own recorded performances. The object was not to produce phonetically exact narrow transcriptions, but to ensure that no

gross errors were made in details of pause and stress, or where alternative pronunciations of a vowel or a word existed, that the poet's pronunciation was followed. The fourth poem is Ezra Pound's haiku-like *In A Station of the Metro*; the poet's recorded voice was available, but not reading this particular work. It has been included because its extreme brevity makes it a particularly good specimen for a short study like this, and because of its beauty and fame.

### 3. FINDINGS

Two factors were focused on: feature frequencies and distributions. Poem frequencies were compared with those of the journalistic prose piece, (see below, Table 1).

**Table 1 Feature/phoneme frequencies: Hentoff, *Village Voice* Feb 5-11 2003**

total phonemes	1187
<b>consonant</b>	769
voiceless	250
voiced	519
obstruent	470
sonorant	717
sonorant consonant	299
anterior	176
coronal	530
posterior	63
fleeting	306
durative	463
stop	252
fricative	200
nasal	130
liquid	133
glide	36
labial	129
<b>vowel</b>	418
front	217
central	96
back	91
upper	164
lower	127
rounded	65
free	142
checked	202
affricate	18
/ə/ (schwa)	70
/=/ (syll. Sonorant)	65

Many of the distributions examined showed neither clusters nor gaps; one example is the even distribution of labials as seen in Figure 1 (Pound's *Station*). But here we note that the frequency of Pound's labials is high

compared to the prose control sample (26% in Pound, compared to 17% in Hentoff). We can see a restricted frequency and distribution of posterior consonants in the Pound piece (2 out of a total of 161: 5%, compared with 8% for the Hentoff posterior consonants). These are located almost at the ends of lines; at the same time, we can see an equally restricted distribution of (stressed) back consonants in a similar location, but with frequencies much lower than the control (Pound 18%, and Hentoff 30%). The rarity and placing of the alliterated back vowels give the poem a near-rhyme scheme; it can be argued that they and the posterior consonants imitate the 'fixing' of the immediate impression of many people in the subway (represented by the more normal high-frequency anterior and coronal consonants, and the front vowels) in the memory, and into a hard black-and-white image.

**Figure 2: Pound's *In a Station of the Metro*: Stressed back vowel and posterior consonant distributions**

*Diy %Ap^“rIS=n — ^v %Diyz “feyslz — In D^ “[k]r**aw**d*  
 ...  
*“pEt=lz — an ^ “wEt — “blA[k] — “**aw***

An interesting distribution is also found with the fricatives (all but one are confined to the upper line, reinforcing the whispering, ghostly, 'apparitional' quality of the initial experience). We also see upper vowels confined to the first line (corresponding, I would argue, to lightness, pallor, insubstantiality), while all but one of the lower vowels is confined to the second line (corresponding to the deepening or fixing of the immediate impression into a black and white memory-image).

Even though the Pound piece is so short, I believe that more meaningful patterning may be observed, but time and space limitations do not permit further discussion here. Regarding the other data, many more tests remain to be run: even short poems do not give up their mysteries easily. Distributions and frequencies were observed Stevens' *Bantams* which definitely suggested an iconic use of its phonetics. For example, the central verse (in which the subject, a small fierce rooster, talks more intimately to the poet, rather than simply threatening him) is characterized by an absence of posterior consonants and by one line (unique in the poem) with uninterrupted voicing. Uninterrupted voicing is also seen in the refrain lines (initial and penultimate) in Hughes' *The Negro Speaks of Rivers*, and the poem's low frequency of voiceless consonants (47 out of 218: 14%, as compared with 33% for the prose control piece). Its subject is the flow of rivers and the poet's sense of continuity, and historical depth: we note that vocal vibration is made by the articulator deepest in the vocal tract and most primary in language use. Thus, continuity of voicing in the refrain (and the relative restriction of its

interruption throughout) can be seen as nicely embodying Hughes' themes.

#### 4. CONCLUSIONS

The present study is a but a faint scratch on the surface of a huge subject. However, the pilot findings already suggest that there are such things as sound-patterns which are not discernible to a normal audience under normal conditions, but which do play a role in communicating poetic meaning. Our instrument, the Pattern-Finder still needs refining, a larger control corpus is required, and many more poems for input data. But the indications are that we are unearthing a phenomenon which is worth further excavation.

#### REFERENCES

- [1] Burke, Kenneth, *The philosophy of literary form*, Berkeley: University of California Press, 1973
- [2] Jakobson, Roman, Fant, C. Gunnar M. and Halle, Morris, *Preliminaries to Speech Analysis: the Distinctive Features and their Correlates*, Cambridge: MIT Press, 1969
- [3] Marchand, Hans, "Phonetic symbolism in English word-formation," *Indogermanische Forschungen* **64**, pp146-68, 256-277, 1957
- [4] Newman, Stanley S., *Further experiments in phonetic symbolism*, *American Journal of Psychology* **45**, pp 53-75, 1933
- [5] Plato, Cratylus, *The dialogues of Plato* Vol.1, Oxford: Clarendon Press, 1892
- [6] Sapir, Edward, A study in phonetic symbolism, *Journal of Experimental Psychology* **12**, pp 225-239, 1929