

# Sound symbolism in Elvish

Roman Rausch

first presented: *Omentielva Cantea*, Aug. 12th 2011, Valencia, Spain

published online: Jan. 6th 2014

The human phonetic system is a small- ranged instrument (compared with music as it has now become); yet it is an instrument, and a delicate one.

---

A Secret Vice

## Contents

1	Magnitude symbolism	5
2	Distance symbolism	7
3	Closed-set symbolism	9
4	Clustering or abstraction	11
5	Luminosity and colour symbolism	12
6	Temperature symbolism	19
7	Summary	21

## Introduction

The concept of sound symbolism<sup>1</sup> is that of a natural relation between sound and meaning, that sounds may invoke a corresponding idea beside the purely conventional meaning of a word. It is related to the well-known phenomenon of onomatopoeia, i.e. the imitation of the sounds of animals, like *cuckoo*; or sounds of any other kind, as *click*, *ding-dong* or *boom*. Much more interesting, however, and the topic of this investigation, is the imitation of non-acoustical properties, crossing over into the realm of synaesthesia.

Japanese is a language famous for such synaesthetic words which are integrated into standard speech and fill hundreds of pages in dictionaries. Apart from *phonomimes* (onomatopoeic words) there also are *phenomimes* – sound-symbolic words for any perception other than audial, and *psychomimes* – sound-symbolic words describing emotional states<sup>2</sup>. For example, *kirakira* describes something sparkling, voicing the consonants to *giragira* makes it more intense, 'dazzling'. *Kosokoso* describes something done stealthily and secretly, *nurunuru* something slippery, *odoodo* a very nervous state of mind, *shonbori* a despondent, downhearted state. *Potapota* is the dripping of a small amount of liquid, *botabota* of a large amount.

It has been noted that the sounds in these words are not completely random, but that there are certain tendencies. For example, voicing of a consonant increases the intensity as well as the size and bulkiness of the object involved, /s/ tends to describe quiet states and motions, /r/ something fluid or smooth, /o/ something negative with respect to psychology ([1], pp. 50-56); as seen in the above examples. Could such correlations exist in ordinary words as well, and in languages which otherwise do not make use of explicit symbolic words in the same way?

---

<sup>1</sup>also *phonetic symbolism*, *phonosymbolism* or *iconism*

<sup>2</sup>altogether also called *ideophones*

Dabblings into this subject go a long way back: Plato's *Cratylus*, for example, is a discussion between Socrates and Hermogenes about the nature of personal names and words in general. Socrates explains dozens of words by comparing them with others of similar sound. For example, he proposes that the gods are called θεοί because the ancient Greeks before him, like the barbarians of his time, once worshipped the sun and moon which run across the sky, so that θεοί can be associated with θέοντες 'runners'. Such an approach, while on the right track in principle, is of course only pseudo-linguistic since it lacks any systematic, scientific methodology. In any case, having explained some words in terms of others, the question arises how the basic, not further associable words were created. And here Socrates turns to sound symbolism. *Cratylus* is often mentioned in the context of this topic, but a detailed account is rarely given. Socrates mentions the following ([14], 426c-427c):

- ρ — motion (ῥεῖν 'to flow', ῥυμβεῖν 'to whirl')
- ι — something subtle (lit. τὰ λεπτά) which can easily penetrate things, hence motion ἰέναι 'to go')
- φ, ψ, σ, ζ — ?shaking (ψυχρόν 'cold, shivering', ζέον 'boiling, seething', σείεσθαι 'to shake')
- δ, τ — binding, rest in place (δεσμός 'bond', στάσις 'standing')
- λ — smoothness (ὀλισθάναιν 'to glide', λιπαρόν 'sleek')
- γλ — sticky substance (γλυκύ 'sweet', γλοιῶδες 'gluey')
- ν — inwardness (ἔνδον 'inside')
- ο — roundness (γόγγυλον 'round')

Similar ideas were expressed by Wilhelm von Humboldt in 1836, who wrote (among other things) that the German consonant *w* [v] in the words *wehen* 'to blow', *Wind* 'wind', *wirren* 'to move chaotically' and *Wunsch* 'wish' symbolizes a vacillating, wavering motion ("schwankende, unruhige Bewegung"; [8] p. 79). However, it soon becomes very clear that one can get carried away by such sentiments. If matching criteria are sufficiently weak, many words will be found to follow them just by coincidence. In fact, Humboldt remarks that generalizing such notions to a system of principles is a 'slippery path' ("einen [...] schlüpfrigen Pfad"; [8] p. 80).

There could be many German words describing vacillating motion which do not start with *w*-. On the other hand, many words which start with *w*- do not describe any such idea. So the foremost question actually is: Does sound symbolism exist at all, beyond subjective and romantic impressions? Clearly, statistics has to be included into considerations somehow in order to make the path treadable again.

In the course of the 20th century a number of psychological experiments were carried out aiming to uncover correlations between sound and meaning. In 1929, Edward Sapir [2] offered a group of participants invented word pairs differing by one vowel only, like *mīl* and *mal*, said to mean 'table'; or *glupa* and *glopa*, said to mean 'horse'. But one word of the pair was said to mean 'large table', the other one 'small table' (or 'horse'); and the participants were left to decide which was which. Sapir then gathered data about how large the particular vowels were perceived, which was later statistically evaluated by Newman who also performed a series of follow-up experiments [3]. It turns out that the subjects were very consistent in their judgements of the magnitude, in fact so consistent that Newman was able to assign a neat scale to the vowels. His results can be presented in a condensed form as follows:

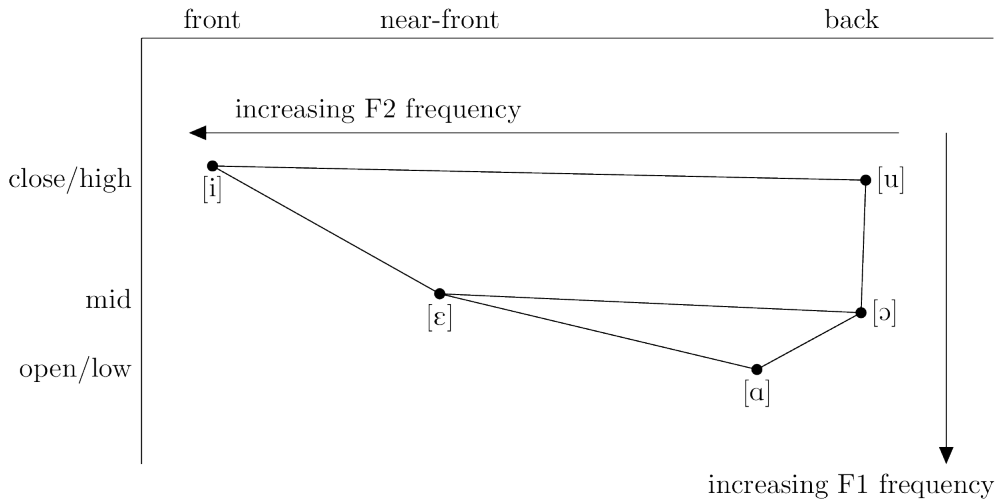
$$[i] \ll [e] < [\varepsilon] \ll [\text{æ}] < [u] \approx [a] < [o] \approx [ɔ]$$

This means that [i] was consistently judged to signify small size much more often than any other vowel. For example, [i] < [e] was decided by 70%, [i] < [u] by 83% of the participants on the average. On the other hand, [a] > [u] was decided by only 52.5%, making these vowels approximately even in size, so to speak.

Such consistency indicates that there must be a strong reason behind these judgements. To understand it, let us look at the acoustic spectrum (the amplitude over the frequency) of vowels. There, one observes several characteristic peaks, called *formants*. The first two formants F1 and F2 determine the quality of the vowel, the others characterize the speaker's voice. The frequencies of the five principal vowels, as measured from recordings of American English speakers, are approximately as follows ([7], p. 172):

- F1: [i] ≈ [u] 0.3 kHz < [ε] ≈ [ɔ] 0.6 kHz < [a] 0.7 kHz
- F2: [i] 2.3 kHz ≫ [ε] 1.8 kHz ≫ [a] 1.1 kHz > [ɔ] 0.9 kHz ≈ [u] 0.9 kHz

The following diagram is drawn proportionally to the above values:



As can be seen, the scale of F1 reflects vowel closeness (close vowels have the smaller frequency) while the scale of F2 reflects vowel frontness (front vowels have the higher frequency). We can also observe that the vowels [i] and [ɛ] are clearly separated from [ɑ], [ɔ], [u] on the F2 scale, and that the variation in F1 frequency is of a much smaller width ( $\sim 0.4$  kHz) than the variation of F2 ( $\sim 1.4$  kHz).

If one compares the F2 frequency scale to Newman's scale above, one sees that they are in almost perfect agreement: A low-frequency vowel is associated with large size, a high-frequency vowel with small size. This is probably based on everyday observations that smaller objects produce more high-pitched sounds when they fall, and the voices of smaller animals (as well as human beings) are of a higher pitch as well. The F2 scale seems much more suited for this mapping than the F1 scale because of its greater width. One thus could say that the order

$$i < [ɪ] < [e] < [ɛ] < [æ] < [a] < [ɑ] < [ɔ] \approx [o] \approx [u]$$

forms a natural 'musical scale of speech'.

This leads to the question of how these vowels are actually represented in the words for 'big' and 'small' across languages. These English words themselves disobey the determined tendency (one would expect [i] < [ɔ:]). On the other hand, 'little' and 'large' fit the pattern, as well as the emphatic long [i:] in 'teeny' and the diminutive suffix *-y*. Greek  $\mu\upsilon\kappa\rho\acute{\omicron}\varsigma/\mu\alpha\kappa\rho\acute{\omicron}\varsigma$  would be a perfect example in agreement with the above scale where only one vowel is changed (/i/ < /a/). Latin *parvus/magnus*, on the other hand, seem neutral, having the same vowel (although the comparative and superlative *minor, minimus* again show /i/), but in Romance languages this is altered to French *petit/grand*, Spanish *pequeño/grande*, once more in agreement with the scale. Japanese shows *chiisai/ookii* where the long vowels /i:/ and /o:/ of the initial syllables are also in agreement with the pattern. Newman evaluated words from Roget's Thesaurus under the categories 'largeness' and 'greatness', with independent judges striking out words which in their opinion did not belong into these categories. Using the previously determined scale of magnitude he calculated the 'average size' of both groups of words. The result was that words from the largeness category had an average magnitude of 0.98, words from the smallness category of 0.92. The difference was so small as to suggest that there is no sound symbolism in these English words – the vowel distribution is what one would expect from two randomly selected lists of words.

