OED Online's Single-Quotation Entries: an Analysis Giles Goodland

The OED is a dictionary founded on 'historical principles', created as a guide to trace the history of individual words 'with regard to each individual word, when, how, in what shape, and with what signification, it became English; what development of form and meaning it has since received; which of its uses have, over time, become obsolete, and which still survive; what new uses have since arisen, by what processes, and when [etc.]' (Murray. Preface to Volume I of *New English Dictionar)*. For a small proportion of entries in the OED any such history might be seen as problematic, since they only have a single recorded instance. It is, of course, impossible to say with certainty if this instance records the only actual contextual use of a word, or if the word is just *rare* in written or published use.

In this paper, I have isolated all of these single-quotation headwords between the start of R to REZ. This section of the OED has been recently edited, and shows a total of 548 single-quotation entries[1]. 79 of these headwords were upgraded in the course of editing from sublemmas or derivatives sections in OED2. 14 of the headwords are newly-created entries. The remaining 465 are from OED2[2]. This compares to a total of 10443 headwords in the section of the dictionary R to REZ, in other words, the number of single-quotation headwords is less than 4.5% of OED Online headwords: a small proportion, but a significant number in aggregate.

There are fewer single-quotation entries in OED Online than there are in OED2. The 1989 ed. OED2 for the range R-REZ has 1,406 single-quotation entries. Adding to this the four 1933 Supplement entries, makes 1, 410. Disregarding the upgraded and newly created entries on our wordlist of OED online words, makes 456 OED2 single-quotation entries remaining on OED Online. Thus 68% of OED2's single-quotation entries in this range have been contextualised with the addition of further quotations. This contextualisation will have been through an intensive process of searching for examples of these words on a large number of databases, as well as sifting through the OED's own collection of slips, regional and historical dictionaries, and in-house databases. The number of such databases has grown massively over the last two decades. The size and accessibility of databases such as EEBO, Literature Online, Googlebooks and Nexis (to name only a few) has led to a radical change in the availability of contextual examples of word-use.

[1] I am excluding RECIPROCALTY, which is now noted as a 'Compositorial error for *reciprocall tye* in 2nd to 6th editions (1624–51) of R. Burton *Anat. Melancholy'*.

[2] OED2 here refers to the 1989 edition of the OED, which is the full text of the *New English Dictionary* with the addition of the Supplements, apart from some material in the 1933 Supplement. OED Online, fully re-edited, also incorporates the entries from the 1933 *Supplement* that were not previously included.

This paper concentrates on this remaining 32% of OED2 single-quotation entries, plus the upgraded and new items. The most obvious feature that these words all have in common is that they are rare, at least in printed use. Single quotation-entries are interesting in the same way that many rare things are interesting: their uniqueness says something about the conditions under which words are formed, certainly about the way vocabulary enters the dictionary, but also perhaps about how words can enter the language**.**

Having stated what these words have in common, the other notable thing about this wordlist is that the words are remarkably varied**:** ranging in date from 1225 to 1970, with an impressive range of sources and registers, regionalities and areas of discourse, authors and origins. Looking at the spread of dates (Graph One) it is clear that an even spread is combined with a spike in the Early Modern period, and a gradual decrease as we approach recent years. The reason for these features should become clear as the paper progresses.

Also, a high proportion of the words are unusual in origin: this can be seen from the etymologies. 50 entries start their etymology sections with *apparently*, there is one *could alternately*, 15 *eithers*, one *humorous alteration*, one *irregular formation*, five *irregularlies*, 3 *perhapses*, 29 *probablies*, as well as 27 *origin uncertains*, and 15 *origin unknowns*. In the section R to REZ as a whole, there are 58 *origin unknowns*, and 183 *origin uncertains*, or 2.3% of total headwords. This compares to 7.6% for these two etymological descriptions for the single-quotation entries. This reflects the fact that unusual or one-off words are often more hard to analyse. The less evidence that there is, the more need there is conjecture. So for the rest of this paper I am going to ignore the conjectural elements in the entries, and stick as much as is possible to what is definitely recorded about them in their OED online entries.

Apart from the simply odd or uncategorisable words, which include many of the earliest in the list, as well as many of the *origin unknown/uncertain* words, there appear to be four main categories.

The first one is the most obvious, and also the most limited diachronically. These are the words that the *New English Dictionary* took from earlier dictionaries, especially those in the seventeenth century. There are two 1598 *Florio* quotes, and three more from 1611. There are four from *Cotgrave*, seven from *Cockeram*, two from *Blount*, one from *Coles*, five from *Chambers*, and one from a 1657 *Physical Dict*. (See Table 1). These are all a heritage from OED2 (all, in fact, from the *New English Dictionary*), and range in date from 1598-1876 (although there are only four examples after 1723, and these all from specialist lexicons). A high proportion of these words are clearly polysyllabic Latinate constructions.

