

Mapping out semantic spaces in Elvish

Roman Rausch

1 Introduction

When learning a foreign language, any learner sooner or later discovers that concepts from his mothertongue do not exactly match up with concepts from the language he is learning. In fact, a perceptive learner will soon discover that even the simplest concepts do not quite match up if one goes into detail deeply enough.

Take for example the English word ‘drink’ and compare it with its Japanese counterpart which (as any dictionary will tell you) is *nomu*. Hence *mizu-wo nomu* ‘water-ACC drink’. However, ‘take medicine’ can be rendered by the same verb: *kusuri-wo nomu* (literally ‘drink medicine’); and it is even possible to say *tabako-wo nomu*, literally ‘drink cigarettes’ for ‘smoke cigarettes’ (Suzuki, 1984 [1])¹.

And the further one goes, the deeper it gets: *ashi* is both ‘foot’ and ‘leg’, while *te* ‘hand’ and *ude* ‘arm’ are separate; *koshi* is both ‘hip’ and ‘lower back’. There is a special word for ‘hot water’, *yu*; and *tsumetai* ‘cold to the touch’ is distinct from *samui* ‘cold of weather’. An elephant’s trunk is simply *hana* ‘nose’, there is no special word for ‘worm’ (*mushi* ‘insect’ is used both); similarly ‘bees’, ‘wasps’ and ‘hornets’ are all subsumed under *hachi*, while ‘nuts’ and ‘berries’ are subsumed under *kinomi*. The verb *oru* corresponds to either ‘fold’ (as in *origami*) or ‘break’ (usually of bones).

As long as one only deals with two languages, all of this can be learnt on a case-by-case basis, but chaos and confusion are expected to grow exponentially once more languages are added to the mix. However, exactly the opposite is actually the case. Human cognition luckily tends to go along similar channels and as more languages are compared, these universal patterns start to come to light and their variability actually turns out to be more limited than one might expect.

Namely, if one makes a survey across many languages, one will discover that the expression ‘drink cigarettes’ is used in a number of them: Boyeldieu (2008) cites several spoken across Africa: Arabic, Bedja, Hausa, Kasim, Laal, Ngbandi, Ngbugu, Sar, Yulu [5]. One can also add Modern Greek *πίνω (τσιγάρο)* to the list [8]. Thus, ‘drink’ in these languages might be defined as ‘swallow without chewing’ (which also explains the usage with respect to medicine, which comes in form of pills or liquids), while in English it might be defined as ‘swallow a liquid meant for hydration or taste’ (Suzuki, 1984 [1]).

But actually, we do not even have to define anything. We may simply represent this result in a graphical fashion in terms of a *semantic map* (Anderson 1982 [4], Haspelmath 2003 [2]). The recipe for this is the following:

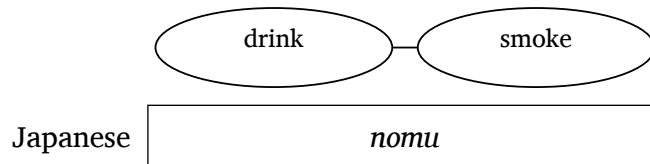
1. If a concept is distinguished in a number languages, we put it on different nodes of a graph².
2. If it is not distinguished in others, we connect them by edges.

To talk about the latter case, we may say that the concepts are “colexified” or that they are “polysemous”³. In the case of ‘drink’ vs. ‘smoke’ we get the following minimap:

¹This might be somewhat old-fashioned in the Japanese of today, but the example still serves to illustrate the point.

²‘Graph’ rather than ‘map’ would actually be the mathematical term for what is drawn here.

³There has been a considerable debate regarding how such semantic fields are represented in the brain. Roughly, the monosemist point of view proposes that there is one concept which is specified according to context. The polysemous point of view proposes that a word can really have different ‘meanings’ or ‘uses’ that stay separate. (The term ‘homonymy’ is usually used for accidental sameness.) Interested readers are referred to Pinker (2007) [9] and Klein, Murphy (2002) [10]. In order to step away from this debate, some authors also switch up the terminology: from ‘polysemy’ to ‘colexification’, from ‘meaning’ or ‘use’ to ‘function’. In any case, the power of the semantic map technique actually permits us to draw insights without committing to any of the views.

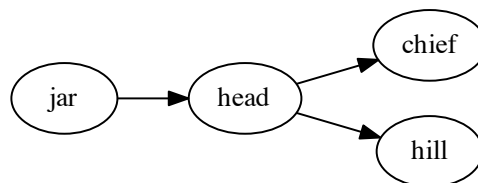


In the simple case of pairwise polysemy, we may also just use pointy brackets in the text to indicate it:

<drink, smoke>

But a graphical representation becomes much more useful once we get more nodes. The fact that not every node is connected to every other one teaches us along which lines human cognition proceeds.

This technique can also be extended in many ways. One can use roots with some leeway rather than exact words, put weights on the edges according to how common certain polysemies are, or add arrows to make the graph directed to indicate diachronic semantic shifts. For example, the word ‘jar, cup’ often becomes the word ‘head’ (as Latin *testa* ‘jar’ > Italian *testa*, French *tête* ‘head’ or Latin *cuppa* ‘cup’ > German *Kopf* ‘head’). Furthermore, the word ‘head’ is often used in the sense ‘hill’ or in the sense ‘chief’ [8]. A graph reflecting that would look like this:



The aim of this article is to apply the semantic map technique to Tolkien’s languages and see what insights can be gleaned from it.

In doing so, one is inevitably confronted with the question whether the Elvish languages can be counted as human languages at all. It should be obvious that they are at least almost the same, sharing most of their characteristics. But we cannot expect Elvish cognition to exactly match the human one, and so we expect the semantics to be different as well.

The comparison is partly carried out with published papers on polysemy, partly with my personal observations based on the languages familiar to me or examples encountered during my research. The work of List and al. (2013) [6] has proven particularly useful, being a large compilation of polysemies that can be searched with a query script. Another large compilation is *The Catalogue of Semantic Shifts* [8]. Finally, another very helpful source was William Annis’ meta-compilation of semantic maps for the purpose of conlanging [7].

The article is divided into three parts: Case studies of vocabulary, case studies of grammar and a particularly interesting semantic map of the word ‘breath’.

2 Case studies: vocabulary

2.1 Tree, wood, forest

The first semantic map concerns the words ‘tree’, ‘wood as material’, ‘firewood’, ‘small forest’ and ‘large forest’. They constitute a linear map, where neighbouring words can be colexified. Spanish distinguishes all of the five concepts, while other languages distinguish only two to three. This observation goes back to the work of Hjelmslev (1943) [3], thus way before the semantic map technique was formalized as such (Anderson 1982 [4], Haspelmath 2003 [2]).

	tree	wood (material)	firewood	small forest	large forest
Danish	<i>træ</i>		<i>skov</i>		
German	<i>Baum</i>	<i>Holz</i>	<i>Wald</i>		
Russian	<i>дерево</i>		<i>дрова</i>	<i>лес</i>	
French	<i>arbre</i>	<i>bois</i>			<i>forêt</i>
Spanish	<i>árbol</i>	<i>madera</i>	<i>leña</i>	<i>bosque</i>	<i>selva</i>

Comparing this to Elvish, it would seem that a slight modification needs to be made: There appears to be a frequent distinguishing of ‘small tree’ (Q. **alda**) and ‘large tree’ (Q. **orne**) rather than of small and large forests. Otherwise, we see indeed that ‘firewood’ has separate words in Goldogrin and early Qenya. Later, an explicit gloss ‘firewood’ is not found, it could be that Q. **toa**, S. **taw** ‘wood as material’ can be used in this sense.

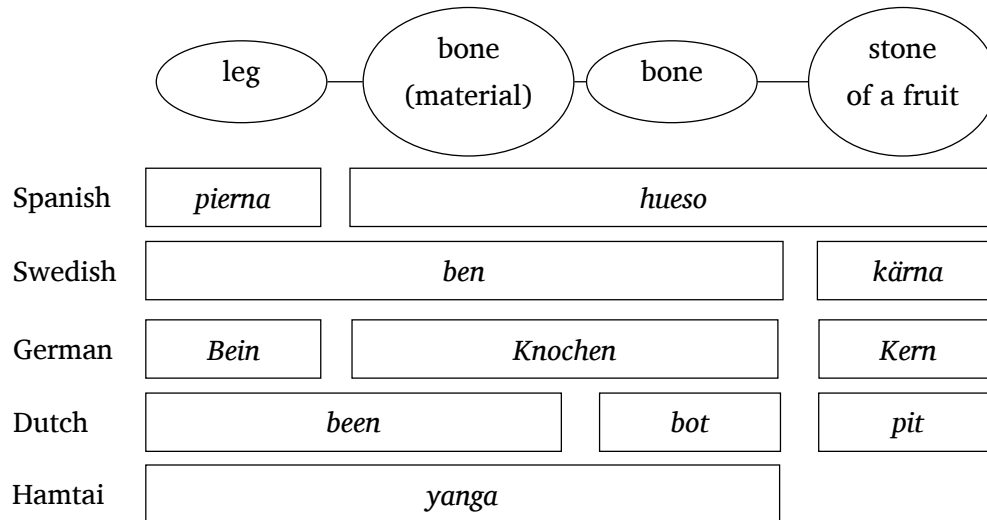
	tree	wood (material)	firewood	forest
	broad tree			
	tall tree			
G.	orn, †alwen	âl	tan, lôg	tavros, taur, †alos/aloth, goloth
EQ {	alda	turu	tundo	mālos, taure, tauno
	taule			
N./S.	galadh	orn	taw	tawar, taur, glad, eryn
			?	
Q.	alda	orne	toa	taure

References:

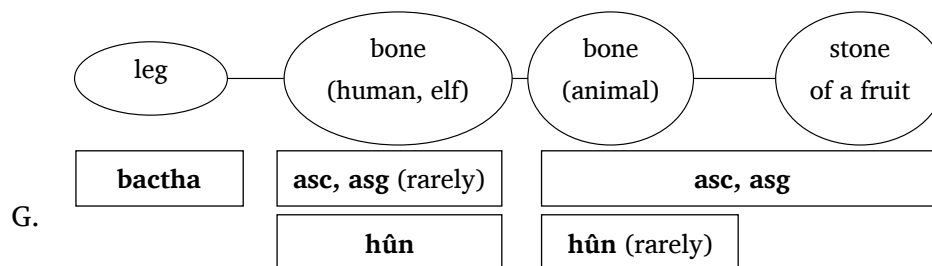
- Goldogrin: **orn** ‘tree’ (PE11:62), **†alwen** ‘tree’, **âl** ‘wood (material)’ (PE11:19), **tan** ‘firewood’ (PE11:69), **lôg** ‘fuel, firewood’ (PE11:52), **tavros** ‘forest, wooded land’ (PE11:69), **taur** ‘dense wood or forest’ (PE11:69), **†alos**, **†aloth** ‘forest’ (PE11:19), **goloth** ‘forest’ (PE11:41)
- early Qenya: **alda** ‘tree’ (PE12:29), **taule** ‘great tree’ (PE12:90), **turu** ‘properly = firewood, but used of wood in general as material’ (PE12:96), **tundo** ‘firewood, fuel’ (PE12:96), **mālos** ‘forest’ (PE11:41), **tauno** ‘forest’ (PE12:90), **taure** ‘forest’ (MC:213)
- Noldorin/Sindarin: **galadh** (Silm:*alda*, PE17:25), **orn** ‘tall tree’ (PE17:25), **taw** ‘wood as material’ (PE17:115), **tawar** ‘forest’ (UT:319), **taur** ‘forest, wood’ (Silm:*taur*, PE17:82), **eryn** ‘forest’ (PE17:33), **glad** ‘wood’ (UT:153)
- Quenya: **alda** (Silm:*alda*, PE17:25), **orne** ‘tall tree’ (PE17:25), **toa** ‘wood as material’ (PE17:115), **taure** ‘forest’ (Silm:*taur*, PE17:80)