However, it must be objected that Newman's lists were too long. The largeness category contained words like *dimension, terrible, world, elephant, perfect*, the smallness category words like *microbe, puppet, embryo*, and, ironically enough, even *elf*. While they might indeed activate a sense of largeness or smallness, they also activate a lot of other ideas. 'Microbe', for example, may be associated with 'dirty', 'infection', 'life', 'parasite'. There is no universal rule that a language has to classify a microbe under the category SMALL prior to any others.

A more careful key study across language families was carried out by Russell Ultan in 1978 [20]. He examined 136 languages (however, with a bias towards American Indian, due to a better availability of descriptive grammars) regarding patterned sound symbolism, in particular looking for any systematic usage of sounds for diminutives and augmentatives. Unfortunately, he is not too specific about the choosing criteria, but judging from the given examples a 'patterned sound symbolism' seems to imply a minimal variation in word pairs, ideally just one vowel or consonant. His conclusion was that only 37 languages made use of such a symbolism with respect to size at all. The commonest technique turned out to be an alteration of consonants, e.g. palatalization, which was however limited to American Indian and thus appeared due to weighting in their favour. The next most common method was vowel ablaut, and the pattern observed here was a contrast of high front vowels for the diminutive versus low back vowels (in other words, /i/ vs. /a/). Very rarely, a high back versus low back

contrast (/u/ vs. /a/) was observed. Some languages were also observed to use high tone for the diminutive, confirming the significance of frequency, which is naturally higher when the pitch gets raised. Ultan's conclusion was included into the Universals Archive [16] (a database which gathers proposed universal features in the world's languages) in the entry 1001: "Front vowels predominantly correspond to diminutive and associated categories".

Two languages are reported to contradict this tendency: Korean and Bahnar (the latter spoken in Vietnam). They have ideophonic vocabulary very similar to Japanese described above, and use a very consistently structured vowel ablaut to alter the meaning. But both, according to the authors of the papers, use /i/ for the LARGE category and /a/ for the SMALL category [18, 19].

However, I must note that the words used to demonstrate the contradiction are not strictly diminutive or augmentative, but always additionally connected with intensity, e.g. Korean

- [piŋkʷɔl] '(turn) round and round',
- [pɛŋkʷɔl] 'round and round (the circle involved is smaller and the movement faster)'.

So while the circle gets smaller, the intensity of the action increases which complicates things beyond a single scale. While one may associate greater intensity with a greater size, one could also take the universal MORE IS UP metaphor as basis: One speaks of 'numbers going up', 'prices being high', something 'going through the roof' and so on. It is a recurring pattern in the world's languages and is probably derived from basic observations like the water level rising in a cup when poured in. So with regard to amount and intensity, the mapping from meaning to sounds could go either way and still make sense: 'intense' = 'being high' = 'producing a high tone' (hence symbolized by a high or front vowel), or 'intense' = 'being big in size' = 'producing a deep tone' (hence symbolized by a low or back vowel). In my opinion many of the investigations of sound symbolism largely disregard such possibilities of metaphorical reasoning behind the mapping, and generally seem to put 'size' into the same category with 'importance', 'power', 'amount' or 'intensity'.

No more than about a third of the languages investigated by Ultan show patterned sound symbolism in the first place. It has probably largely to do with diachronic sound changes affecting vowels and consonants in a regular fashion. These changes are clearly a much more dominant force than sound symbolism. However, as noted by Otto Jespersen, sound symbolism may also have its share of influence ([4], pp. 406-409).

For example, sound-symbolic words may resist changes, as English *cuckoo*, the first vowel of which remains [u], resisting the change to [ʌ] as in *cut*. A drift of meaning may appear due to sound-symbolic influence, so that the word *miniature* becomes a 'small picture' rather than 'image painted with minium (red lead)' which was its original meaning. It stands to reason that if a word sounds right by sound-symbolic criteria, it will be more stable, i.e. more unlikely to be replaced by another. In a similar way, foreign words are more likely to be loaned if they have a suggestive sound, e.g. *jungle* from Hindi *janḡal* is easily acceptable as an English word due to its association with *jumble*, *tumble*, *bundle*, *bungle*, according to Jespersen. Finally, sound-symbolic words are constantly reinvented, like *teeny*, *teensy* as a modification of *tiny*.

And this brings us to constructed languages, part of which is invented vocabulary. How does one invent vocabulary? Human beings are notoriously bad at doing something at random, they are always influenced by something, even if unconsciously. I believe therefore that sound symbolism plays a much larger role in constructed languages, and should in fact reflect the preferences of their creators, unless being consciously avoided.

Making a statistical statement is normally impossible for a single natural language because it will usually show no more than a couple of words for one conception. In the case of Tolkien's languages, however, we have their extrafictional development to help us out. A statistical statement will therefore be made by listing *all* roots for a conception from Goldogrin, Noldorin, Sindarin, Qenya and Quenya.

A consideration of roots rather than individual words is motivated by the following observation: The onomatopoeic word **buzbō** 'large fly' becomes **puspo**, **pupso** in Quenya (PE19:101) and loses its suggestive buzzing in the process; **kwāenē** 'small gull, petrel', becomes **pāne** in Telerin (VT45:24) by regular changes and loses its suggestive wailing. We can expect the same to happen to sound-symbolic words, although the opposite case, where a root is reverse-engineered to fit a symbolic meaning, cannot be always ruled out.

However, if one compares root vowels, the early sources pose a difficulty, as they frequently contain syllabic sonorants, spelled **Ṛ**, **Ṙ**, **Ṝ**, and even syllabic **Ṣ**. In these specific cases it stands to reason to consider the derivatives instead. Note also that the roots given in the Qenya Lexicon do not quite reflect Primitive Elvish, but already contain sound changes of Qenya, so that it is sometimes necessary to compare them with the Goldogrin derivatives.

I should also remark that a general distinction between broad, phonemic transcriptions; and narrow, phonetic transcriptions will be made. For the former, I use slanted brackets /./., for the latter square brackets [..]. The

realization of the vowel phoneme /a/, for example, might be [a] or [ɑ], and might also vary among the daughter languages, but /a/ can be conveniently used if this difference is of no particular importance to the point being made.

## 1 Magnitude symbolism

Let us first turn to magnitude symbolism, as described in the introduction. As mentioned, Sapir's and Newman's result can be presented in the following fashion:

- [i] ≪ [e] < [ɛ] ≪ [æ] < [u] ≈ [ɑ] < [o] ≈ [ɔ]

Conducting the same experiment with varying consonants, Newman found the following scale:

- [t] ≪ [ʃ] ≈ [p] < [n] < [d] ≈ [f] < [s] ≪ [h] < [k] < [tʃ] < [θ] ≪ [b] < [l] ≪ [v] < [g] < [ʒ] < [m] < [ð] < [ʤ] < [r] < [z]

From here, one can extract the following tendencies:

- dentals < labials < velars,

so that, for example, [k] feels larger than [p], which in its turn feels larger than [t]. Also:

- voiceless < voiced,

so that, for example, [b] feels larger than [p], and [g] larger than [k].

Turning to Elvish, one finds the following roots and derivatives for 'small', 'little', 'tiny':

- INI- 'small': Q. **inya** 'tiny' (QL:42), G. **inig, ineg, inc** 'small' (GL:51)
- \*LIK-: N. **ligen** 'small' (PE13:125)
- \*MIKI-: G. **migin** 'little' (GL:57)
- MINI-: Q. **minwa** 'small', **min (d)** 'a bit, piece' (QL:61), N. **minedh** 'thin, fine, small' (PE13:164)
- MIT- 'small': Q. **mitsa, minte**, N. **mint, mithren** (VT45:35)
- MIW-: N. **mîw**, \*Q. **míwa** 'small, tiny, frail' (VT45:35)
- NIK-: Q. **níka, ninki** 'small (good senses)' (VT48:18), #**nitya** 'little' (PM:365)
- NIP-: Q. **nípa, nimpi** 'small (usually with connotation of weakness)', S. **nimp** 'small & frail, or small', **niben** 'petty' (VT48:18)
- TIT-: Q. **titta**, N. **tithen** 'little, tiny' (LR:394)
- PEYE-: Q. **pia, pikina, pinke, pitya** 'little' (PE17:115)
- PIK-: N. **pigen** 'tiny' (LR:382), Q. #**pitya** 'little' (PM:353)
- PIKI-, PINI-, PĪ-: Q. **pin, pink** 'a little thing, mite', **pīnea** 'small' (QL:73), **pinilya** 'small' (MC:220), G. **pinig** 'tiny, little', **pī** 'anything very small, a bit, mote' (GL:64)

One can find the following diminutive endings:

- S. **-in, -ed** in **noegin** 'little dwarf', **noged** 'Petty-dwarf' (PE17:45-46)
- G. **-chi, chin(t)** (GL:25)
- G. **-inc** in **iolinc** 'little lamb, lambkin' (**iol** 'lamb') (GL:51)

- **-inkī -ikki, -iksi, -si, -ensi, -ki** (VT48:18)
- G. **-ing, -(r/l)in, -lim** in **Ilfing, Ilfin** 'little one, little heart' (**ilf** 'heart (espec. used of feelings)') (GL:50) and deleted **elfrin, elfelin, elfim** 'little heart' and **Elfriniol** 'Littleheart' (GL:32)
- Q. **-īne, -īnea** (QL:42)
- Q. **-īne, -(w)inte, -linte** (PE14:49)
- Q. **-itse** in **melitse** 'darling' (QL:60)
- Q. **-ke** < **\*-ki** from **kilinke** 'a small bell' (**kilin** 'bell')<sup>3</sup> (QL:46)
- Q. **-le** < **-li** in **telelle (-ī)** 'little elf' (QL:91), **qimelle** 'little woman' (QL:77)
- Q. **-li-tse** (perhaps a combination of two suffixes) in **inwilitse** 'little fairy' (QL:42), **miqilitse** 'a little or tender kiss' (**miqe (e)** 'a kiss') (QL:61)
- N. **-ol** < **\*-l** in "diminutive form" **naugol (naugl-)** (Etym:NAUK-)
- Q. **-pi, -pit, -pin, -pinke** (QL:73)
- G. **#-(th)li** in **rogli, rothli** 'little pipe, especially musical' (**rod** 'tube, stem', **ross** 'pipe, especially musical') (GL:65), **tilithli** 'a little elf' (**Tilith** 'an elf [...]') (GL:70), **puthli** 'little baby' (**pui** 'child') (GL:64)

Conversely, the roots for 'large', 'big', 'great' (only considering size, not 'great of power') are as follows:

- **ÁLAT-**: 'large, great in size': Q. **alta, alat-** 'large, great in size' (LR:348)
- **BEL-, MBEL-**: Q. **melek-**, S. **beleg** 'large, great' (PE17:115)
- **FABA-**: Q. **fauka** 'fat, large' (QL:37)
- **\*GWENE-**: G. **gwen** 'large, big, fine' **gwent, gwenn** 'big, large — fine' (GL:45)
- **KHAWA-**: Q. **hoa** 'large, big', **haura** 'huge' (PE17:115)
- **MBOL-, BOL-**: Q. **molda** 'big, large' (PE17:115)
- **\*POL-**: Q. **pol** 'large, big (strong)', **polda** 'big' (PE17:115)
- **\*UM-**: Q. **úmea** 'large' (QL:97), G. **úmi** 'large' (GL:75)
- **UR-**: Q. **úra** 'wide, large, great' (LR:396)
- **YNT-(D)YNTN-**: Q. **'yanta** 'large' (QL:106), G. **gant** '(lit. increased.) larger, greater, more' **ganthan** '(large) waxen, grown big, grown up, adult' (GL:37)

And there is also a number of augmentative suffixes, some derived from the LARGE roots above:

- S., Q. **-on** (PE17:27)
- G., S. **-os** in **doros** 'throne' (**dorn** 'seat') (PE11:29, PE17:151)
- S. **-oth** in **nogoth** 'big dwarf' (PE17:45)
- Q. **-ūme** (QL:97, PE14:48) (see **\*UM-** above)
- G. **-wen** (see **\*GWENE-** above)
- Q. **-(y)ando, -(v)ando** (PE14:48)

---

<sup>3</sup>In this specific word, sound-symbolism coincides with onomatopoeia; cf. the counterpart **kalonga** 'a large bell' (ibid.)