Although it is not possible to be certain that these words never occurred in contextual use, the number of instances here certainly suggests that these are for the most part 'dictionary words', anglicizations of Latin words with no currency in actual discourse. This is not to say that they are completely without interest. They tell us something about the production and values of early modern dictionaries, in which, one might speculate, lexicographers of the period actually created a proportion of the words that they included in their dictionaries. The entries are also interesting in that they show how the *New English Dictionary* was often reliant on earlier lexicons as a source of vocabulary, even if no corroborating evidence was available.

The second category is that of rare words, where there is some evidence that the word did have a specialist currency at one time, or in one specialist field. These have subject labels attached to them, ranging from *Anat*. to *Zool*. (see Table 2). With these, it can be suspected that they had some currency that may have been chiefly in speech or ephemeral writing rather than in published, retrievable sources; certainly some of the labels suggest a specialism that would have been more used in oral communication, or in restricted, or hard to retrieve types of writing. In other cases, the words are from glossaries, and should perhaps be taken with the same kind of scepticism with which we treat the words in the first category (indeed, a few are from the same sources). It is also of course possible that some of the other OED Online entries not given subject labels may also have had an isolated currency, especially if the entry is of a technical nature: one so specialized that there is no other trace of them in the records available to the OED research team.

Another set of rare words is that of regional varieties; five entries in our sample have an *English regional* label, one has a *U.S.* label, and 24 are Scottish. Here (Table 3) it is more clearly the case that the word was used in speech, and although certainly rare in its written form, may have been a widely recognised term in some localized communities (for some of the Scottish words, the label does not preclude isolated use in a literary or poetic context). A few of these regional words, when not marked *Obs*., may still be in use now. This is quite a low title, and may reflect the fact that in the sample, Latinate words tend to predominate, due especially to the number of words starting with the Latinate prefix RE- (dialect words tend to be non-Latinate in origin). Again, it is also the case that some entries not given a regional label may have had a geographically localized provenance.

The third category is that of nonce words. The term *nonce-word* was itself coined by James Murray in 1884, 'Words apparently employed only *for the nonce*, are, when inserted in the Dictionary, marked *nonce-wd*.' [*New English Dict.* p. x]. This is especially interesting category since it suggests that the user

of the word is consciously coining in order to fit a particular occasion. The key and paradoxical quality to nonce words is that they depend on rarity. Once a word becomes known or recognized, it is called to mind by the user as an existent word, and ceases to be nonce. A nonce word is only nonce so long as the user of the word is consciously coining it at the moment of usage (it can of course be coined independently many times, as long as the coiners are unaware of previous coinages). They are also dependent on particular occasions, since they are formed to avoid a longer periphrasis, or because a new word is the only way that the user can express a particular thing, or for a humorous purpose. For this reason, you would not expect a nonce-word with an obvious synonym, or a one-word paraphrase. This context-dependence also has implications for the morphologies of nonce words, since (unlike some poetic formations) they should have a transparent and easily recognisable derivation. 'Generative' principles of word-formation operate here, since these new words should be instantly decodable: just as with generative grammar, both producer and recipient are able to understand the rules that enable the production of the new word.

The headwords explicitly labelled as 'nonce' in this range comprise 27 words, forming only 4.9 % of the total of entries in the OED Online sample (see Table 4), but it would be easy to make a case that many entries not explicitly marked as ‘nonce’ may still come under the definition of a nonce-word.

Looking at the contexts of some of these words (Table 5), it is easy to see how these usages are context-driven. The need to communicate the circumstances of an unusual state of affairs causes the writer (quite often in a humorous way) to coin the word instead of using some longer phrase. These words can be seen to be morphologically simple (either straightforward and easily interpretable affixation, or zero-derivation) combined with a fairly complex dependence on context, often one in which the same affix is used more than once, or (as with the Wodehouse quotation) the verbal form can be seen from its proximity to the nominative, to be created from it. In several of these, there is a playful and conscious awareness of the possibilities of affixation or of zero-derivation. The complexity or uniqueness of the context may give a clue as to why these words are so rare (how often is someone made an admiral twice, or imputed to have become patriotic twice?). In other cases, a fine distinction seems to be intended, for instance 'reluctantism' may seem a stronger or more emphatic use than its obvious near-synonym, reluctance.

The final and largest category is that of literary coinages, in poetry and verse-drama. Since Murray has already given us the precedent of the consciously-coined nonce-word, I think that I am secure, given the rarity of these items (reckoned against the easy accessibility of specifically literary and poetic texts on various databases), to presuppose that these words are mostly conscious formations. Even though the

formation process is also contextual and conscious, as with nonce words, I would suggest a distinction between these formations and nonce-words, in that the poetic text functions in a different way from the communicative act presupposed in most of the texts in the list of nonce-words.

Poets in this list are given in Table 6. This is only a partial list of the many quotations that come from poetic or verse-dramatic sources in the sample, since I have ignored anonymous and indeterminate texts. The total of entries which are from poetic sources in the sample is roughly 157, or 28.6% of the total. This is clearly a high figure, but it is hard to reckon how much of the OED's source material, on average, is from poetry.