2.2 Leg, bone, stone

This map is based on a personal observation that some languages colexify the ‘stone of a fruit’ (of peaches, plums etc.) with the word for ‘bone’ (or its diminutive). On the other hand, ‘bone as material’ may be separate from the physiological bone. Finally, some languages show a colexification of ‘bone’ and ‘leg’. This is a particular feature of Germanic languages, but also seems to appear in such an unrelated language as Hamtai (Papua New Guinea) [11] and as a semantic shift between Archi and Aghul (spoken in Dagestan) [8].



In Elvish, Goldogrin **asc, asg** shows the same colexification of ‘bone’ and ‘stone of a fruit’. Furthermore, there is a distinguishing between animal bones (**asc, asg**) and humanoid bones (**hûn**, of Elves and Men) that I could not find in natural languages.



References:

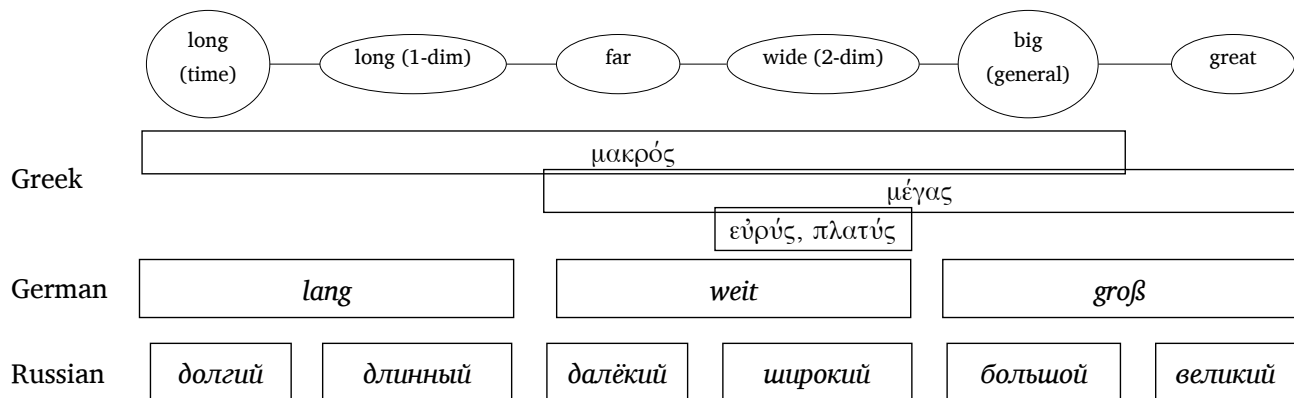
- Goldogrin: **bactha** ‘leg’ (PE11:21), **asc, asg** ‘(1) bone (especially of other animals, rarely of man), (2) stone of fruit’ (PE11:20), **hûn** ‘bone (especially of men and eldar)’ (PE11:49)

2.3 Long, far, wide, big, great

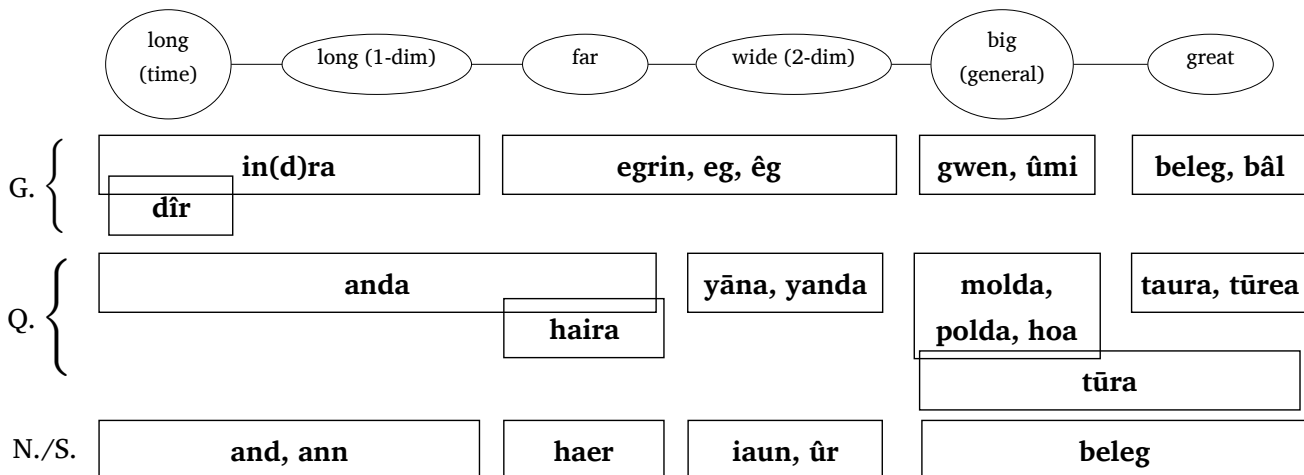
This map is based on personal observations as well, and deals with extensions in various spatial directions. We have the three basic concepts ‘long’ (one-dimensional), ‘wide’ (two-dimensional) and ‘big, large’ (general). Normally, here seems to be no word that is limited to three-dimensional objects, the word for ‘big’ can be used for any kind of shape. Since time is conceptualized as one-dimensional, it is natural to colexify spatial length and duration. Furthermore, the concept ‘great, of high value or importance’ may be colexified with ‘big’. Finally, the concept ‘far’ tends to be between ‘long’ and ‘wide’.⁴ The following figure shows a comparison of some languages⁵:

⁴Interesting extensions of the map that are not considered here would be the addition of ‘high/deep’, ‘tall’ and ‘loud’ (see below for the latter).

⁵The Greek words are somewhat difficult to untangle, but it is clear that *μακρός* tends to be on the left side of the map and *μέγας* on the right, with the usages getting more marginal the further one goes towards the middle.



We find similar patterns in Elvish: While ‘long (time)’ and ‘long (space)’ are usually colexified, Goldogrin also has a separate word **dîr** for the former. It also colexifies ‘far’ and ‘wide’. And even though there is a separate word for ‘far’ in Quenya, **anda** shows a colexification with ‘long’. Finally, Q. **tûra** and S. **beleg** colexify ‘big’ and ‘great’. The notes given on page PE17:115 are particularly interesting and show how Tolkien is busy dividing the semantic space of ‘wide’, ‘big’ and ‘great’ for the Quenya and Sindarin words.



References:

- Goldogrin: **in(d)ra** ‘long (also used of time)’ (PE11:51), **dîr** ‘long (of time only)’ (PE11:30), **egrin** ‘wide, vast, broad, far’, **eg, êg** ‘faraway, wide, distant, far off’ (PE11:32), **gwen** ‘large, big, fine’ (PE11:45), **ûmi** ‘large’ (PE11:74), **beleg** ‘mighty, great’ (PE11:21), **bâl** ‘worthy, important; great, mighty’ (PE11:21)
- Quenya: **anda** ‘long, far’ (PE17:90), in respect to time: **andave laitivalmet** ‘long will we praise them’ (LotR), **haira** ‘remote, far’ (Etym:KHAYA), **yāna, yanda** ‘wide’, **molda** ‘big, large’, **polda** ‘big’, **hōa** ‘large, big’, **tûra** ‘big, great’, **taura, tûrea** ‘mighty, masterful’ (PE17:115)
- Noldorin/Sindarin: **an(d)** ‘long’ (Silme:an(d)), in respect to time: **andreth** ‘long suffering, patience’ (PE17:182), **haer** ‘remote’ (PE17:25), **iaun** ‘roomy, wide, extensive’ (PE17:42), **ûr** ‘wide’ (Etym:UR), **beleg** ‘mighty’ (Silme:beleg), ‘large, great, big’ (PE17:115)

2.4 Big, loud

To a native speaker of English, the word ‘loud’ might appear as a basic adjective that can be found in other languages, but this is not necessarily the case. Latin, Ancient Greek and Japanese are among the languages that do not have a word for ‘loud’, but still express the concept in essentially the same way using ‘big’: Latin *magnā vōce*, Greek *ἐν φωνῇ μεγάλῃ*,

Japanese *ōkii koe de* ‘with a big voice’. Sindarin does seem to have a word for ‘loud’, ***brui** from **Bruinen** ‘Loudwater’ (LotR), but it is a hapax legomenon, with no other attestation or root to corroborate it. No word for ‘loud’ is attested in Quenya.

2.5 Beard, chin

A peculiar colexification can be found in Goldogrin, where **beg** is glossed ‘beard, chin’ (PE11:22). A search in List et al.’s database shows that this polysemy is shared by a few languages, notably within the Slavic family: Bulgarian *bradā*, Czech, Slovak, Slovenian *brada*, Eastern Yiddish *bord*, but also Albanian *mjekër*, as well as non-Indo-European Chamalal (NE Caucasian) *muzamIḡ*, Quechua *k’aki* and Rotuman (Austronesian) *kum-kumu*, *kum-kumi* [6] – all meaning both ‘beard’ and ‘chin’.

This polysemy is so specific that I suspect that Tolkien may have encountered it somewhere in a Slavic dictionary.