The SMALL roots show a remarkable consistency: All but one contain the vowel /i/, while the LARGE words never do. The only exception PEYE- still yields /i/ in Qenya. The LARGE roots have a higher vowel diversity, but the back vowels /a/, /o/ and /u/ appear most often. The vowel /e/ comes next to /i/ in terms of F2 frequency and we might expect it to be in the SMALL category, but it does appear in LARGE roots. However, LARGE roots containing /e/ also contain voiced stops (\*GWEDE-, (M)BEL-) which could, in a sense, make up for the vowel. Unfortunately no real measure exists as to how various opposite tendencies would affect the overall perception of a word.

With regard to consonants, one observes a strong preference towards voiceless stops [p, t, k] (never voiced), nasals [m, n] and in some cases the approximants [w, j]. Interestingly, [k] never appears as the initial consonant which might agree with Newman's result that velars are felt to be the 'largest'; and also the positioning of [t, p] on the lower end of the scale. On the other hand, according to Newman's result, [ʃ] is also on the far lower side of the scale, yet [s], the most closely corresponding sound in Elvish, never appears in the above roots.

Note also that the behaviour of nasals is much less predictable: In Newman's data, [n] is slightly 'smaller' than [d], while [m] is 'larger' than [b]. In the Elvish roots, the nasals do not seem to follow any preference and appear in both groups of words, although usually as intensifying prefixes in the LARGE roots (MBEL-, MBOL-).

The association of the vowel /i/ does not seem to be limited just to small size, but seems to include frailness and weakness. Consider the roots (N)DIP/B- and (N)DUP/B-, said to refer to bending and drooping (PE17:168)<sup>4</sup>. While (N)DUP/B- yields general notions like Q. **lúva** 'bend, bow, curve' and **numba** 'bent, humped', (N)DIP/B- yields **nimpa** 'drooping, ailing' and **limpa** 'frail, slender and drooping'. This association may account for the presence of /i/ in such roots as SLIN-, CE **slindi** 'fine, delicate', SLIW- 'sickly' (LR:386), NIN-DI- 'fragile, thin', or MIL-, Q. **milya** 'soft, gentle, weak' (VT45:34).

On the other hand, it seems to be present in roots like LINI-, Q. **linda** 'gentle, kind — soft' (QL:54) or PHIM-, S. **fim**, Q. **fimbi-** 'slim, slender' (PE17:23) with positive connotations of 'gentleness' and 'tenderness'. In this sense it seems to be associated with the feminine gender, as in **fimbrethil** 'slender princess, birch' and the roots QIMI-, (QL:77), NIS- (LR:378), INI- (LR:361) 'female'. In fact, Ulta already remarks that there tends to be a diminutive-feminine category in languages, i.e. both tend to be marked in the same way ([20], p. 528).

If one looks at how the vowels are correlated with quantity, then we find a consistent usage of /i/ for high amounts, as in the Qenya plural **-i**, partitive plural **-li**, or the roots RIM-, LI- 'many' (LR:383,369). This runs against Ulta's observation of a correlation between singular and diminutive ([20], p. 547), but may be based on the MORE IS UP metaphor, as mentioned above.

On the other hand, N. **odog** 'much, abundant', **mig**, **tidig**, **migligen** 'little (of amount)' (PE13:125) as well as Q. **olya** 'much', **mitya** 'little (in quantity)' (PE14:48,80) do use front /i/ for small amount in contrast with back /o/. Looking at the various singulative endings in Noldorin and Sindarin, one finds that they are sometimes identical to diminutive endings, although the back vowel /o/ is used just as often:

- **-ig, -eg**: **lhewig, gwanunig, lotheg** (LR:368, WJ:367, VT42:18); **-ig, -il, -el, -iel** (PE13:124)
- **-od, -og**: **filigod, lothog, glamog** (LR:381, VT45:29, WJ:390); **-as, -on, -og** (PE13:124)
- **-l, -ion** (PE13:124)

So it would appear that the two metaphors MORE IS UP and MORE IS BIG are engaged in a tight competition in Elvish.

## 2 Distance symbolism

The basic contrast between front and back vowels is not limited to magnitude symbolism. Another fundamental scale is the one of distance from the speaker (or more properly, the deictic center), usually manifesting itself in two forms of demonstrative pronouns: a proximal one ('this/here', position near the speaker) and a distal one ('that/there', position away from the speaker). A significant number of languages has a three-way contrast which is often speaker-listener-oriented, meaning that the second form means 'near the listener' and the third one 'away from both the speaker and the listener'. Sometimes, there is just a three-way contrast into distance. Having no contrast at all, or a four- or more-way contrast is very rare [17]. Following [6], I shall call the three categories THIS, THAT and THAT YONDER.

In his work, Ulta has also investigated demonstratives. It turns out that one third of the languages in his sample shows a patterned distance symbolism and that the predominant pattern is the one of front or high vowels corresponding to the proximal form. The result is in fact much clearer than in the case of magnitude

<sup>4</sup>The latter of the pair is only given as NDUP/B-, but **lúva** requires a D-variant.

symbolism ([20], p. 546).

Hartmut Traummüller [5] investigated a sample of 37 languages, chosen to be unrelated, and found that /i/ and /e/ appear to 92% and 72% respectively in the proximal forms, while /a/, /o/ and /u/ appear to 74%, 88% and 75% respectively in the distal forms.

Nancy Woodworth [6] undertook a very careful study with regard to the involved patterns. She investigated 26 languages (again chosen to be unrelated) with respect to the pronouns 'this/that' and the adverbs 'here/there'. Based on the previous results, a hypothesis can be formulated that proximal forms are expressed by vowels of a higher F2 frequency compared to the distal forms. If the words had one differing vowel, they were assigned as either confirming or contradicting the hypothesis. If the distinction was made by different consonants, while vowel remained the same, the language was termed 'neutral' with respect to the hypothesis. If several vowels were different, the language was termed 'not clear'.

English (although not considered in this study) would be a language conforming to the hypothesis: *this* [i] < *that* [æ], *here* [i] < *there* [ɛ], as would be French and German if one counts the adverbial demonstratives *-ci* < *-là* and *hier* < *da/dort*<sup>5</sup>, Tongan (an Austronesian language) *ni* 'this' < *na* 'that' or Kannada (spoken in Eastern India) *illi* 'here' < *alli* 'there'.

The result with respect to pronouns was that 13 out of the 26 languages supported the hypothesis, 2 opposed it, and 8 were not clear. Of the 22 languages with known place adverbs 9 supported the hypothesis, 1 opposed it, 3 were neutral and 9 were not clear. Overall, the hypothesis can be regarded as confirmed, as it is supported by 13 out of 15, and 9 out of 10 clear cases respectively.

This has been included into the Universals Archive, entry 1000: "Vowel ablaut is the commonest iconic means (especially for expressing distance)". Apart from the obvious relation to measurement of distance (small distance = small size), it has been suggested as a rationalization that faraway objects tend to be larger than nearby objects [5]. Overall, distance symbolism is backed up by the best evidence with regard to what is actually employed as a contrastive pattern within languages.

Let us now turn to Elvish. We find that demonstratives and place adverbs are always derived from the same roots, so no distinguishing is required. Looking at Goldogrin and Early Qenya at first, we find the following roots in the THIS category:

- SI(N)-: **sitha** 'this', **sith** 'hither' (GL:67)
- KI- 'this by me' (QL:46): G. **cî** 'here' (GL:26)
- I- 'here it is': Q. **ike** 'this' (QL:41)
- HYA- 'this by us': Q. **hya** 'this by us', Q. **hyâ** (QL:41)
- TA-, TAMA-: Q. **tāma** 'this' (QL:87)

One root can be assigned to the THAT category:

- E-: Q. **en-**, **ek-**, **et-**, **ena**, **en** 'that by you' (QL:34), G. **en** 'that by you, that already mentioned (by you), that past', **entha** 'there by you, thither' (GL:32)

One root is in the THAT YONDER category:

- \*HAY-: G. **hai** 'there (by him)' (GL:47)

And finally, one root seems to be distance-neutral:

- SA-: Q. **san** 'then, at that time', **sa-** = **e-**, **en-**, **ta** (QL:81)

It probably refers back in time, as seen in the poem *Earendel*: **San ninqeruvisse lútier...** 'Then upon a white horse sailed...' (PE16:100, MC:216), but is at the same time equated with both TA- and E-.

We see that Goldogrin follows the frequency scale with **cî** /i/ < **entha** /e/<sup>6</sup> < **hai** /a/. In Qenya, however, more roots appear which do not quite fit, in particular TA- in the THIS category.

---

<sup>5</sup>The demonstrative pronouns Fr. *ce* and Ger. *dieser* are actually distance-neutral. The contrast between *dieser* and *jener* is on the verge of becoming obsolete in German.

<sup>6</sup>It seems reasonable to use the stressed initial vowel /e/ in **entha** for comparison, rather than the final vowel.



Next we look at the *Early Qenya Grammar* where the following demonstrative roots with a number of derivatives appear (PE14:54-55):

- **qi** (**si-**, **ni-**) 'this'
- **tye-** (**ke**) 'that by you'
- **sa** (**ta**) 'that (there)'
- **en-**: distance-neutral

They are now perfectly following the frequency scale /i/ < /e/ < /a/. Note also that TA- was moved from THIS to THAT YONDER. It will always remain in THAT or THAT YONDER in later conceptions. The roots **tye/ke** and **sa/ta** may also be used with respect to time – the former of something in the mind of the listener, the latter of something generally in the past.