From the authored quotations on the list, it is easy to break down some of the words in terms of constraints and probable or at least partial determinations. An examination of the context of the quotation is often helpful in order to show that the word has been crafted in order to fit a metre or to make a rhyme. The easiest category to isolate is that of rhyme-words. In these examples, the word seems likely to have been coined for the sake of a rhyme. See Table 7.

Deliberately-formed metrical examples are harder to isolate, since a metrical poem is in metre, so *all* the words in the list should fit the metrical pattern, hence, occurrence in a metrical context is not in itself evidence that the word was coined to fit the metre. However, in some cases a more common synonym with a different metrical pattern is clear, and, we assume, would have been the preferred choice if it *did* fit the metre. Table 8 gives the metrical contexts where a common near synonym (often a version of the word that is actually used) has apparently been rejected in favour of the otherwise unrecorded word. I have put the more usual word in square brackets after the quotation, with its varying stress.

The morphology of these words is interesting, since in contrast to the transparent affixation we saw in the case of nonce words, there is also a degree of clipping, elision, merging, and non-transparent derivation. Also, in this list, it is interesting to note that they come later in date than the words formed for the sake of rhyme. As has often been noted, a strict iambic pattern did not become an established feature of English poetry until the late Sixteenth Century. In addition, we can see clear-cut examples of this kind of usage dying out quite rapidly—with only one doubtful nineteenth century example.

Easily paraphrasable words, or those with a common synonym, are clearly not nonce-words, since in a communicative context, there would be no need of them. But also in poetry we find words that are the opposite: there is no clear meaning outside the poetic context. These could be nonsense words of the

Lewis Carroll type, or poeticisms chosen for some euphonious or echoic sound-quality. Such words are also not nonce words, since the aim is not to communicate a contextual meaning outside that of the literary text as a whole.

Words of this type should normally be comprehensible, but one of the qualities of poetic language can be a certain elusiveness. Their sound quality can add for instance to the assonance of a line, while the unfamiliarity of a word may enforce a slow or more careful reading. The consistent use through time shows that poets writing in English were not afraid of the effects of coinage. Estrangement and defamiliarization are later terms for this, but pre-modern writers were certainly aware of the power of strangeness and unfamiliarity, in adding another layer of effect to their texts. Perhaps surprisingly, words in this category (Table 10) often show a similarity to the apparently manufactured 'Dictionary Words' we saw in Table 1. Many of them are either Latinate, or have a polysyllabic Latinate quality to them. Some are metrically replaceable with near-synonyms (such as REGARDIVELY, which might be replaced with 'regardfully'). Others are not easily paraphrasable, hence having something in common with nonce-words, such as RECURB, which requires a 14 word definition ('The curved shape of an oscillating surface produced at the extreme point of oscillation'). Words of this nature do not have to be restricted to poetry, of course. Examples from Nashe and Joyce show that they may also be used in prose.

A special mention should be made of a creative use of the RE- prefix in poetry, in this section of OED Online. RE must have often seemed irresistible in metrical poetry, since as a single syllable which could take either a stressed or unstressed positions, it solves many prosodic problems at a single stroke. (See Table 11). A few poets seem to have discovered this useful shorthand and used it frequently.

The English language does not lack terms for recurrence, so although some of these uses of the RE- prefix have the appearance of nonce words, a word like 'again' or' twice' would have worked just as well in giving the prose meaning. Although a single word could not replace the sense in these examples, it is not, apparently, the uniqueness of the communicative context that determines the coined word, but the chosen sound (especially metrical) pattern.

To say that these words are there for the sake of rhyme or metre of course does not *explain* them, but does help to account for their formation. Particularly in highly crafted language such as poetry, there is unlikely to be a single determinant. In a few cases, of course the same term comes into both the metre and the rhyme lists (showing perhaps a certain inevitability in the choice). The different and more fluid

registers of poetry allow for a wider range of choice than in other discourses. Clearly, this includes the licence to invent words, as the context demands. Many of the words in the list are not easily paraphrasable, and need a fairly lengthy sentence to define them.

In conclusion, nonce words are formed through contexts relating to an immediate need to communicate a thing in a specific context. They are more frequent in less formal discourse, such as letters. They usually have a simple and easily decodable derivation. Literary coinages tend to be more premeditated, and may show determination through metre or rhyme. They may have a more complicated or less easily ascertainable etymology. In poetry and verse drama on the other hand, intelligibility (or its lack) is but one of an array of effects.

Through these single-quotation entries, OED Online traces through time the fact that there is a creative, generative aspect to word-formation that can produce rare new words when an isolated context demands it. Each individual act of creation builds up a diachronic picture or word-creations, each in their own narrow and rarefied context. In particular, words from poetry influenced by a rhyme or metrical context peak in the sixteenth century and decline soon after, Ezra Pound perhaps sealing their fate with his famous attack on the pentameter. Other diachronic aspects of rare lexis is the brief rise and decline of dictionary words. It is clear that there are diachronic trends in conscious word-formation, and OED Online helps to trace what these words are.