2.6 Clouds

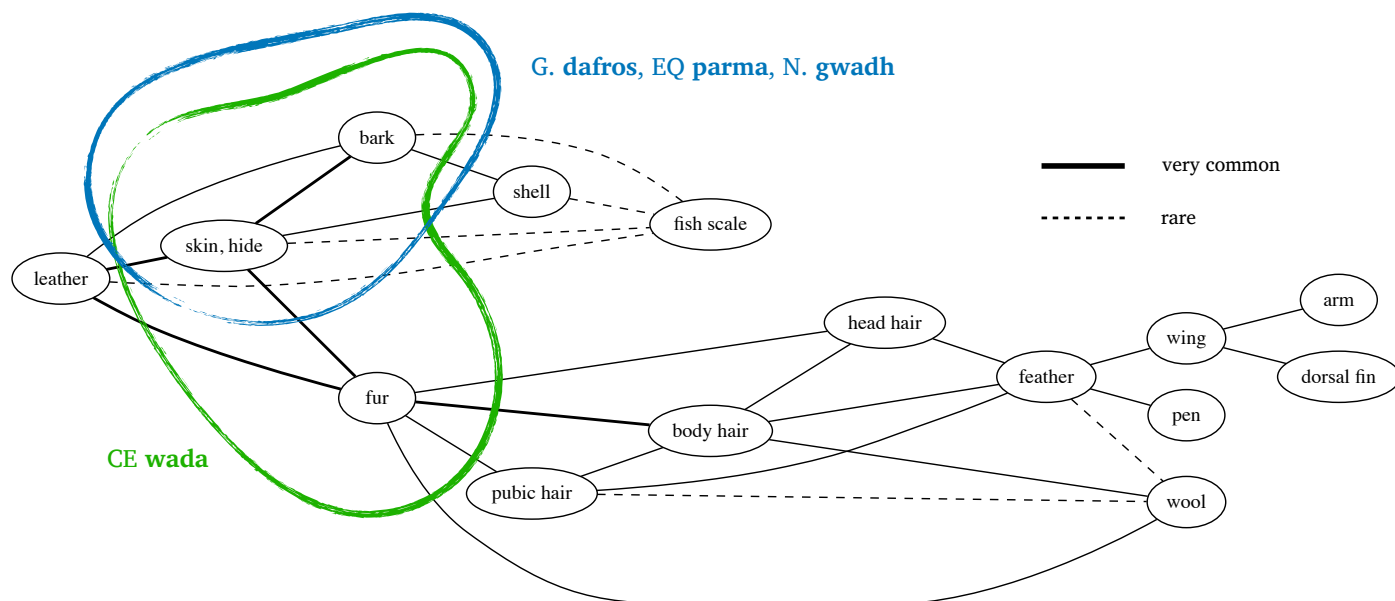
This is another colexification pair that is based on a personal observation. Russian has a separate word for ‘black cloud’ (*муча*), referring to stormy clouds which bring the rain. The general word for ‘cloud’ is *облако*, and it becomes ‘white cloud’ in contrastive usage. A similar separation is found in literary French (*nuage* (white) vs. *nuée* (black) [8]) and in Maori, which in general has a very large amount of words referring to clouds.

	white cloud	black cloud
English	cloud	
Russian	облако	муча
Maori	ao	atiru, paroro,...
Quenya	fanya	lumbo

Quenya seems to make the same distinction: **fanya** is glossed ‘(white) cloud’ (PE17:26), said to be ‘used [...] only of white clouds [...] not used of storm clouds’, while **lumbo** is glossed ‘dark, shade’ (PE17:168) ‘gloom’ (PE17:72), but is used in the sense ‘cloud’ in the *Markirya* poem (MC:223), where it is clear from the context that storm clouds are meant. This is more explicit in the early sources, where we find the translation ‘dark lowering cloud’ (PE12:57).

2.7 Outer shells

This is a much larger semantic map (List et al. 2013 [6]) of what could be called ‘outer shells’, namely words like ‘skin, bark, leather, fur’. Via ‘fur’ this is also connected to various words for ‘hair’ and ‘wool’, which in their turn connect to ‘feather’ and from there also to anatomical ‘wing’ and utilitarian ‘pen’. Of course, no single word from any language spans across all of these semantic labyrinths.



In early Elvish, we find that G. **dafros** ‘bark, skin, peel’ (PE11:28), EQ **parma** ‘skin, bark; parchment; †book, writings’ (PE12:72), N. **gwadh** ‘bark, skin, peel’ (PE13:146) show the polysemy between ‘skin/hide’ and ‘bark’. ‘Peel’ unfortunately does not appear in List et al.’s data. Furthermore, if we take the reflexes of Common Eldarin (s)**wada** (Q. **fara** ‘fur, fell’, T. **suada** ‘hide’, Ilk. **swat** ‘bark’ (PE13:146)), one can also add ‘fur’ to the mix. Interestingly, neither ‘shell’ nor ‘fish scale’ nor ‘leather’ appear here.

The <skin, bark> polysemy is exhibited by languages that are notably non-Indo-European as well as non-European, including Hawaiian *?ili*, Maori *hiako*, *peha*, Turkish *kabuk*, Paraguayan Guaraní *pire*, Ket (Siberia) *inolt*.

Another interesting point is the polysemy <bark, book> that reappears in later writings: Q. **parma** ‘bark, parchment, book’ (PE17:171). This is clearly peculiar to the history of Elvish writing where bark was the first writing material. Note that the polysemy <paper, book> is common in natural languages (List et al., 2013 [6]).

2.8 Wh-words

Interrogatives form an interesting area of semantic comparison. First of all, different languages may choose to lexicalize different semantic domains. For example, German *wievielt-* asks for a certain rank (for instance applicable to days in a month or sequential bus stops), a category which is not exactly mirrored in English. According to Cysouw’s rough analysis [12], the four most commonly lexicalized categories are PERSON (who?), THING (what?), SELECTION (which?), PLACE (where?). Less common are QUANTITY (how much?), TIME (when?), MANNER (how?). Other categories are incidental. In Quenya, the following interrogative categories are attested:

- THING, PERSON **man(a)** ‘who, what?’ (LR:63, MC:221, PM:395; PE23:99, 101–105)
- PLACE **masse** ‘where?’ (PE22:124–126; PE23:111)
- MANNER **manen** ‘how?’ (PM:395; cf. **mal(de)** in PE23:110–111)
- DIRECTION **manna** ‘whither?’ (PE22:124; PE23:112)
- ORIGIN **mallo** ‘whence?’ (PE22:124; PE23:112)
- QUALITY **manima** ‘what sort of?’ (PE22:124; PE23:106–107)
- EXTENT **malka** ‘how great?’ (PE22:124; cf. **man(de)** in PE23:110–111)

A great deal more were published in PE23, sometimes very fine-grained:

- POLAR QUESTION **ma** ‘whether?, is it the case that?’ (PE23:99, 101–104)

- SELECTION **mana** ‘which?’ (PE23:104–105)
- BINARY CHOICE **matwa** ‘which of two?, whether?’ (PE23:107)
- NUMBER **malli(ni)** ‘how many?’ (PE23:106, 108)
- SMALL EXTENT **manikka** ‘how small?, how little?’ (PE23:108)
- TIME **mallume** ‘when?, at what time?’ (PE23:109–110)
- DATE **manar(yas)** ‘on what day?’ (PE23:109–110)
- AGE, DURATION **mallúmea** ‘how old?, of what age?’ (PE23:108–109)
- MEANS **m(an)ainen** ‘by what means?’ (PE23:111)
- LIKENESS **manandon** ‘like what?, after what fashion?’ (PE23:111)
- REASON **maro** ‘why?, for what reason?’ (PE23:112)
- PURPOSE **mámen** ‘why?, for what purpose?’ (PE23:112)

Having fine-grained language-specific interrogatives that are often hard to match exactly between languages is not an unusual situation. Compare the impressive German list based on *wo* ‘where?’: *wobei, woraus, wodurch, worin, wofür, worüber, wogegen, worum, wohinter, worunter, womit, wovon, wonach, wovor, woran, wozu, worauf(hin), wozwischen*.

It turns out that in natural languages, the category **QUALITY** is actually seldom lexicalized and usually derived from other interrogatives. A notable exception is Latin *qualis* (Cysouw also cites Wardaman as another). One can see, however, how the frequent adjectival **-ima** can easily interact with interrogative **ma(n)-** in Quenya. Similarly understandable are **manna** and **mallo** that naturally follow from Quenya’s case system. This has parallels in natural languages with large case systems, such as Lezgian *hinaj* ‘from where?’ and *hiniz* ‘where to?’, continued with other cases as *hinal* ‘where on?’ (superessive); or Basque *nora* ‘where to?’, *nondik* ‘from where?’, *norantz* ‘in what direction?’.

Furthermore, Cysouw claims that the very presence of an interrogative element like Quenya **ma(n)-** or English *wh-* is not very common across the world’s languages except for Europe, even though it is not unheard of (cf. Japanese *d-*). Thus, many languages are completely happy with taking interrogatives from various unrelated sources without common initial sounds. On the other hand, a statistical analysis using the concept of entropy by Slonimska and Roberts [14] shows that interrogatives still tend to look more alike than randomly chosen words across a large sample of languages. It would seem that it is of practical use to have easily recognizable interrogatives as hints of an upcoming question.

Another point worth attention is that Quenya also appears to colexify the question words <who?, what?> (although this has to be taken carefully, as the examples come from different sources):

- Q. **Man kenuva fáne kirya?** ‘Who shall see the white ship?’ (MC:221)
- Q. **E man antaváro?** ‘What will he give indeed?’ (LR:63)

This is a highly unusual feature, as even languages which normally do not care much about animacy, usually make the distinction in this case. In fact, a distinction between <who?> and <what?> was labeled as ‘nearly universal’ by Ultan [15]. However, there is still a number of natural languages which do exhibit this polysemy, most notably Latvian *kas* and Lithuanian *kàs*, but otherwise non-Indo-European languages like Akhvakh (NE Caucasian) *чвеба*, Budukh (NE Caucasian) *uu*, Khvarshi (NE Caucasian) *зуба*, Maca (Paraguay) *lek*, Mocoví (Argentina) *nege?*, Toba (Argentina) *’nege?t*, Waurá (Brazil) *ka-i’ca*, *ka* (List et al. 2013 [6]).

Finally, according to Cysouw, the category **REASON** is almost universally derived from **THING** (cf. French *quoi, pourquoi*), less commonly from **MANNER**, **PLACE** or **TIME** and is lexicalized only in about 40% of languages. Before the attestation of **maro**, Neo-Quenya, Q. ***manen** ‘how?’ was often used in the sense ***‘why?’**. According to List et al.’s data, this is shared just by a few natural languages: Araona (Bolivia) *k’ipa*, Tsimané (Bolivia) *’hojnka*, Hmong Daw (SE Asia) *ua li cas*. With the attestation of **maro** we in fact have **-ro** as a specifically reasonal suffix; and furthermore **mámen** rather than ***manen** asking for **PURPOSE**. According to Cysouw, it is in fact not uncommon to have various words for **REASON** with slight differences relating to goal, cause, motivation et cetera.

What is interesting to point out is that the Quenya interrogatives seem to be completely consistently derived from a single root **ma-**. In natural languages, even if there is a common interrogative element like *wh*, one often finds at least one oddball form that breaks the pattern, like English *how*.