In later writings, the most frequently encountered contrast is between the two roots SI- 'this' and TA- 'that', again agreeing with the frequency scale:

- SI- 'this, here, now' (LR:385), 'this (by me)' (VT49:18): Q. **sina** 'this' (UT:305), **sīs**, **sisse** 'here' (VT49:18) etc.
- TA- 'that': Q. **ta** 'that, it' (LR:389), Q. **tande**, **tanna** 'thither', **tanya** 'that' (PE16:90,96), **tās**, **tasse** 'there' (VT49:11) etc.

But there is also the root SA- which only refers back to the past. As the details are still unpublished, it is not clear how it contrasts with other demonstratives:

- SA- 'a time already discussed or under discussion', 'a place already discussed': Q. **sana** '\*that one mentioned' (PE16:96-97)

Furthermore, two roots from the THAT YONDER category are found in *The Etymologies* and once more refer to the past and future:

- EN- 'over there, yonder' [future]: Q. **enta** 'that yonder' (LR:356), S. **ennas** 'there' (SD:129-131)
- YA- 'there, over there; of time, ago' [past]: Q. **yana** 'that (the former)' (LR:399)

The presence of the near-front vowel /e/ runs against the pattern.

Overall, the Eldarin languages seem to exhibit distance symbolism in the same way as natural ones, although often with restrictions. At least one can point out that /i/ never appears anywhere else than in the THIS category, while /a/ and /e/ are always in the THAT or THAT YONDER category with the exception of TA- in Early Qenya. It would appear that Tolkien's need to create demonstratives referring specifically to time had a somewhat disrupting effect on the pattern.

Interestingly, the contrast in Elvish is always between high front and low back, as also is most common in natural languages, rather than simply front versus back. The reason for this might be sonority – low vowels have a higher sonority, meaning that they are pronounced somewhat louder which may reinforce their iconic magnitude.

### 3 Closed-set symbolism

Personal pronouns in the first and second person singular exhibit a common trait across Eurasia: The first person very often has the labial nasal [m], the second person a dental sound like [t], [d], [ʃ], [tʃ] or [s]. In Indo-European they are *me* and *tū* (with variations), in Finnish *minä* and *sinä*, in Georgian *me* and *šen*, in Yukaghir (spoken in northeast Siberia) *met* and *tet* [17].

Trautmüller has suggested the following sound-symbolic rationalization for this phenomenon: When producing a nasal, the nasal cavity is used as a resonance chamber, creating the impression that the sound remains within one's own head. <sup>7</sup> In producing sounds where a puff of air is ejected from one's mouth, the speaker has the

---

<sup>7</sup>I might also add that it could account for the 'mhm' interjection to imitate the process of thinking, that /n/ is observed to have a personal and subjective association in Japanese, and that the 'inwardness' of *v* was already pointed out by Socrates in *Cratylus*.

impression to project something from himself towards another person. This sound-symbolic association could have increased the fitness of survival of the M-T pronouns. Thus the languages which share them might be ultimately related, but this relation is probably so ancient that it cannot be recovered by the comparative method.

However, M-T pronouns are only limited to Eurasia. In the American Indian languages another pattern is dominant where [n] appears in the first, and [m] in the second person singular [17]. This seems to disprove a universal tendency of having an M-T contrast, although there is still a nasal in the first person form.

Johanna Nichols proposed a different, and a more sophisticated explanation: She noticed that there is the same kind of M-T (or M-P) contrast in the words 'mama' and 'papa' (or 'mom' and 'dad') across the world. For example, in Basque they are *amma* and *aita*, in Quechua *mama* and *tayta*, in Swahili *mama* and *baba* and so on. She has proposed the term *closed-set symbolism* meaning that pairs of closely related words will show some kind of assonance, like rhyme or alliteration, with small variation. This variation tends to be a change in place of articulation (dental versus labial) or manner of articulation (nasal versus stop) or both. Personal pronouns in the first and second person seem to be such a closed sets of words, just like the words 'mama' and 'papa'. Alliteration, for example, appears in Avar 1st sg. *di*, 2nd sg. *du-*, while rhyme appears in Finnish (see above). Notable is that such patterns often cannot be traced to a single ancestral pattern, but seem to arise independently on their own, so that there appears to be some kind of a universal tendency after all.

In Elvish we find the following:

- Goldogrin 1st sg. **ni-**, **im** (GL:53,97), 2nd sg. {**fi-**}, **thath** (dissim. **thas**) (GL:13,35)
- Early Qenya 1st sg. **ni-**, 2nd sg. **ke-** (PE14:52)
- CE 1st sg. **NI-**, 2nd sg. **KI-** (VT49:50), **KE-** (VT48:25)
- S. 1st sg. **-n**, 2nd sg. **-g** (PE17:132)
- S. 1st sg. **im**, 2nd ?sg. **ech** (PE17:46)
- Q. 1st sg. **-n**, 2nd sg. **-tye** (VT49:16)
- Q. 1st sg. **ní**, 2nd sg. **tyé** (VT49:51)

We see that it follows the Eurasian pattern, if we assume it to be just a nasal versus stop/fricative contrast. Regarding the place of articulation, the most common pattern is an N-K contrast. Things vary in the particular languages: Only Goldogrin and Sindarin have [m] in an independent pronoun, but contrast it with [f] and [x] respectively. On the other hand, late Quenya has [ç] (properly a palatal stop which one can count as a dental in a rough approximation) in the second person, but contrasts it with [n] in the first person.

The type of assonance in Common Eldarin is rhyme: In the first table in VT49:50 only NI- and KI- show the same vowel /i/, while the other pronoun roots show /e/. On the other hand, assonance disappears when the 2nd sg. independent pronoun becomes **kje** > **tyé** in Quenya, versus **ní**; although a kind of assonance seems to be kept in Sindarin **im** versus **ech**, which are both of a VC syllable type.

Tolkien makes an interesting remark regarding the origins of the root KI-, saying that it '*linguicized*' a *click* (*inbreathed withdrawn t*) used in impatience or to draw attention of someone (*esp. subordinate*) to oneself (VT49:50). Interestingly, the idea of inbreathing as pointing to oneself is similar to Traunmüller's rationalization of the T-forms in the second person as pointing towards the interlocutor with the ejection of air. It appears that Tolkien wanted to create some sort of motivation behind the choices of at least some of the original Primitive Quendian sounds. It therefore seems entirely possible that he thought of a nasal in the first person as having a sound-symbolic quality as well.

Note also that the vowel in the first person is constantly /i/, which is quite a natural choice considering the significance of /i/ in proximal forms described above.

Considering the hypocoristic words for 'mother' and 'father', we find the following:

- AMA-, \*ATA-: Q. **ama** 'childish word for "mother"', **atta** 'child's word, "father"' (QL:30,33), **(m)ami**, **(m)ama** 'affectionate or childish diminutives = mummy, etc.', **(n)adi**, **(n)ada** 'affectionate or childish diminutive, father, daddy' (GL:17,19)
- \*?BAP-, \*?PAP-: G. **babi**, **baba** 'mummy, mama, etc.' (GL:21)
- NAN-, ATA-: N. **nana** '(hypocoristic) mother' (LR:374), Q. **atto** 'father' (hyp.) (LR:349)

- \*AT-, \*EM-: Q. {(a)taryo} 'daddy', {(a)milye} 'mommy' (VT48:4), atto/atya, emme/emya (VT47:10, VT48:6), S. atheg, emig (VT48:6)

We see an M-T, sometimes an N-T contrast where Goldogrin/Noldorin/Sindarin alter **t**, **tt** > **d**, **th**. Both words are always equally long and there is an assonance which takes the form of an alliteration on the vowel /a/ (cf. **ataryo**, **amilye**), or the same vocalic sequence /a-/-i/ in Goldogrin. Notable is the change of the vowel in the late sources: The word for 'mommy' now has /i/ or /e/ instead of /a/ which can be again seen as evidence of an association of high front vowels with the feminine category. It might also be the reason for the asymmetrical use of diminutive suffixes in Sindarin **atheg**, **emig**, as **-ig** would have led to S. \***ethig** by i-umlaut (compare also **honeg** 'little brother', but **nethig** 'little sister' and **gwinig** 'little baby').

Note also that N. **inw** 'female' is reformed after **anw** 'male', forming an assonant pair, instead of regular \***ein** < \***inyā** (Q. **inya**) (Etym:INI-, 3AN-).

## 4 Clustering or abstraction

One should say that it is possible to argue in the same manner as Socrates and Humboldt have done, by giving a list of examples, as long as their amount is statistically significant and the semantic domain clearly definable. One such commonly noted sound-symbolism in English is the cluster *gl-* appearing in words like

*glitter, glisten, glow, gleam, glimmer, glare, glint, glance, ...*

apparently invoking an idea of visibility and luminosity. However, looking at it diachronically, we find that the meaning of *glance* was 'to strike obliquely', later shifting to 'look obliquely'. *Glean* might belong here as well: Its original meaning is 'gather grain or other produce left by reapers', but it was extended to 'gather information or material bit by bit', so that it is now often used in a visual context like reading.

What happens here is that the association of *gl-* with visibility is recognized and affects other words which change their meaning in its favour. This results in a language-specific symbolism, sometimes called *clustering*. The original cluster of words might consist out of distantly related derivatives, or might simply arise by accident.

The same clustering effect occurs in Elvish languages as well, except that Tolkien calls it 'abstraction':

By 'abstraction' some **part** of [a] word *originally integral* is taken from it and added to other words, as an indicator of their sharing in some part of the significance or character of the source word. [...] Thus the association (observable in Eldarin) or **r** in suffixal elements with *masculinity*; but **s** with *femininity* may in original depend, at least in part, on the accident that the most ancient basic word for *man* ended in **r**, and the equivalent word for *woman* in **s**. Several of the most ancient bases expressing number, frequency, or abundance contain **l** (**r**) or **m**, and the appearance of *l*, *r* and *m* as plural inflexions may be connected with that fact. (PE19:103)

The unmentioned words are undoubtedly (N)DER- 'man' and N(D)IS- 'woman', as well as LI-, RIM- 'many'. Even though Tolkien talks about this phenomenon in the context of grammar, where inflexions arise due to it, this is basically nothing but sound symbolism. Compare for example the English words 'father' and 'man' – they share a common idea of a male person, but there is nothing in the word-shapes [fɑ:ðə(r)] and [mæn] which would encode that kind of relationship. On the other hand, Q. **atar** and **nér** share the common element **r**. In the former case, it is a derivational suffix applied to ATA-, but in the latter case it is part of the root. Such elements are usually called *phonaesthemes*, assuming an intermediary place in a hierarchy between morphemes and phonemes.

We also encounter a drift of meaning akin to the English *gl-* words at various instances of cross-influence between roots. For example, the Sindarin prefix **gor-** 'hard, difficult' (< GUR-) acquires a sense 'very painful and horrible to do' by an association with ÑGUR- 'death', ÑGOR- 'terror, dread' (PE17:154) and leads to a phonaestheme **gor/gur**. It is a whole morpheme in the case of the abovementioned prefix, but only part of a morpheme in words like **goroth** 'horror' (< ÑGOROTH-, LR:377).

A systematic investigation of phonaesthemes in the Elvish languages is not the purpose of the present study, however.

## 5 Luminosity and colour symbolism

By the end of the 19th century, several experiments were undertaken to determine the association of vowels with colours. Fechner handed out questionnaires to native speakers of German, Claparède to native speakers of French, Bleuler and Lehmann to native speakers of Swiss German. In his 1893 book [9], Flournoy compiled the results into his tables II and III. I present them here in a condensed form, giving only the most frequent responses (usually ranging between 20-30%) in table 1.

	Fechner	Claparède	Bleuler and Lehmann
/i/	yellow	red-white	white
/e/	yellow-green	yellow-blue	yellow
/a/	white-red	white-red-black	black-red
/o/	red-blue	yellow-red	yellow-red
[u]	black-brown	brown	red-black

Table 1: Responses about the colour of vowels from three questionnaires; only those with frequency above 19% included. Note that [u] is spelled <u> in German, <ou> in French.

There is undoubtedly a lot of bias in the results. Some of it cannot be helped, as speakers tend to be influenced by the vowels in the names of the colours in their own native languages. We see, for example, that /e/ is predominantly yellow for the speakers of German where the native word is *gelb*, and /i/ is predominantly white where the native word *weiß* is pronounced with a long [i:] in Swiss German. Additionally, the vowels seem to have been presented in written form, which allows for influence by the shape of the letters. For example, the letter *o* encloses a white space while *i* is a straight black line, *e* is silent at the end in French which might make it seem grey. The presence of the letter *i* in French *noire* 'black' might be an influence, even though it is not pronounced.

And overall, we see that the assigned colours vary wildly. If a tendency can be derived, it would be that /i/ and /e/ tend to be yellow or white (even for French speakers where the native words *jaune* and *blanc* do not contain them), and [u] tends to be black or brown. Otherwise, no correlation is apparent.

However, in his data, Claparède had also gathered general attributes assigned to the vowels. Flournoy evaluated them by devising the three categories CLEAR, MEDIUM and DARK (*claire, moyen, foncé*) and distributing the judgements among them. Into the CLEAR category he put attributes like 'bright', 'brilliant' and so on, as well as inherently bright colours, like white, yellow and orange. Into the MEDIUM category he put all the other colours except for black, and attributes which could not be qualified as 'clear/bright' or 'dark'. Into the DARK category he put 'black' and all attributes associated with gloom and darkness. The result is shown in table 2 and is compared with the questionnaire of Bleuler and Lehmann which already had a CLEAR/MEDIUM/DARK category.

	Claparède	Bleuler and Lehmann
	CLEAR : MEDIUM : DARK	
/i/	58:33:9	90:6:4
/e/	54:36:10	72:20:8
/a/	37:32:31	30:30:40
/o/	37:37:26	22:46:32
[u]	20:42:38	8:34:58

Table 2: Procentual proportions of judgements in the CLEAR, MEDIUM and DARK categories from two questionnaires. (From [9], table III.)

One can see that while the numbers differ, the overall tendency is the same: When going from front to back vowels, the vowels increasingly change from the CLEAR into the DARK category (/o/ seems actually to come before /a/, but the difference is marginal). Flournoy even calls it 'the law of clarity' (*Loi de Clarté* [9], p. 68). Also, the CLEAR category is much more strongly defined, in the data from Bleuler and Lehmann the percentages reach 70% and even 90%.

The results can thus be summarized as such: There seems to be a tendency to match the front-back scale (equivalently, the F2 frequency scale) of vowels with brightness or clarity, where front vowels appear to be clear or bright and back vowels appear to be dark. The particular hue does not seem to be very important, except that yellow and white may be perceived as inherently clear/bright, and brown and black as inherently dark.

These results agree with other experiments without any reference to hue.

Newman's experiment, as described above, also included bright/dark judgements with the following result:

- bright [i] < [ɪ] < [æ] < [ɑ] < [ɔ] < [u] dark
- bright [k] = [s] ≈ [l] ≪ [h] < [p] < [ɕ] ≪ [n] < [g] < [b] < [r] ≪ [gl] ≪ [d] ≪ [m] < [gr] < [br] dark

Lawrence Marks [12] carried out an experiment to determine the association between brightness and sound. Participants were presented slips of grey paper and could tune the frequency of a tone played to them, until it would synaesthetically match the degree of brightness of the grey surface. The result was that higher frequency would correspond to brighter surfaces.

Turning to Elvish, there seems to be a quite similar, but at the same time different picture. We first take a look at primary roots of colours, i.e. those which seem to point to a colour by themselves, and where the meaning is not derived from words like 'fire', 'sky' or 'blood'. It is not always easy to determine whether a root is primary or not. When in question, I have treated them as primary.

The roots with the vowel /i/ are as follows:

- + {GLINDI- 'pale blue': Q. **ilin**, N. **glind**, **glinn** (LR:359)}
- + ISI-: Q. **iska** 'pale, white' (QL:43)
- + KR̄N-: G. **crintha** 'rosy, pink' (QL:48, GL:27)
- + MITH-: N. **mith** 'grey' (LR:373)
- ± NINI-: Q. **ninwa** 'blue', **nin(d)** 'blue colour, blueness, blue green' (QL:12), G. **nim** 'pale blue, blue green' (GL:60)
- ± \*NIN-D-: Q. **ninya**, N. **neinn**, **nainn** 'blue' < **nindyā** (PE13:150,164) (cf. also later {NIN-DA- 'blue'} without derivatives, changed to NIN-DI- 'fragile, thin' (VT46:4))
- + NIQI-: Q. **ninqe** 'white' (QL:66), G. **nimp** 'pallid' (GL:60), N. **nimp** 'wan, sickly', 'pale, wan' (PE13:151,164)
- + SINI- 'pale blue ??': **sine** 'pale (blueish) green' (QL:83)
- + THIN-: CE **thindi** 'pallid, grey, wan', Q. **sinde** 'grey', N. **thind**, **thinn**, Ilk. **thind** 'grey, pale' (LR:392)
- + {WIN-/WIND-: CE **windi** 'blue-grey, pale blue or grey', Q. **vinde**, N. **gwind**, **gwinn**}, CE **windjā** 'pale blue', Q. **win(d)ya**, **vinya**, N. **gwind** (LR:399, VT46:21)

There are only a few roots with the vowel /e/:

- ? GWEN-: Q. **wenya** 'green, yellow-green, fresh', Ilk. **gwene** 'green' (LR:359)
- ? SENE-: Q. **senna** 'red-brown, red' (QL:83), G. **sen** 'brown-red, russet', **sentha** 'russet' (GL:67)
- ? *sleiwa*: Q. **laiwa**, T. **líva**, Ilk. **slíw**, N. **lhui** 'pale' (PE13:149)

Then follow the roots with the vowel /a/:

- + AWA- 'burn, be parched, yellow, warm': Q. **aürqila**, **auriqilea** 'golden, sunlit' (QL:33)
- ± BARÁN-: Q. **varne**, EN **baran** 'brown, swart, dark brown' (LR:351), S. **baran** 'brown, yellow-brown' (RC:343)
- + GAY-: Q. **aira**, N. **gaer**, **goer** 'red, copper-coloured, ruddy' (LR:358)
- + GLAW(-R)-/(G)LÁWAR-: Q. **laure**, N. **glaur** 'gold (of colour and light, not the metal)' (LR:358,368)
- + GWALA-: Q. 'walna, 'walin(a) 'brown' (QL:103)
- + GWAY-: Q. **waina**, **vaina** 'blonde, fair of hair', S. **baen** 'fair-haired' (PE17:154-155)
- – KAMA-: Q. **kanwa** 'dark red' (QL:44)
- + KARÁN- 'red': Q. **karne**, N. **caran** 'red' (LR:362)

- + KR̄N-: Q. **karne** 'red', **karmalin** 'russet, orange-red' (QL:48), G. **carnin** 'scarlet' (GL:25)
- + LÁYAK-: Q. **laiqa** 'green', N. **lhoeb** 'fresh - 'green' only in Lhoebenidh' (LR:368)
- – MAD-: Q. **marya**, N. **meið**, **maið** 'pale, fallow, fawn' (LR:371)
- ± MALA-: Q. **malwa** 'pale, yellowish', **malina**, **malin(d)** 'yellow', **malikon (d)** 'amber' (QL:58), G. **malon** 'yellow', **maltha** 'golden yellow, rich, mellow' (GL:56)
- ± \*MAL-: N. **malw** 'yellow', **malen** 'wan, pale', 'pale, yellowish, sickly' (PE13:150,163)
- + \*PHAN- > FAN- 'white': Q. **fanya**, S. **fain** 'white and shining' (PE17:26)
- + SMAL- 'yellow': Q. **malina**, N. **malen** 'yellow' (LR:386)
- + SPAN- 'white': Q. **fána** 'white', N. **fein** (LR:387)
- + WAN-: Q. **vanya**, **vana** 'fair-haired (yellow to golden)', S. **gwain** 'blonde' (PE17:150)

And finally, roots with the vowels /o/ and /u/ (some are left out and will be discussed in the next section):

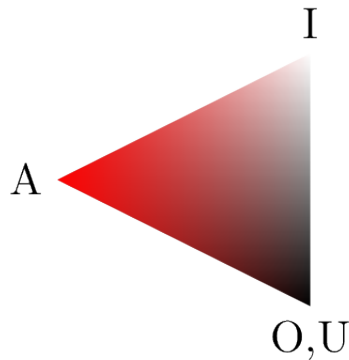
- + DUN-: 'dark (of colour)': Dor. **dunn**, Dan. **dunna** 'black', N. **donn** 'swart, swarthy' (LR:355)
- – \*LOS-: G. **gloss** 'white, clear white' (GL:40)
- ? LUG-: Q. **lúne**, N. **lhûn** 'blue' (LR:370)
- ? LUY-: Q. **luine**, S. **luin** 'blue' (VT48:23)
- ? \*LU'U-: Q. **lûne** 'blue, deep blue' (QL:57), G. **luim** 'blue' (GL:55)
- ? \*HY(UN)U-: Q. **hyúne** 'deep blue' (QL:41)
- + MOR-: Q. **more**, N. **†môr**, **morn** 'black' (LR:373)
- + MORO-: Q. **morna**, **morqa** 'black' (QL:62), G. **morn** 'dark, black' (GL:58)
- + ÑGOL- 'dark-hued, dark-brown': OQ **ñolda** 'dark-haired' (PE17:125)
- – RUSKĀ-: ON **ruska**, N. **rhosc** 'brown' (LR:385)
- + \*SROB-: Q. **hróva** 'dark, dark brown' (PE17:154) [presumably of hair]

I propose the following interpretation: There seem to be three main categories of colour roots based on the root vowel, let us call them PALE, PURE and SOMBRE.