3 Case studies: grammar

3.1 Dative functions

This chapter follows Haspelmath (2003) [2]. English has two prepositions expressing the dative: *to* and *for*, but their functions are noticeably different. The former can be used to express the categories PURPOSE ('to see'), DIRECTION ('to London'), RECIPIENT ('gave it to me'), EXPERIENCER ('seems good to me'). It cannot be used to express the function BENEFACITIVE ('can you do it for/**to me?'). French, for example, has the prepositions *à*, the direct dative (pronouns like *lui*) and the preposition *pour*. While *à* mostly overlaps with English *to*, it is also used for the function PREDICATIVE POSSESSOR (*Ce livre est à moi* 'This book is mine'). The direct dative can be used for EXTERNAL POSSESSOR (*On lui a cassé la jambe* 'They broke his leg'), while *pour* is only used for the function of JUDICANTIS (*trop lourde pour moi* 'too heavy for me').

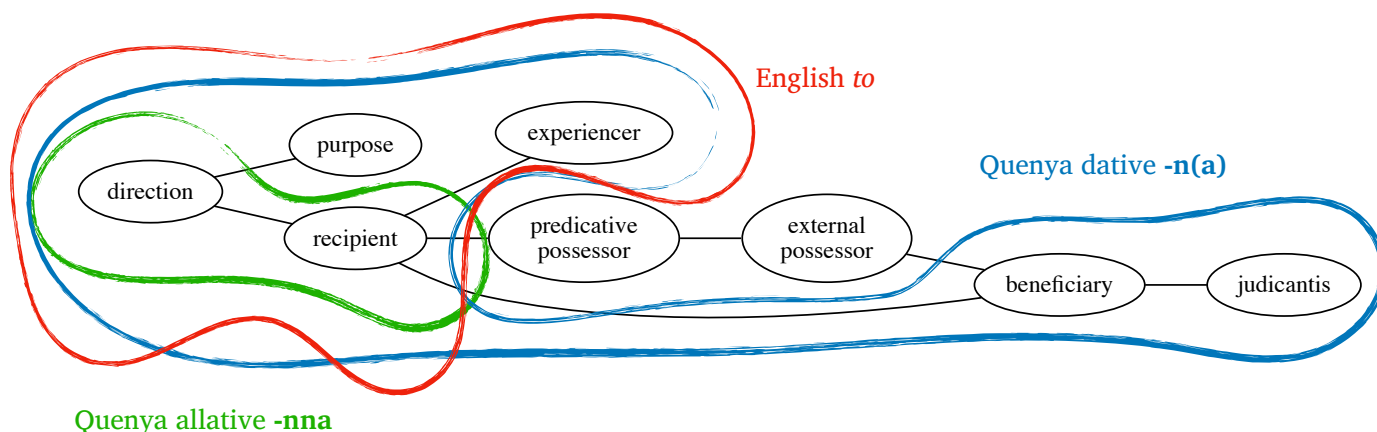
The following list shows the attested functions of the Quenya dative **-n(a)** and allative **-nna** cases:

- DIRECTION: **Endorennā utúlien** 'to Middle-earth I am come' (LotR, PE17:103), **antanen hatal sena** 'I cast a spear at him' (VT49:14)
- PURPOSE: **alkar enyalien** 'to commemorate the glory' (UT:305)
- RECIPIENT: **tultauvas sa nin** 'he will send it to me' (PE22:156), **Antanen parma sen(na)** 'I gave a book to him', **veryanen senna** '*I married/joined to him/her' (VT49:45)
- EXPERIENCER: **umbe nin** 'I have a feeling' (PE22:168), **ú tén(a) ulca** 'was not to them evil' (VT49:7)
- BENEFICIARY: **Man i yulma nin enquantuva?** 'Who shall refill the cup for me?' (LotR, RGEO)
- JUDICANTIS: **Nás mára nin** 'I like it', '*it is good for me' (VT49:30)

The categories EXTERNAL POSSESSOR and PREDICATIVE POSSESSOR do not appear and might be typically Indo-European.

The role of the ending **-na** is difficult to discern: It might be a longer variant of the dative, a shorter variant of the allative or a new suffix altogether. In **hríve ūva vēna** 'winter is drawing near to us' (PE22:167, VT49:14) one may argue that allative **-nna** is shortened after the long vowel **ē**. However, the vowel is short in **sena** (see above), so that **-na** seems to be a longer variant of the dative here, perhaps preferred for monosyllabic pronouns (compare also the change from **téna** to **tén** in the *Ambidexter Sentence* (VT49:14), suggesting they are equivalent) and apparently can also be used for the category of DIRECTION. Conversely, **senna** in the above examples shows the allative in the function of RECIPIENT, otherwise typically occupied by the dative case.

On the semantic map, the ranges of the two case endings can thus be delimited as follows:



Note that the allative also tends to occupy the neighbouring category of PURPOSE, as seen in the deleted **nie karienna** 'I am for/towards making it' (said to be a close equivalent of English 'I am to make') (PE22:123).

3.2 Coordination

This section follows the monograph *Coordination Relations in the Languages of Europe and Beyond* by Caterina Mauri [13]. She identifies the following categories for coordinating conjunctions:

- SEQUENTIAL: “I opened the door and went away”
- SIMULTANEOUS: “He is dancing and clapping his hands”
- ATEMPORAL: “Doctors are rich, and lawyers marry pretty girls”
- OPPOSITIVE: “I bought shoes, whereas Sue found a skirt”
- CORRECTIVE: “He did not run, but walked slowly”
- COUNTEREXPECTIVE: “John is tall, but he is not good at basketball”

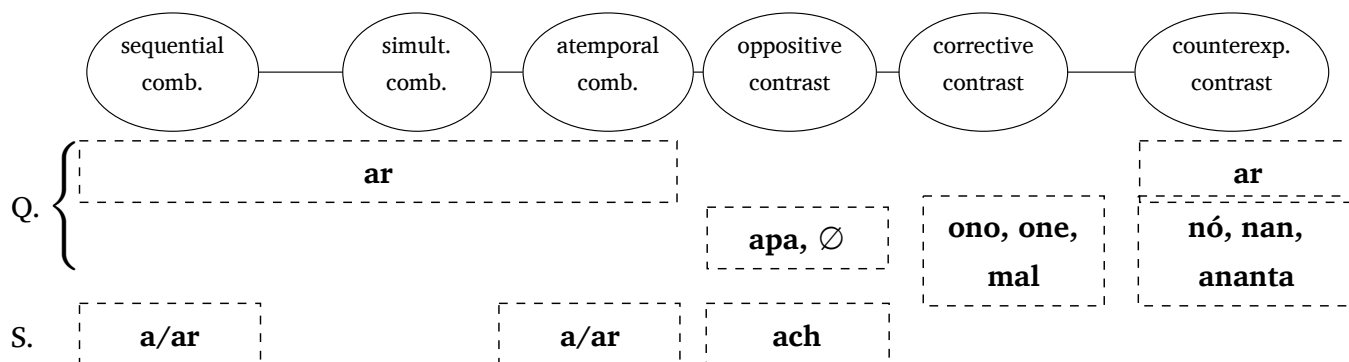
They form a one-dimensional semantic map, with some examples from natural languages given below (the reader is referred to the comprehensive original work for further details). The empty set symbol $\emptyset$ means that the language simply uses no word at all for a given function:

	sequential comb.	simult. comb.	atemporal comb.	oppositive contrast	corrective contrast	counterexp. contrast
Mangarayi	$\emptyset$					<i>gana</i>
Tuvaluan	$\emptyset$	<i>kae</i>				
Hausa	<i>kuma</i>					<i>amman</i>
English	<i>and</i>		<i>whereas</i>		<i>but</i>	

We find the following in Quenya and Sindarin:

- SEQUENTIAL: Q. **ar Sauron túle...** ‘and Sauron came...’ (SD:246), **Ilu Ilúvatar en káre eldain a firimoin ar antaróta...** ‘the Father made the World for Elves and Mortals and he gave it...’, S. **Pedo [...]** **a minno!** ‘Speak and enter!’ (LotR)
- SIMULTANEOUS: Q. **Ámen anta [...]** **massamma ar ámen apsene [...]** ‘Give us our bread and forgive us [...] (VT43:12)
- ATEMPORAL: **Savin Elessar ar i náne aran Ondóreo** ‘I believe that Elessar really existed and was a king of Gondor’ (PE22:158, VT49:27)
- OPPOSITIVE: Q. **malle téra lende númenna, ilya sí maller raikar** lit. ‘road straight went Westward, all now roads bent’ (SD:310), Q. **melinyes apa lá hé** ‘I love him but not him’ (VT49:15), S. **i glinn hen agorer Edain [...], ach hí in Ellath îr ed epholar** *‘Men made this song [...], but now only the Elves ?remember it’ (VT50:16)
- CORRECTIVE: **Álame tulya úsahtienna mal áme etelehta ulcullo** ‘Do not lead us into temptation, but deliver us from evil’ (VT43:12) [Earlier versions of the *Pater Noster* have **ono, one.**]
- COUNTEREXPECTIVE: Q. **ore nin caritas, nó namin alasaila** ‘I would like to do so but judge it unwise’ (VT41:13), **nan úye sére indo-ninya símen** ‘but my heart resteth not here’, **ananta úva táre fárea** ‘but yet it will not be enough’ (LR:72), **i-ner ne raiqa ar sí aphastat** ‘the man was angry, but now is in good humour’ (PE22:124)

Here we are faced with the problem that an attestation in a given context does not tell us about other possible uses, so that the boundaries on the map are not known exactly. To indicate this, I draw dotted lines:

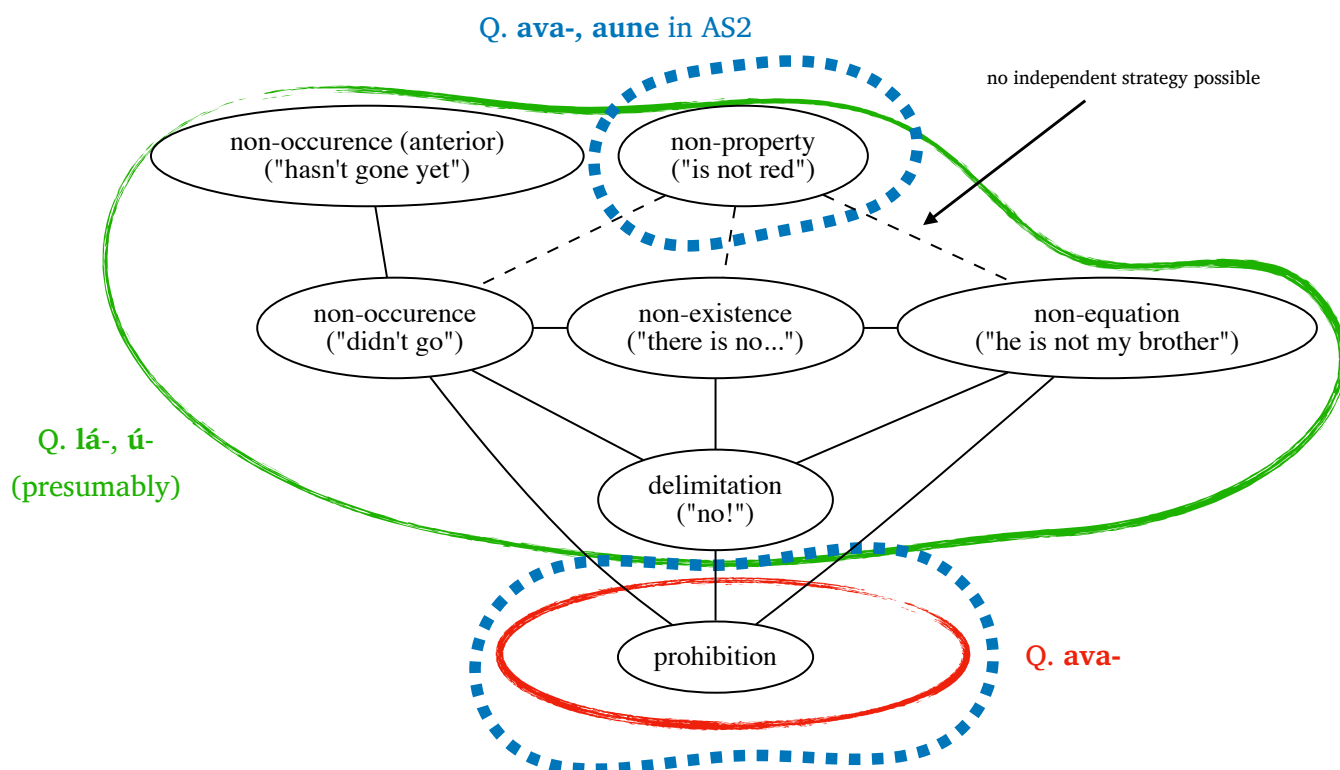


Interestingly, Q. **ar** is attested with the first three functions as well as the last one, COUNTEREXPECTATIVE. Given the power of semantic maps, we would therefore expect that it also has all the functions in between – that is if we believe that Quenya behaves like a human language in this respect.

There is also an implication for Neo-Quenya, where one has to bridge the attestation gaps. This can be done in a number of ways, but in order to remain consistent with the semantic maps, we have some additional restrictions. On the other hand, we do not have to relex the English system. Hence, there may be more than one word for ‘and’ and more than one word for ‘but’, neither of which exactly coincides with the English words in its usage. So for example, Q. **apa** could be confined to OPPOSITIVE only, or OPPOSITIVE and ATEMPORAL, or OPPOSITIVE and CORRECTIVE overlapping with **mal**, and so forth.

3.3 Negation

A consistent idea that Tolkien implemented into Elvish is that prohibition (root BĀ/ABA-) should be expressed differently from other forms of negation (roots Ū or LA). The following semantic map of negation is taken from Bond (2009) [16] and we find indeed PROHIBITION as a separate node, meaning that other natural languages may also single out this function.



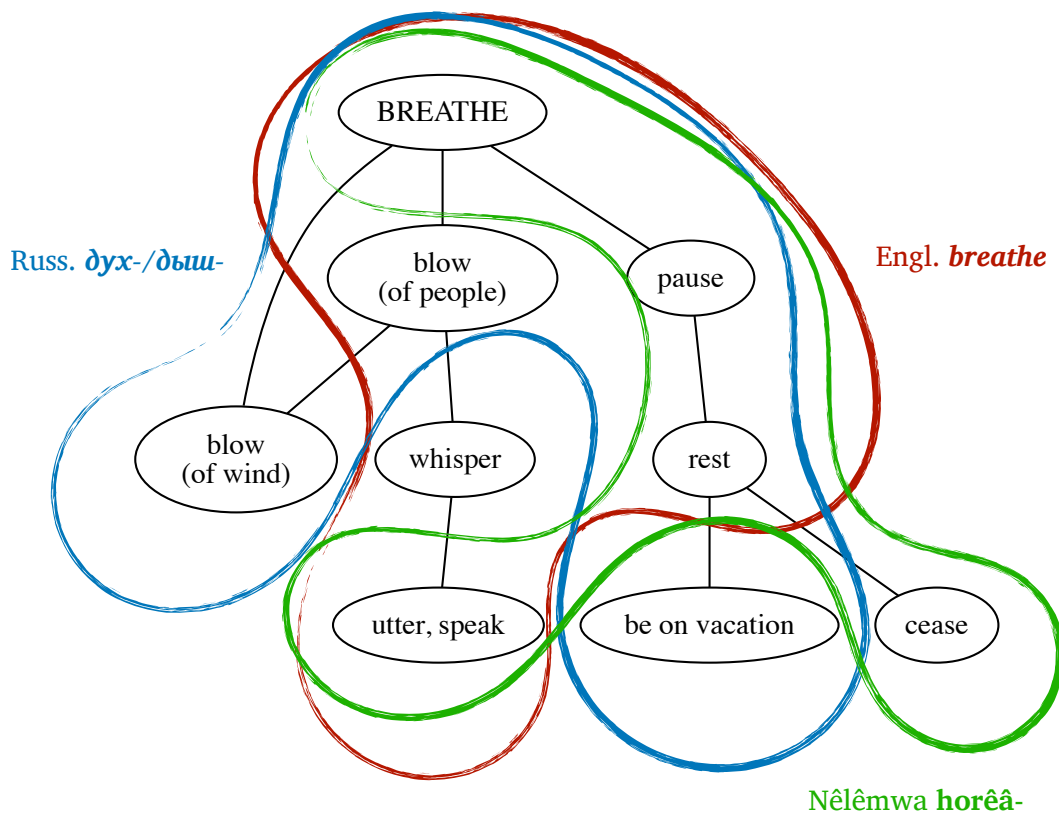
Though ABA- or Q. **va-** is predominantly used for PROHIBITION (**áva care** ‘don’t do it’ (WJ:371)), we interestingly also find the root in the function of NON-PROPERTY in *Ambidexter Sentence 2* (VT49:6): **hyarmen aune** “**sinister**” ‘left had no sinister connexions’ (with **aune** probably being the past tense of **ava-** (**abnē*), (VT49:6)), suggesting that Tolkien was at some point contemplating loosening the semantic constraint. AS NON-PROPERTY is a node that is far away from PROHIBITION however, we would expect that some of the intermediate functions should be covered as well in this case.

4 The semantic map of ‘breath’

There is one particular semantic map that is of special interest in regard to Tolkien’s languages, namely the semantic map of words related to ‘breath’ or ‘breathe’. As it happens, it is also a very large map that spans many categories (François, 2008 [17]). To appreciate its scope and significance, I will first discuss it in greater detail than the maps before.

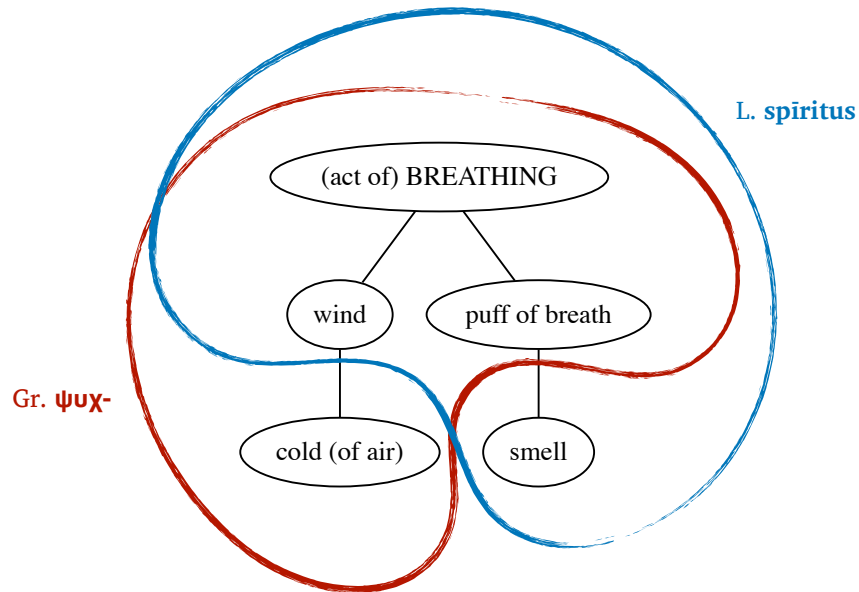
4.1 Natural languages

Let us consider the verb first. On the one hand, the action of breathing can be linked to the blowing of wind. On the other, it may be connected to ‘whispering’ and from there to ‘speaking’ in general (as English ‘to (not) breathe a word’, however marginal the usage). In a third extension, breathing can start to signify a pause (English ‘a breather’), rest in general and then even ‘(to be on) vacation’ (as Russian *отдых*). Finally, resting can also start to mean ‘ceasing’, as is the case in Nêlêmwa (South Pacific), where ‘stop beating the dog’ is literally expressed as ‘breathe beating the dog’. This part of the map is illustrated in the following figure:

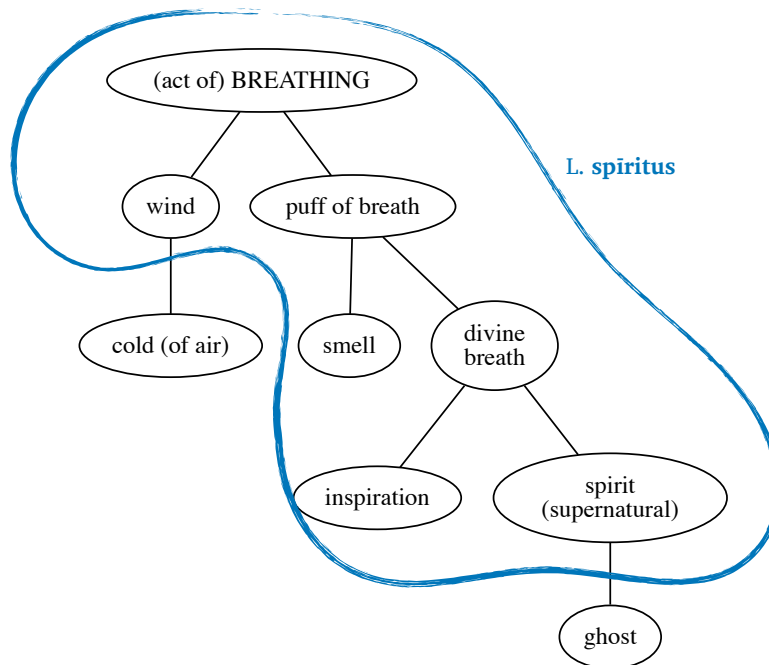


Let us now turn to the noun. The development of ‘breathing’ can be nicely demonstrated by the highly influential Latin words *spīritus* and *anima*, as well as the Greek root *ψυχ-* and the noun *πνεῦμα* – all literally ‘breathing’, but also so much more beyond that.

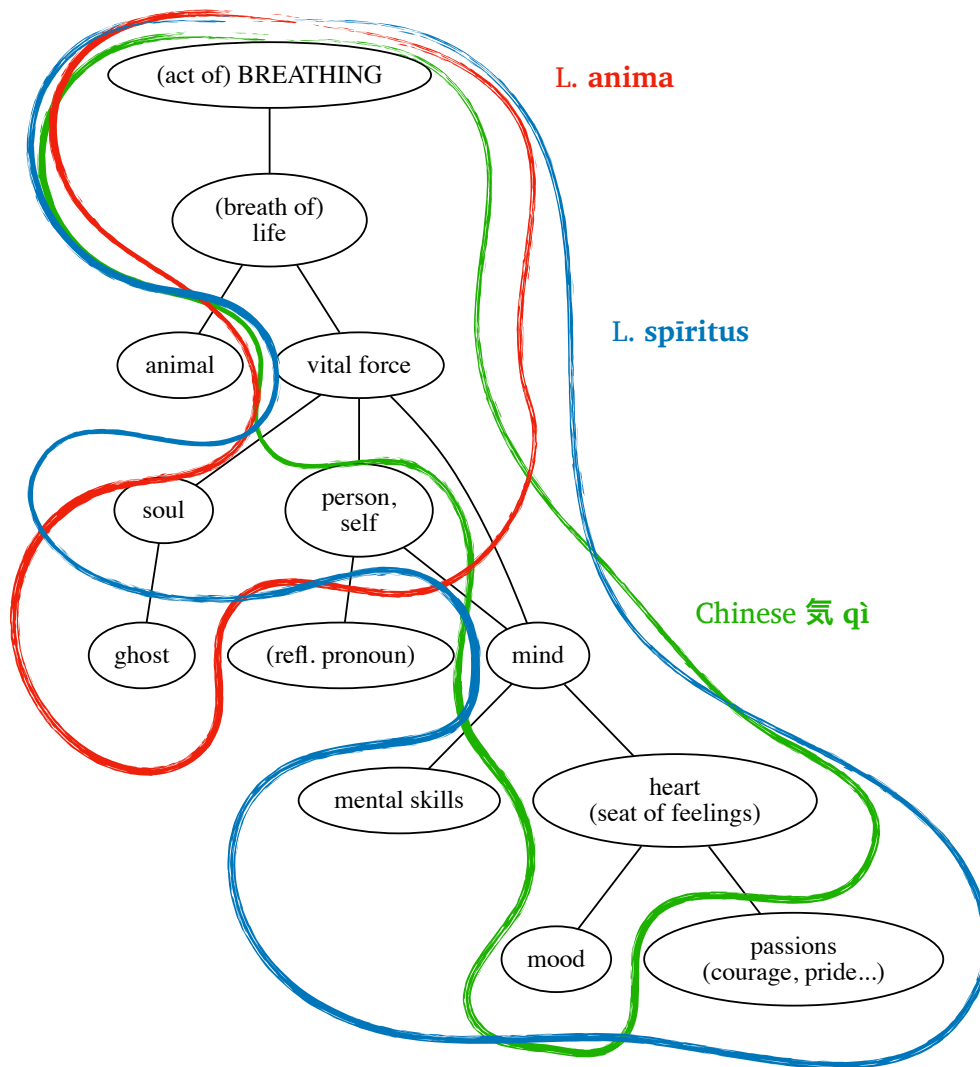
In this part of the map, ‘breath’ can be linked to ‘wind’ as well (and from there to ‘cold’). Likewise, the act of breathing can develop into ‘smell’. The following figure demonstrates it:



However, this is just the beginning, as 'breath' can also start to signify 'divine breath', a source of 'inspiration' on the one hand and related to supernatural 'spirits' or 'ghosts' on the other. The extended range of usage of Latin *sp̄iritus* is demonstrated in the following figure (and the obvious English loans 'inspiration', 'spirit' and 'sprite' serve as a further illustration):

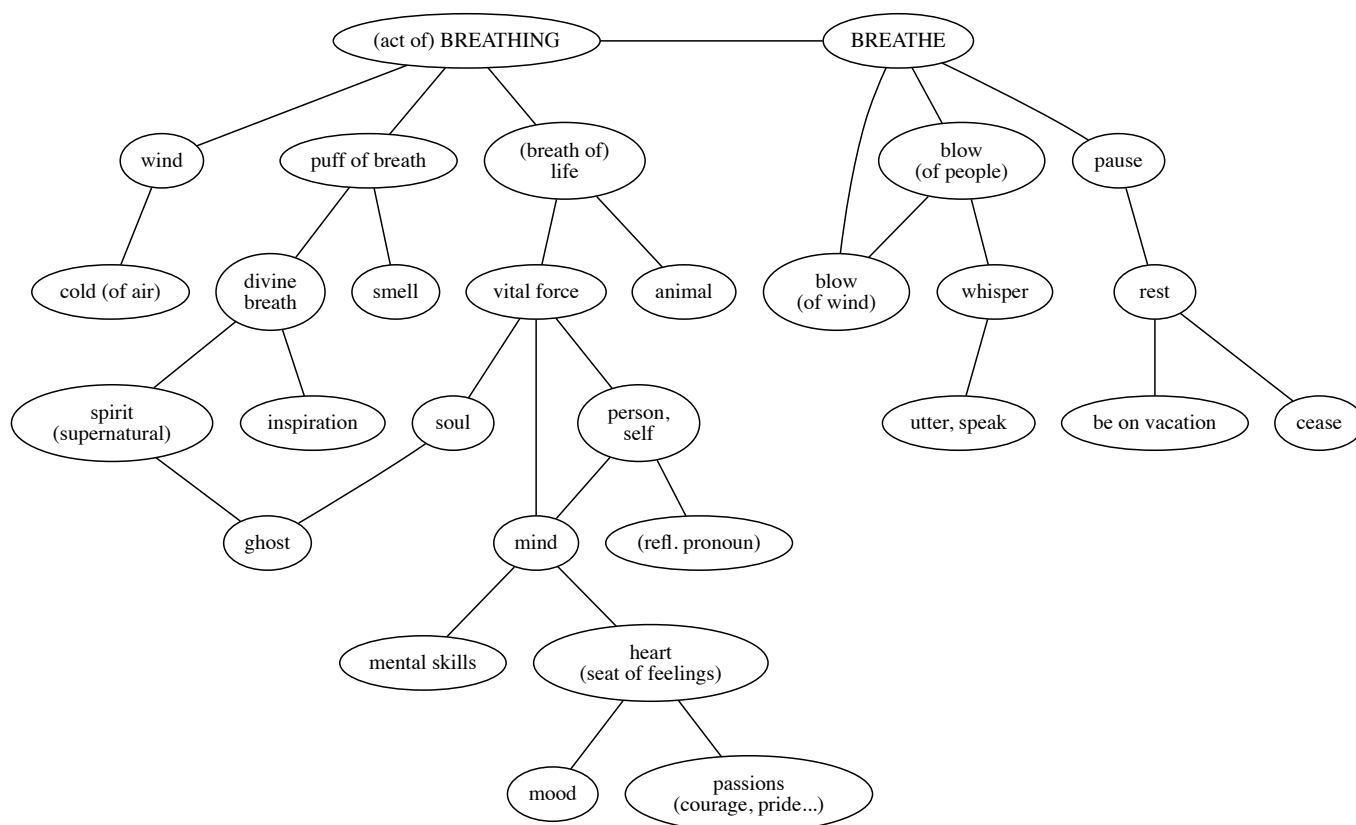


An even larger part of the map stems from breath being an integral part of living, leading to the extension to 'breath of life'. This can then be used in the sense 'animal' or acquire a more general meaning 'vital force'. (For example, chi, the mental energy from Chinese culture (Pinyin transcription *qì*; in Japanese *ki*) also goes back to 'breath'.) From here, it can start to mean 'soul' and in particular a disembodied 'ghost' (joining with the other branch). Another sub-branch goes via 'self' to the reflexive pronoun (as Sanskrit *ātman*). Finally, from the significance 'mind' it can branch out into various mental skills, feelings, passions and descriptions of mood. The ranges of Latin *sp̄iritus*, *anima* and Chinese *qì* are demonstrated in the following figure. Greek *πνεῦμα* occupies a similarly large part, as does the Arabic root *rūḥ*-, but they are left out for reasons of space.



Finally, all the sub-maps can be stitched together into a complete one⁶:

⁶I wonder if more nodes are waiting to be discovered. For example, the development of 'spirit' to mean 'alcohol' as well as Russian *дыху* meaning 'perfume' might not be isolated usages. Another extension might be 'to die' from 'breathe out, expire', as is found in Russian *издыхать* and Q./S. **fir-** 'expire, fade, die' (VT43:34, MR:250).



4.2 Tolkien's languages

The semantic map of 'breath' is of particular interest for Elvish because of the work of Owen Barfield who in 1928 wrote a book called "Poetic Diction" [18]. The book is philosophical rather than linguistic in nature, so a summary that does it justice is very difficult to give. I would say that it draws a parallel between polysemy (though never named like this, of course) and poetry. One idea is that poetry allows for a glimpse into someone else's mind and the way that person views the world. And a different language with a different division of semantic spaces also reflects a particular worldview one can find poetic pleasure in.

Another idea is the claim that languages underwent a "crystallization", by which he seems to mean a shift from high to low polysemy. And this is demonstrated using the word "breath":

We must, therefore, imagine a time when 'spiritus' or πνεῦμα or older words from which these had descended, meant neither breath, nor wind, nor spirit, nor yet all three of these things, but when they simply had their own old peculiar meaning, which has since, in the course of the evolution of consciousness, crystallized into the three meanings specified [...]

Barfield was also a casual Inkling and knew Tolkien, though the two were not close friends. According to C.S. Lewis, Tolkien told him that Poetic Diction "had modified his whole outlook" [20], [19]. This has been taken to mean a lot and a reprint of the book even had the tag "The seminal text that inspired Tolkien and C.S. Lewis" on the cover (to boost book sales, I presume). It is, however, difficult for me to see what was exactly modified in Tolkien's approach. He clearly understood semantics very well even before his contact with Barfield and long before "Poetic Diction" was published, as evidenced by the rich semantic spaces in the earliest examples of Qenya and Goldogrin. It will be seen that 'breath' does indeed enter the semantics more prominently at a later point than in the earlier material, but I would say that this is more like an addition of a stroke to a landscape rather than a complete redrawing of the landscape itself.

Be it as it may, I leave this matter to philologists and historians. I would just like to note that Tolkien was very well aware of the semantic complexity associated with the word 'breath' and it is interesting to see how this has reflected in his own work⁷.