The PALE category is associated with the front vowel /i/ and is either grey, or a mixture of a colour with white. The hue is not particularly important, we see 'pale blue', 'pale green' and even 'pale red' (i.e. 'pink') if G. **crintha** rather than the root KR̄N- is to be counted. The PURE category is associated with the central vowel /a/ and contains all saturated, pure colours 'red', 'yellow' (including 'gold' and 'blonde'), 'white' and 'green'. The SOMBRE<sup>8</sup> category is associated with the back vowels /o/, /u/ and contains black as well as mixtures with black. The vowel /e/ is almost never used in colour roots.

This can be summarized in the following diagram:

<sup>8</sup>Unfortunately, it is not easy to find a word meaning 'a colour mixed with black' other than 'dark'. I would like to avoid 'dark' and leave it for the luminosity (bright/dark) scale without a reference to hue.



In the above list, I have marked all colours adhering to this division with a plus sign: When being white, grey or qualified by 'pale' in the PALE category, when being a saturated colour in the PURE category, or when being black or qualified by 'dark' in the SOMBRE category. Non-adherence is marked by a minus sign. When glosses are competing, a plus-minus sign is used. Somewhat unclear is the position of brown in this division, as typical brown is always a mixture with black. In terms of perception, however, it may be regarded as a pure colour. I have treated it as a SOMBRE colour when qualified as 'dark brown', and as PURE colour when the gloss is just 'brown'. The appearance of blue in the SOMBRE category is labeled with a question mark (to be discussed further below), and so are the roots with /e/ which are too few to support a statistical judgement.

At this point, let us also take a look at colours from non-primary roots:

- – DUS- ?'burn': N. **dost** ?'brown' (VT45:11)
- ? ʒEL- 'sky': Q. **helwa**, N. **elw** '(pale) blue', Dor. **gelu** 'sky-blue' (LR:360)
- + GAL-/KAL- 'shine/light' (PE17:153,156): S. lenited **lân** 'white' (UR:390, LotR)
- – GOLÓS- \*'snow': N. **gloss** 'snow-white' (LR:359)
- ? HELE-, Q. **helin** 'pansy': Q. **helinille** 'violet', **helinwa** 'of pansies, of colour = "a blue-violet"', **helinqila** 'mauve' (QL:39)
- ? HESE- 'wither': G. **heth** 'white, pallid, wan' (QL:40, GL:49)
- + KAL- 'shine (general word)': N./S. **calen** 'bright-coloured = green' (LR:362, PE17:153)
- + KHIS-/KHITH- 'mist, fog': Q. **hiswa** 'grey' (LR:364)
- ? KYELEP- 'silver': Q. **telepta** (LR:366, PE17:71)
- + LAÿA- 'be alive, flourish': Q. **laiqa** 'green' (QL:52), G. **laib** (GL:39)
- – LOS- 'snow': S. **glos**, **glosui** 'snow white' (PE17:161)
- + NAR- 'flame, fire': N. **narw**, **naru** 'red' (LR:374)
- + NIK-: S. **nimp** 'pale, pallid', **nim** 'white (usual word)' (PE17:168)
- + NIK-W- \*'snow': Q. **ninqe** 'white', N. **nimp**, **nim** 'pale' (LR:378)
- ? TEL(E)PE- \*'silver': Q. **telpia** 'silvern' (QL:91)
- + TORO- \*'bake': Q. **torwa** 'baked : dark (rich) brown' (QL:94), G. **torn** 'sunburnt, swart, dark-brown' (GL:71)
- + YAR- 'blood': Ilk. **arn** 'red' (LR:400)

One can see that the front vowel /i/ is again used very consistently, it almost always describes pale colours (with the exception of NINI- in an early source which yields a general word 'blue', and the abolished NIN-DA- 'blue'). KAMA- as 'dark red' in the PURE category is an isolated example, and MAD- seems to be an *a posteriori* attempt at explaining the already existing name **Maidros** as 'pale-glitter'. With regard to the vowel /e/, GWEN-, SENE- and HELE- would belong into the PURE category, *sleiwā* (where the root is unclear), ǞEL- and HESE- into the PALE category (KYELEP-/TEL(E)PE- may be put into either of them), so that it would seem that it assumes a position in between, according to its position between /i/ and /a/ on the F2 frequency scale.

Note that the colour 'white' appears twice: If one proposes a PALE category as a mixture with white, white itself has apparently to be in it. On the other hand, it can be regarded as a pure colour in terms of perception. There is a notable difference between the roots, however: NIK(W)-/NIQI- often tends to mean 'pale'. And while SPAN-/PHAN- is associated with clouds and dimmed light ('veil' is also the basic meaning of the root in some conceptions), in Quenya it is associated with the bright, majestic and shining raiments of the Valar (an idea later borrowed into Sindarin from the exiled Noldor) (PE17:26, RGE0). Similarly, S. \***glân** (from **Curunír 'Lân** 'Saruman the White') seems to be derived from 'bright, shining' (Goldogrin had **glan** 'clean, pure' from 'bright' (GL:39)).

It seems contradictory to see the colour 'blue' in the SOMBRE category. A way to explain it would be to fine-tune the categories. A common perceptual contrast is the one between warm and cool colours. Experimentally, one observes that if presented with a choice of dividing the colours white, black, red, yellow, orange, green, blue and violet into two piles, most people would put white, yellow, red and orange on one side; and green, blue, violet and black on the other. It has been furthermore determined that orange is felt to be the warmest colour, and blue the coolest [11].

For reasons which will be clear later, I would like to avoid the warm-cool metaphor, however. Let us just say that red and yellow go through the same perceptual channel and have an affinity to white, while green and blue go through another and have an affinity to black. This would explain the appearance of blue in the SOMBRE category. Green, on the other hand, seems to have an affinity to brightness in the Elvish languages, contrary to human perception. For example, **calen** 'bright' becomes 'green' in Noldorin/Sindarin. At all stages, the roots KAL-/GAL- 'shine' or 'light' seem to be closely connected with GALA- 'grow' (yielding words for 'tree', 'plant' and the like). This relationship has been investigated in detail by Helios De Rosario Martínez [13].

Also interesting to note is the appearance of GOLÓS-/LOS- from the earliest to the latest sources with a back vowel, but with the meaning 'snow' and the colour 'white' derived from it (in Gnomish, it might also be a primary root). Compare Newman's brightness scale listed above: The consonants [k], [l] and [s] were judged to be the brightest in his experiment, which might counteract the association of the vowel /o/ with darkness. It is also notable that the root for 'bright/shine' has always been KAL- throughout the external history of Elvish, involving the same consonants (see below).

Furthermore, the usual root for 'black' in Elvish is MOR-. Correspondingly, [m] and [r], along with [d], are far on the side of darkness in Newman's results (see above). In fact, disregarding the clusters [gr] and [br], [m] remains as the darkest single consonant. Regarding [d] and [r] we might suspect an influence from the English word *dark*, however.

The SOMBRE category has been defined as a mixture with black, but apart from black itself (and the already discussed blue) it only really contains dark brown and the generic gloss 'dark-hued'. The reason for this is probably perceptual: Russian is often cited as having a separate word for 'light blue', which is голубой (not derived from синий 'blue'), and English parallels it with 'pink' for a mixture of red and white. Mixtures with white are *new sensations* and are often culturally distinguished by separate words. Mixtures with black, on the other hand, are not new sensations, but are simply perceived as dark variations of the respective hue. The only exception might be dark brown, where Russian incidentally also shows a separate word бурый, distinct from коричневый 'brown'.

In fact, a system of classifying colours as hues mixed with white and black does already exist and is called *Ostwald's system*<sup>9</sup>. The colours are presented in the same triangle diagram as above, one for each hue. It is divided into fields to which coordinates are assigned. Furthermore, this system was used in a psychological experiment as well, conducted by Annelies Argelander [10] where participants listened to notes played to them on the piano and had to point at the appropriate colour in Ostwald's system which synaesthetically matched the tone. The result once more was that there was a lot of variation with respect to hue, but much less variation with respect to the position on the triangle itself, i.e. the proportions of black and white. Argelander gives the following result of matches with the notes ([10] pp. 38-43, 24):

- d<sup>h</sup>ʰ, eʰ, fʰ — PALE

<sup>9</sup>In the context of digital display it is simply referred to as HSB (*hue, saturation, brightness*), beside the RGB system (*red, green, blue*).



- a', h', c'', d'' — PURE
- f', g', g# — SOMBRE

This means that high notes correspond to mixtures with white, medium notes to pure colours, and low notes to mixtures with black; in complete agreement with the F2 frequency of the vowels /i/, /a/ and /o, u/ in Elvish and their association with the same categories. Argelander also mentions that the dull colours in the middle of the triangle which contain both white and black were almost never chosen ([10], p. 39).