⁷The last part of "Poetic Diction" deals with the semantics of the English word 'ruin' and its historical development, but an interesting parallel

4.2.1 Qenya and Goldogrin

Looking into the *Qenya Dictionary* (PE12), we find two roots related to air or aerial motion that show an extent not unlike the map for ‘breath’. The first one is *gwili-* (PE12:104) with the following derivatives:

- Q. **wili-** ‘sail, fly, float’, **wilin** ‘bird’, **wilwarin** ‘butterfly’

The second one is *vili-* (PE12:101) with the following derivatives:

- Q. **vīle** ‘breeze - gentle’, **Vilya** ‘air (lower)’, **vilina** ‘airy, breezy, light’

In the *Gnomish Lexicon* (PE11), we even find three related roots, the first being *gwil-* (PE11:45):

- G. **gwil-** ‘sail, float, fly’, **gwilbrin** ‘butterfly’, **gwilith** ‘breeze’, **gwail** ‘air’, G. **gwailtha-** ‘to air, expose to air’

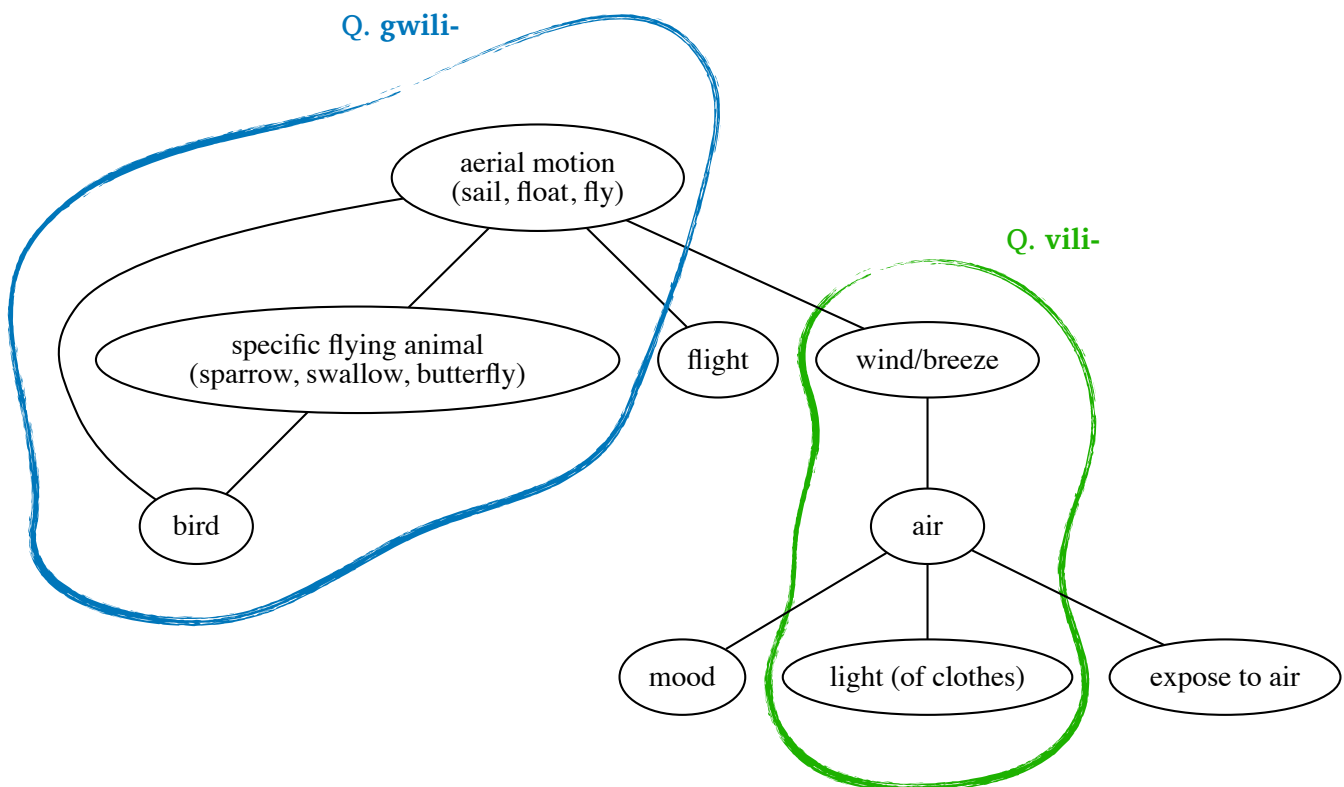
The second is *duil-* (PE11:31):

- G. **duila-** ‘to fly’, **duil** ‘flight, power of flight’, {**duil** ‘swallow’}

And the third is *bil-* (PE11:23):

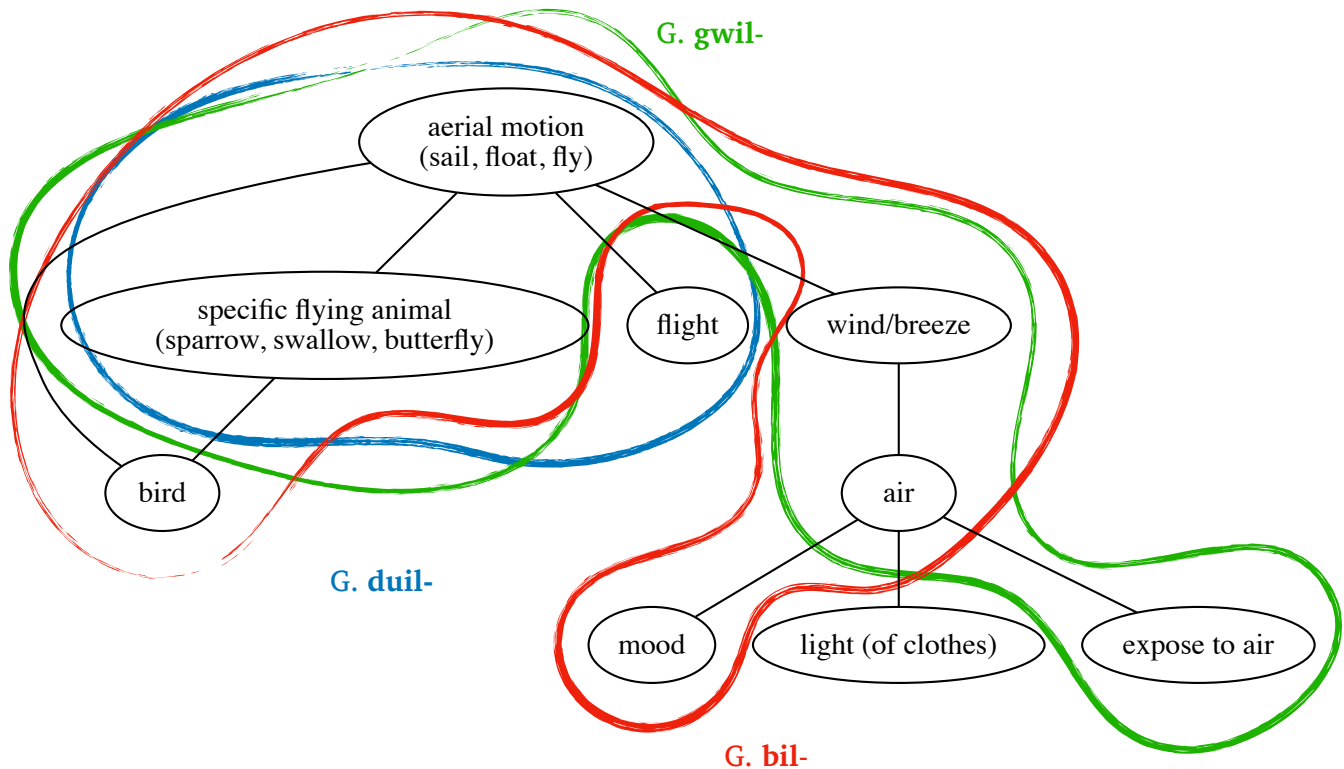
- G. **biltha-** ‘flutter, flit’, **bilinc** ‘small bird, especially sparrow’, **blith** ‘air, breeze, zephyr’, **blaith** ‘spirit’, cf. **gum-laith** ‘weariness of spirit, depression, grief’

Trying to extract nodes from these meanings, I have come up with the following map, first showing the ranges of Q. *gwili-* (aerial motion) and *vili-* (movement of air):



Similarly, G. *duil-* is also positioned within aerial motion, while G. *gwil-* additionally spans into ‘movement of air’. Finally, *bil-* has a very wide range of meanings:

to Elvish is not obvious, except maybe for the frequent appearance of the root TALAT-, *atalat* ‘slip right down, fall in ruin’ (PE18:85) in examples (PE17:186, PE22:134-135 etc.).



Interestingly, ‘spirit’ does appear as the translation of G. **blaith**, but I take it from the example **gumlaith** ‘weariness of spirit, depression, grief’ that the meaning ‘mood’ is meant here. It seems to be a wordplay on the first lines of Percy Shelley’s poem ‘To a Skylark’: *Hail to thee, blithe Spirit! Bird thou never wert [...]*.

However, ‘breath’ does not appear in the above roots. The locus of meanings in Quenya and Goldogrin is rather ‘aerial motion’ and there is no branching into ‘life’, ‘soul’ and the like.

Curiously, however, we also find a number of words like G. **gwilm** ‘quiet, peace, rest – cessation’, G. **gwilb**, **gwilw** ‘quiet, peaceful’ (Q. **qildi-**, **qilda** etc.) related to rest and ceasing. They sound similar, but are not related to *gwil-* above. Remember that ‘rest’ is related to ‘breath’ in natural languages via the necessity to breathe after physical labour.

4.2.2 Late Quenya

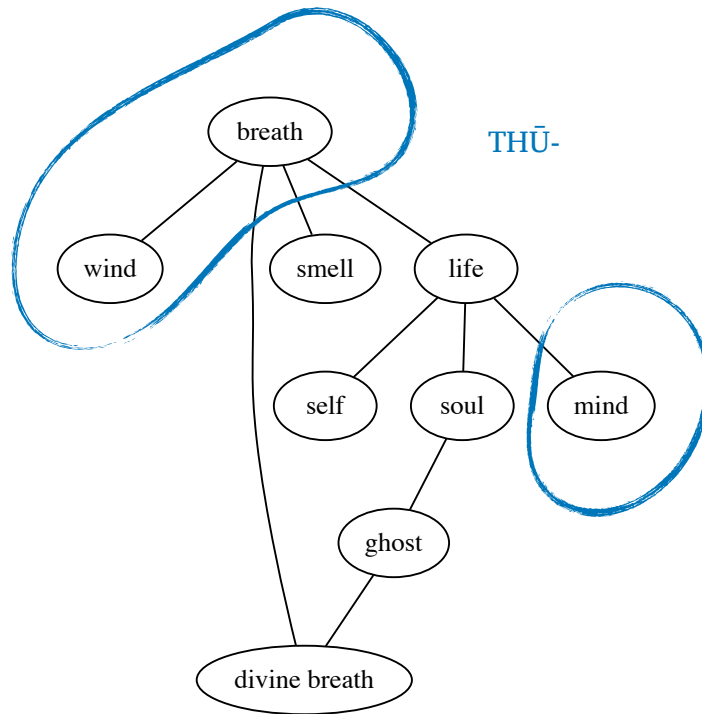
In later Quenya, however, we find the following glaring examples:

- **ksaräre psare sūle** ‘longing frets the spirit’ (PE22:119)
- **sūle** ‘breath’, with ‘spirit’ above it with pencil (PE22:50)

And in addition:

- **sūre** ‘wind, breeze’ (PE17:62,135)

It would appear that Tolkien has now integrated the semantic relation between ‘breath’ and ‘spirit’ into Quenya, but we do not know what kind of nuance of ‘spirit’ the word **sūre** or **sūle** corresponds to. We may therefore simplify the semantic map from natural languages by subsuming more closely related concepts into clusters. According to PE22, the root THŪ- would occupy the three nodes related to ‘breath’, ‘wind’ and ‘mental faculties’:



However, even later (PE17:124), there is a drastic change of concept, whereby:

Eldar did not confound ordinary “breath” of the lungs with “spirit.” The particular spirit indwelling in a body they falled fëa [fāyā]; spirit in general as a kind of being they called fairë. [...] But the Eldar held that “spirits,” the more as they had more native inherent power, could emit their influence to make contact with or act upon things exterior to themselves [...] To describe this they used the √THŪ- [or √SŪ-]. In addition Manwë, the most powerful spirit in Arda, in this respect was Lord of Air & Winds, and the winds were in primitive Eldarin thought to be especially his emission of power for himself.

What we now have is:

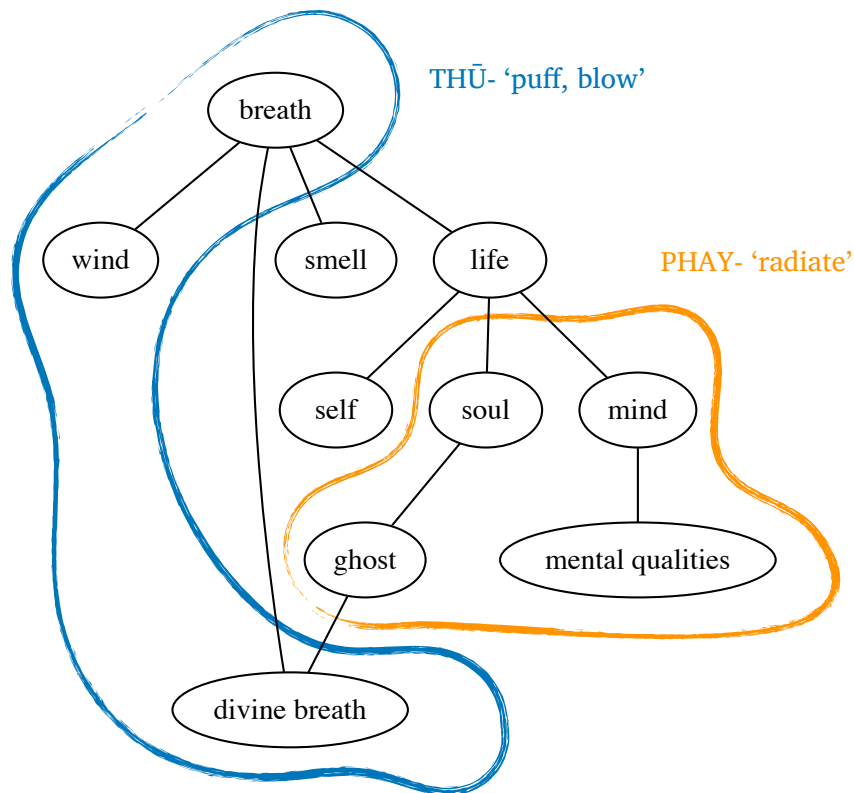
- CE *thūle* ‘blowing forth’, ‘spirit = the emission of power (of will or desire) from a spirit’
- Q. **súre** ‘wind’

However, there is now a new player in town, as the semantic space related to spirits, souls and mental faculties is claimed by a different root PHAY- ‘radiate’:

- Q. **fea** ‘spirit, soul’ (VT41:14)
- Q. **faire** ‘spirit (disembodied), phantom’ (PE17:124)
- Q. **faila** ‘fair-minded, just, generous’ (PM:352)
- Q. **Feanáro** ‘Spirit of Fire’ (MR:217, PM:343)

Note that **Feanáro** was previously glossed ‘Radiant Sun’ in the *Etymologies*. The change of the name’s meaning is therefore consistent with a reworking of the semantic understructure. In fact, associating the meaning ‘radiate’ rather than ‘breathe’ with ‘soul’ is part of larger metaphorical complex where the body is seen as a house, while the soul is seen as the light within it, so that we have the polysemy **koa** <house, body> (PE17:107,174) and **koakalina** ‘soul’, lit. ‘houseslight’ (MR:250,470). We may speculate that this concept was inspired by the New Testament and the association of Jesus’ body with the Temple of Jerusalem. To my knowledge, there is no natural language with the exact polysemy <house, body>, though there are related compounds and phrases like ‘bone house’ (Barouch, 2015 [21]).

One should however mention that Q. **Airefea** ‘Holy Spirit’ (VT43:37), a derivative of PHAY-, is used in the sense ‘divine breath’ and may wonder whether Tolkien would have updated it to ***Airesúle** or if the two roots can overlap at the concept DIVINE BREATH. Thus, as far as the attestations go, the revised concepts can be put on the map as follows:



5 Conclusion

Semantic maps are an illustrative way to demonstrate universal patterns in the semantics of natural language and provide an alternative way to gain insights into Tolkien's languages and his creative process. While grammar usually receives a lot of attention as far as analyses go, I would argue that creating the thousands of words in the lexicon and dividing up semantic spaces is where any language creator actually spends most of his time and thoughts. The semantic map technique is one of the ways to classify and understand this material. It also offers a comparison to natural languages, from which one can conclude that while Elvish mostly behaves like a human language, there are subtle and purposeful differences in the semantics that reflect a different world view.

A particularly interesting feature is the semantic complex around the polysemy <body, house> and the soul or spirit being the light of the house, which supplanted the concept of the spirit related to breath that is found in natural languages.

An analysis such as this one is naturally open-ended, as new semantic maps can be discovered and older ones expanded, especially as more material gets published. Some of the topics I left out due to the sparsity of attestations include indefinite pronouns, modality, prepositions of time and the instrumental/comitative.

To conclude, let me go back to the idea of Barfield that polysemy is essentially poetry. Thinking of the metaphor comparing a bird's wings to the branches of a tree (Q. **yēni únótime ve rámar aldaron** 'long years numberless as the wings of trees'), I typed a query into List et al.'s database which produced one example:

- Siona (Colombia & Ecuador): **kap'i** <branch, wing>

While the ejective sound in this word is a bit un-Elvish in terms of phonaesthetics, it serves to illustrate how one language's everyday polysemy may easily be another language's poetry.

Acknowledgements

I would like to thank everyone who participated in the discussion and critique of this work in Helsinki (2013) and Hayward (2017).

References

- [1] Takao Suzuki, *Words in Context: A Japanese Perspective on Language and Culture (Japanese Characters)*, Kodansha International, 1984 (first edition: 1978), translated by Akira Miura
- [2] Martin Haspelmath, *The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison*, The New Psychology of Language: Cognitive and Functional Approaches to Language Structure, Volume 2 (2003), edited by Michael Tomasello
- [3] Louis Hjelmslev, *Prolegomena to a theory of language*, The University of Wisconsin Press (1969), translated from Danish by Francis J. Whitfield (original 1943)
- [4] Lloyd B. Anderson, The “perfect” as a universal and as a language-specific category, *Tense-Aspect: Between semantics & pragmatics* (1982), edited by Paul J. Hopper
- [5] Pascal Boyeldieu, *From semantic change to polysemy. The cases of “meat/animal” and “drink”*, *From Polysemy to Semantic Change: Towards a typology of lexical semantic associations*, 106, 303 (2008)
- [6] Johann-Mattis List, Anselm Terhalle and Matthias Urban, *Using Network Approaches to Enhance the Analysis of Cross-Linguistic Polysemies*, *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013) – Short Papers* (2013)
- [7] William S. Annis, *A Conlanger’s Thesaurus*, *Fiat Lingua*, September 2014, <http://fiatlingua.org/2014/09/>
- [8] *The Catalogue of Semantic Shifts*, <http://datsemshift.ru/shifts>
- [9] Stephen Pinker, *The Stuff of Thought*, Penguin Books (2007), chapter 3
- [10] Devorah E. Klein and Gregory L. Murphy, *Paper has been my ruin: conceptual relations of polysemous senses*, *Journal of Memory and Language* 47 (2002) 548–570
- [11] *Hamtai Swadesh list*, https://archive.org/details/rosettaproject_hmt_swadesh-1
- [12] Michael Cysouw, *Interrogative words: an exercise in lexical typology*, Presentation presented at the Bantu grammar: description and theory workshop, Vol. 13. (2004)
- [13] Caterina Mauri, *Coordination relations in the languages of Europe and beyond*, Walter de Gruyter (2008)
- [14] Anita Slonimska and Seán G. Roberts *A case for systematic sound symbolism in pragmatics: Universals in wh-words* *Journal of Pragmatics* 116 (2017) 1–20
- [15] Russell Ultan, *Some general characteristics of interrogative systems*, *Universals of Human Language*, Vol. 4: Syntax, pp. 211-48, Stanford University Press (1978), edited by Joseph H. Greenberg
- [16] Oliver Bond, *Mapping negation in conceptual space*, *Proceedings of Conference on Language Documentation & Linguistic Theory 2* (2009), edited by Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells
- [17] Alexandre François, *Semantic maps and the typology of colexification: Intertwining polysemous networks across languages*, *From polysemy to semantic change: Towards a typology of lexical semantic associations* 106 (2008): 163.
- [18] Owen Barfield, *Poetic Diction: A Study in Meaning*, Wesleyan University Press (1987), first publication: Faber and Gwyer, Ltd. (1928)
- [19] Verilyn Flieger, *Barfield, Owen (1898-1997)*, *J.R.R. Tolkien Encyclopedia: Scholarship and Critical Assessment*, pp. 50-51, Taylor & Francis Group (2007), edited by Michael D. C. Drout
- [20] Humphrey Carpenter, *The Inklings: C.S. Lewis, J.R.R. Tolkien, Charles Williams and their friends*, George Allen & Unwin (1978)
- [21] Valeria Barouch, *Lifehouses, Houselights and Other Encounters*, *Arda Philology* 5, Arda Society, 2015