Since brightness matters so much for vowels, the question naturally arises: How are the words for 'bright' and 'dark' distributed in terms of vowels? Let us therefore gather all roots whose derivatives contain 'bright', 'brilliant' or 'day' without any particular reference to hue. We find the following:

- + AR- 'day': Q. **are**, N. **ar-** (LR:349)
- + AWA- 'burn, be parched, yellow, warm': G. **auglas**, **auglos** 'bright sunshine' (GL:20)
- + KAL- 'shine (general word): N. **gail** 'bright light', **glaw** 'radiance' etc. (LR:362)
- + KALA- 'shine golden': Q. **kalma** '(day)light', **kalke** 'brilliant' (QL:44), G. **gala** 'light, daylight' (GL:37) etc.
- + LKL-: Q. **alkara** 'brilliant' (QL:30), G. **glâ** 'day, daytime', **glarw**, **glarwed** 'bright' (GL:39)
- – SELE-: Q. **selka** 'bright' (QL:83)
- – UR- 'heat, be hot': S. **aur** 'a whole day', Q. **aure** 'day as opposite to night' (PE17:148, Silm)

Conversely, we gather all roots with 'dark', 'gloom(y)' and 'night':

- – DEM- 'sad, gloomy': Ilk. **dim** 'gloom, sadness' **dem** 'sad, gloomy' (LR:354)
- + DO3-, DÔ-: Q. **lóna**, N. **dûr** 'dark, sombre', **daw** 'night-time, gloom' (LR:354)
- + DOM-: Q. **lôme**, S. **dû** 'night' (PE17:152)
- + FU3U-: G. **fung** 'dark (esp. of colour)', **fu** 'night', **fuin** 'secret, dark – hoarded' (GL:36)
- + hUÝU-: Q. **hui** 'fog, dark, murk, night' (QL:41)
- + LOMO-: Q. **laumea** 'dark, stormy' (QL:51), **lôme** 'dusk, gloom, darkness' (QL:55)
- + LUM-: Q. **lumbe** 'gloom, shadow' (LR:370)
- + LUVU-: Q. **lūrea** 'dark, overcast', **lūre** 'dark weather' (QL:57)
- + MOR-: Q. **móre** 'blackness, dark, night', **morna** 'gloomy, sombre', N. **maur** 'gloom' (LR:373)
- + MORO-: Q. **mōri** 'night' (QL:62), G. **morth**, **moroth** 'darkness', **mûri** 'night, darkness' (GL:58)
- + NDUL-: Q. **nulla**, N. **doll** 'dark, dusky, obscure' (LR:355,376)
- + (N)DUR-: S. **dûr** 'dark, gloomy, 'hellish'' (PE17:152)
- + ÑGUL- '(possibly in origin simply a variant of ÑGOL applied to a darker shade?) dark – with sinister connotations': Q. **ñúla** 'dark, occult, mysterious'
- + PHUY-: N. **fuin** 'night' (LR:354,382)
- + THUS- 'evil mist, fog': S. **thû** 'horrible darkness, black mist' (PE17:183)
- + TUM(B)U-: Q. **tumna** 'deep, profound, dark or hidden' (QL:95)
- + UÑG-: Q. **ungwe** 'gloom', Ilk. **ungol** 'darkness', **ungor** 'black, dark, gloomy' (LR:396)
- – YAĐA-: Q. **yanda** 'dark, gloomy' (QL:105), G. **gand** 'dark & gloomy [...]' (GL:37)

Both lists seem to confirm the hypothesis, the former predominantly shows /a/ (UR- will be discussed further below), the latter /u/ and sometimes /o/. In terms of consonants, one notices that in the latter list, half of the SOMBRE roots contain [m], [d], or [r].

Another typically Elvish trait, however, is the presence of a third light category. Let us call it GLEAM and count all roots meaning or yielding 'gleam', 'glint', 'sparkle', 'glisten', 'twinkle' and so on. They are:

- + GIL-: G. **gil-** 'gleam, shine pale and silver, as of the moon' (GL:38)
- + GIL- 'shine (white)': N. **gael** 'pale, glimmering' (LR:358), S. **geil, gail** 'silver spark' (PE17:152)
- + \*GLIN-: G. **glim** 'shimmering' (GL:39)
- + GLIM- 'gleam, glint (usually of bright slender but bright shafts of light)': S. **glintha-** 'glance (at)' (WJ:337)
- + GWIMI-, GWINI- 'sparkle': Q. '**winwinoite** 'glistening', '**wimpele** 'a twinkling' (QL:104), G. **gwinc** 'a spark. flash', **gwimla** 'wink, twinkle' (GL:45)
- + \*(G)IL- or \*(G)ILIK-: Q. **ilkala** 'gleaming' (MC:221-223)
- + (I)LK-: **ilke** '(a glint) appearance' (QL:42)
- + \*IT- 'sparkle': Q. **itila** 'twinkling, glinting' (PM:363, PE17:112)
- + IT-, ITH-, ?IS-, ?IT?- 'glitter, shine, shimmer, twinkle' (PE17:156)
- - KATYA-: Q. **kate** 'a gleam, ray' (QL:45)
- + MBIRÍL-: Q. **mirilya-** 'glitter' (LR:372)
- - ÑGAL- 'gleam, sheen': PE **ñal(a)dā** 'gleaming sheen' (PE17:59)
- + RIL- 'glitter': Q. **rilma** 'glittering light', **rilya** 'glittering, brilliance' (LR:383)
- + SILI- (PILI-): Q. **sili-** 'gleam, glint' (QL:83)
- + SNTŶN- 'twinkle': **sint** 'a gleam, spark, glisten, drop of dew, etc.' (QL:85)
- + THIL-: N. **thilio** 'to glisten' (LR:392)
- + TILL-: Q. **tilwe** 'a glimmer' (QL:92)
- + TIN- 'sparkle, emit slender (silver, pale) beams': Q. **tine** 'it glints', **tintina** 'it sparkles' N. **tinno** 'to glint', **tinw** 'spark, small star'
- + TINI- 'twinkle': Q: **tint** '(silver) spark' Q. **tinya-** 'sparkle, etc.' (QL:92), G. **tinc** 'a flicker, a gleam' (GL:70), **tim** 'spark, gleam (star)' (LR:393)

Apart from two counterexamples which are isolated (i.e. not repeated otherwise) they all contain the vowel /i/. Even though **rilya** is translated as 'brilliance' and GLIM- is said to be used of bright shafts of light, both seem to refer to point-like light sources, in particular stars, eyes and jewels (notably the Silmarils, of course). These light sources are small in size, so that the presence of the vowel /i/ forms a very natural association with the diminutive character of /i/. It is probably no coincidence either that the root for 'jewel' most of the time takes the form MIR-, with the same vowel.

On the other hand, we have seen that /i/ is also associated with the PALE category; and it terms of light, this seems go further to twilight or dimmed light. Compare HISI-: Q. **híse, histe** 'dusk', **hiswa** 'dim, fading' (QL:40) or KHIS-/KHITH-: Q. **híse**, N. **hith** 'mist, fog', N. **hethw** 'foggy, obscure, vague' (LR:364). 'Twilight' and 'point-like light sources' seem both to convey the idea of 'small light' which is somehow weaker in overall intensity compared to the light of the sun.

The most common words for 'sun' (Q. **Anar**, S. **Anor**) and 'moon' (Q. **Isil**, S. **Ithil**) also fit the above pattern, where the moon is apparently put into the PALE-GLEAM category. Consequently, 'evening' usually belongs into this category as well, e.g. Q. **sinye** < THIN-, Q. **vinya** < WIN-, WIND- (LR:399).

Regarding the stars, there are two common derivations: The one is from GIL- or TIN- (Q. **tinwe**, S. **gil**) with the basic meaning 'spark'. The other one is from EL- (Q. **elen**, S. **êl**) (Etym:EL-, WJ:362-363 etc.). The latter root would still fit into the PALE-GLEAM category if one counts /e/ into the same front vowel group as /i/,

but I do not believe that this is the right association. In PE13:143, we encounter the Noldorin words **elven** 'wonderment, wonder' < **elmenda** and **elvain** 'wonderful, wonder, marvel' < **elménd(i)ya** which seem to be related, and in WJ:360 we are told that **ele** was an explanation 'lo!, behold!' made by the Elves when they saw the stars. So I believe, this is just a different way to classify the same phenomenon – either as a small light source, or as something beautiful to marvel at.

Summarizing, one finds the following categories in Elvish:

- PALE-GLEAM (front /i/): pale and grey colours, white, sparkling point-like light sources, moon and stars, evening
- PURE-BRIGHT (central /a/): saturated bright colours, shining white, brightness/brilliance, day, sun
- SOMBRE-DARK (back /o/, /u/): sombre colours, black, blue, darkness/gloom, night

The association of /a/ with the highest light intensity might again be explained by the high sonority of low vowels.

## 6 Temperature symbolism

There is an additional group of roots with the vowels /o/ and /u/, which, according to the previous discussion, would be expected to denote sombre colours, and yet do not. They are:

- {GUL- 'glow, shine gold or red': N. **goll** 'red'}
- KUL- 'golden-red': Q. **kulda**, **kulina** 'flame-coloured, golden-red', **kuluina** 'orange', N. **coll** 'red' (LR:365)
- KULU- 'gold': Q. **kuluksa**, **kuluvoite** 'golden', **kulurinda**, **kulmarinda** 'orange-coloured' (QL:49), G. **culwin**, **culuin** 'golden' (GL:27)
- \*MBUR-: S. **born** 'hot, red' (Let#347) [see below]
- ROY- 'ruddy, red': Q. **roina** ruddy, N. **gruin** (LR:384)
- RUYU- 'blaze (red)': no colour derivatives [see below]
- (U)RUS- 'used of a varying brownish red from what we should call brick-red to auburn': {Q. **ruska** 'red-brown'}, Q., T. **urus**, S. **urust** 'copper', Q. **rusa** 'red-haired', S. **ross** 'red-haired, copper coloured [...]' (VT41:10)
- YUL- 'smoulder': Q. **yulme** 'red [heat?]' (LR:400)

This seems to be a major contradiction to everything that has been said in the previous section, as the colours do not fit into the SOMBRE category. However, we see a hint at an interpretation by the appearance of 'glow', 'smoulder', 'blaze' and 'flame', so let us call this category GLOWING. One also notes that KUL(U)- and (U)RUS- are used with reference to metals. In fact, in Goldogrin, **culu** is said to be *the name of all red and yellow metals*, as opposed to **giltha** < GIL- which is the name of all white and grey ones. Note also the root RUS- 'flash, glitter of metal' (LR:384).

The above /u/-group therefore seems to be connected with fire and red-hued metals like copper and red gold. To investigate this, let us take a look at all the roots yielding 'cold', 'chill(y)' and 'winter'. They are (with non-primary roots appearing in parentheses):

- + DÝELE-: Q. 'yelwa 'cold', 'Yelin 'Winter' (QL:106), G. **Gilim** 'winter' (GL:38)
- – (FALA- 'bare, nude': G. **folornin** = **Gilim** (QL:37, GL:35))
- – N. **harch** 'chill', 'chill, cold' (PE13:148,163)
- + (HELE-: Q. **helka** 'ice-cold' (QL:39), G. **helc**, **helw** 'ice-cold, icy, cold' (GL:48))
- + (HESE- 'wither': Q. **Hesin** 'winter' (QL:40), G. **Hess** 'winter', **hesc** 'chilled, chill' (GL:49))
- + (KHELEK- 'ice': Q. **helka** 'ice-cold', N. **helch** 'bitter cold' (VT45:21))

- + N. **lhing** 'cool' (PE13:149)
- + NIK-: Q. **niku-** 'to snow, it is chill, it freezes', **ninque** 'white, chill, cold, pallid' (PE17:168)
- + RINGI- 'cold': Q. **ringe**, N. **rhing**
- + RIDI-: Q. **ringa** 'damp, cold, chilly' (QL:80), G. **ring** 'cool, cold' (GL:65)
- + \*SRIB-: Q. **hríve**, S. **rhîw** 'winter' (LotR, PM:136)

We compare them with all the roots yielding 'hot/heat', 'warm' and 'summer'. They are:

- + AS- 'warmth': Q. **āze**, **āre** 'warmth, especially of the sun, sunlight' (PE17:148)
- + AWA- 'burn, be parched, yellow, warm': Q. **aurin** 'warm', **austa** 'summer, high-summer' (QL:33), G. **aust** 'summer' (GL:20), N. **oren** 'hot (weather)' etc. (PE13:151,160)
- + BARĀS-: CE **\*barasá** 'hot, burning', N. **brassen** 'white-hot'
- + KULU- 'gold': G. **cluim** 'warm (cosy)', **Clui** '[...]' used for "Sun", Aur, and for sunny warmth' (QL:49, GL:26)
- + LAW- 'warm': Q. **lauka**, N. **lhaug** 'warm' (LR:368)
- + LAY-: Q. **laire** 'summer', S. **Laer** 'summer' (PM:136, PE17:159)
- + LUSU- 'foment, cherish, warm, bathe': Q. **lūsina** 'warm, glowing [...]' (QL:57)
- + \*MBUR-: N. **boir**, **boer** / **bordh** 'hot, raging', Q. **mure** 'heat, close weather' (PE13:139,160)
- + SAHA-, SAHYĀ- 'be hot': Q. **saiwa** 'hot' (QL:81), G. **sair** 'hot', **saiwen** 'summer, midsummer' (GL:66)
- + (SOVO-, SOWO- 'wash': Q. **sausa** 'bath water, hot water' (QL:86))
- + (TORO- \*'bake': Q. **torqa** 'blazing hot' (QL:94))
- + URU-: Q. **ūrin** 'blazing hot' (QL:98), G. **urin** 'hot, very hot' (GL:75)
- + UR- 'be hot': Q. **urya-** 'be hot' (LR:, PE17:148)
- + URŪ- (QL:98) < \*GUĐU-: G. **gudh-** 'am hot. (intr.)', **gauth** 'heat, espec. of body' (QL:98, GL:38,42)

We see a contrast of the front vowels /i/ and /e/ versus the back vowels /a/, /o/ and /u/. Among these, /e/ and /o/ are rare and do not seem to appear in primary roots with the given meanings – the notions of 'warm' and 'cold' are derived from meanings like 'bake' and 'ice' in these cases. Hence, the main contrast is between /i/ and /a/, /u/. The vowel /u/ might have a more intensified meaning, i.e. 'hot' rather than 'warm'.

Furthermore, the byname of the sun takes the forms **Ur**, **úr**, **Ūri**, **Ūrin**, derived from UR- 'be hot' or similar roots (QL:98, PE13:155, LR:396). This seems to be once more a case of different classification: While **Anar**, **Anor** classifies the sun as something bright, **Ur** classifies it as something hot.

Note furthermore the presence of **ur/ru** in the following roots:

- \*GURU-, cognate to QL URU-: G. **gruithog** 'ferocious, mad with wrath' (GL:41)
- \*MBUR-: N. **bordh** 'heat, rage', **byr**, **buir** 'fire', **boir**, **boer** 'hot, raging' (PE13:139), **boir**, **boer** 'heat, rage', **bordh** 'hot, raging, wroth' (PE13:160)
- PUŔU- 'consume by fire': Q. **purma** 'flame, blaze', (**uru**)**purnie** 'conflagration' (QL:75)
- RUTH- 'rage': S. **rúth** 'wrath' (PE17:183)
- RUYU- 'blaze (red)': Q. **ruina** 'blazing, fiery', **ruive**, **aparuive** 'wild fire, fire as conflagration' (ibid.)
- URU(P)- 'anger, rage, wrath': Q. **rúse** 'wrath', S. **rúth** 'anger' (PE17:183,188)

The connection of anger with a hot, boiling liquid is a common metaphor: Anger is 'burning', at the beginning it is 'brewing', then 'boiling', and unless one 'lets off some steam', one will 'explode with it', just as if the body was a cooking pot.

We can therefore understand the above /u/-group of orange-reddish colours by an association of /u/ with heat and fire. This seems to be yet another case of different ways of classification: One can think of red as a pure colour, or as the inherent colour of fire. Similarly, hotness is an inherent property of fire, so that a chain of associations is formed. This seems to be extended to reddish metals, as their glistening radiance might be metaphorically compared with fire. Note also that G. **cluim** 'warm (cosy)' seems to be derived from the same root KULU- as †**culu** 'gold, class name of all red and yellow metals'.

We can thus conclude:

- COLD (front /i/): cold, winter
- WARM (central /a/): warm, summer, sun
- HOT-GLOWING (back /u/): hot, sun, glowing-red, reddish metals

## 7 Summary

The sound-symbolic properties of vowels were fittingly summarized by Benjamin Lee Whorf in 1941 ([15], p. 267):

In the psychological experiments human subjects seem to associate the experiences of bright, cold, sharp, hard, high, light (in weight), quick, high-pitched, narrow, and so on in a long series, with each other; and conversely the experiences of dark, warm, yielding, soft, blunt, low, heavy, slow, low-pitched, wide, etc., in another long series. [...] Thus the vowels a (as in 'father'), o, u are associated in the laboratory tests with the dark-warm-soft series, and e (as in English 'date'), i (English e in 'be') with the bright-cold-sharp set.

But this is exactly what happens in Elvish! We can sum it up as follows:

- SMALL-FRAIL-FEMININE-THIS-PALE-GLEAM-COLD (front /i/):  
small size, frailness (positive and negative connotations), feminine gender, closeness to speaker, white, grey, colours mixed with white, gleaming point-like light sources, coldness, winter
- near-front /e/ is often in between this category and the next
- LARGE-THAT-PURE(OF COLOUR)-BRIGHT-WARM (central /a/):  
large size, distance from speaker, pure colours, brightness, warmth, sun, summer
- LARGE-SOMBRE-DARK-HOT-GLOWING (back /o/, /u/):  
large size, black, colours with affinity to black, gloom/darkness, heat, sun, glowing-red, reddish metals, anger/rage

The association of these vowels with colours and luminosity is well-exemplified by the triplet GIL- 'shine (white or pale)', GAL- 'shine', and GUL- 'glow, shine gold or red' (Etym).

On a personal and conclusive note, I should say that it matters a lot if a language is formed by a single person as opposed to a whole community. In the community, each change is controlled by many speakers and has to be approved by the majority before it is implemented into the language. A single person, on the other hand, will likely produce something which might be subjectively natural, but which just would not work in real communication. But a single person can make use of the gathered linguistic knowledge and apply it. Whenever Tolkien formed the same grammatical categories, or applied the same diachronic changes as found in natural languages, it is not because he was unimaginative, but because these categories and changes have proven themselves as working.

The chain associations mentioned by Whorf appear in psychological experiments, not in actual languages themselves (although I believe that they would, if it was not for the many destructive forces of language change), and only on the average. One person chosen in particular would not necessarily share all of them. Apart from

that, understanding the impression of colours is quite different from actually painting a picture; and all the sound-symbolic tendencies seem to be indeed woven into the picture of the Eldarin languages.

Of course, if one wants to fill a created language with sound-symbolic content, it is possible to work from knowledge, provided that this knowledge is available. It is very hard to say whether Tolkien knew about the papers and discussions concerning sound-symbolism at all, but Sapir's pioneering experiment was conducted in 1929 which is about 15 years after the composition of the Gnomish and Qenya Lexicons. One notices a development in the sound-symbolic concepts in Elvish over time, as they seem to become more clear-cut in the later writings, but all the foundation had been already laid in the Lexicons.

It would seem to me that Tolkien did not work from knowledge (except maybe a personal sampling of natural languages), but simply was particularly sensitive to sound-symbolism. In any case, its skilfull implementation into his languages allows us to appreciate them from a wholly different point of view.

## References

- [1] Seiichi Makino and Michio Tsutsui, *A dictionary of basic Japanese grammar* (1986)
- [2] Edward Sapir, *A study in phonetic symbolism*, *Journal of Experimental Psychology* 12(3):225–239 (1929), also published in: *Selected writings of Edward Sapir: in language, culture and personality* (1949), edited by David G. Mandelbaum
- [3] Stanley S. Newman, *Further Experiments in Phonetic Symbolism*, *The American Journal of Psychology* 45(1):53–75 (1933)
- [4] Otto Jespersen, *Language – its nature, development and origin* (1922)
- [5] Hartmut Trautmüller, *Sound symbolism in deictic words*, in: *Tongues and Texts Unlimited. Studies in Honour of Tore Jansson on the Occasion of his Sixtieth Anniversary*, pp. 213–234 (1996) [http://www2.ling.su.se/staff/hartmut/symb\\_abs.htm](http://www2.ling.su.se/staff/hartmut/symb_abs.htm) (4-page summary)
- [6] Nancy L. Woodworth, *Sound symbolism in proximal and distal forms*, *Linguistics* 29(2):173–300 (1991)
- [7] Peter Ladefoged, *A course in phonetics* (2001)
- [8] Wilhelm von Humboldt, *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß auf die geistige Entwicklung des Menschengeschlechts* (1836)
- [9] Théodore Flournoy, *Des Phénomènes de Synopsie (audition colorée)* (1893)
- [10] Annelies Argelander, *Das Farbenhören und der synästhetische Faktor der Wahrnehmung* (1927)
- [11] Paul Kay and Luisa Maffi, *Color Appearance and the Emergence and Evolution of Basic Color Lexicons*, *American Anthropologist New Series* 101(4):743–760 (1999)
- [12] Lawrence E. Marks, *On Associations of Light and Sound: The Mediation of Brightness, Pitch, and Loudness*, *The American Journal of Psychology* 87(1/2):173–188 (1974)
- [13] Helios De Rosario Martínez, *Light and Tree* (2005) <http://www.elvish.org/Tengwestie/articles/DeRosarioMartinez/lightandtree.phtml>
- [14] Plato, *Cratylus*, edited by John Burnet (1903) <http://www.perseus.tufts.edu/hopper/text?doc=Perseus:text:1999.01.0171>
- [15] Benjamin Lee Whorf, *Language, thought and reality* (1978)
- [16] *The Universals Archive* <http://typo.uni-konstanz.de/archive/intro/>
- [17] *The World Atlas of Language Structures* <http://wals.info/>, feature 41A: distance contrasts in demonstratives, feature 136A: M-T pronouns, feature 137A: N-M pronouns
- [18] Gérard Diffloth, *i: big, a: small*, in: *Sound Symbolism*, pp. 107–114 (1994)
- [19] Kong-On Kim, *Sound symbolism in Korean*, *Journal of Linguistics* 13(1):67–75 (1997)
- [20] Russel Ultan, *Size-sound symbolism*, in: *Universals of Human Language, Volume 2: Phonology* (1978)