

Word formation in Esperanto

1. Introduction

Word formation is the part of Esperanto grammar which above all gives Esperanto its power and ease-of-use. However the topic of word-formation is one where the original specification of the language was vague, to say the least. If we look at the “**Fundamento**” of Esperanto from 1887/1905, we read:

*“11. Vortoj kunmetitaj estas formataj per simpla kunigo de la vortoj (la ĉefa vorto staras en la fino); ili estas kune skribitaj, sed disigitaj per streketoj. La gramatikaj finaĵoj estas rigardataj ankaŭ kiel memstaraj vortoj. Ekzemple **vapor/ŝip/o** estas formita de: **vapor, ŝip, kaj o** (finaĵo de la substantivo)”*.

This description appears at first glance almost completely unhelpful from a learner’s point of view, providing not a clue to the immense possibilities for word-formation in Esperanto, or guidance on anything other than the most basic example. Some of it even no longer applies, such as the idea of subdividing words into their component parts with little strokes, which people very soon stopped doing. A lot of textbooks of Esperanto instead simply present Esperanto word-formation as a matter of memorising a few dozen affixes and the learner is expected to get on with it, which is not much use either. I shall try to explain below how word-formation **really** works, and hopefully never again will you be “lost for words”, or go “I didn’t know I was allowed to do that!?” when you see a particular example of Esperanto’s word-formation facilities being skilfully taken advantage of by a writer.

The discussion below may, to begin with, involve some technical linguistic points that some of you may be unfamiliar with, but these parts can be skipped and come back to later. Word formation involves some quite complex issues that would take decades of study if learning other languages, but which native speakers of these languages seem to manage to tackle without even noticing (and non-native speakers trip up over). But with Esperanto it isn’t as difficult as it may at first seem, and understanding this aspect of the language is well worth the effort.

2. Compounding

The major point we can take from the initial outline in point 11 of the “*Fundamento*” is that the order of components in forming compound words in Esperanto is defined as (to use the modern linguistic term) *headword last*. A *headword*, or simply, *head*, is that element in a construction which characterises or stands for the construction as a whole. So, in the example here, to form the word “steamship”, we take the stem of the word for “steam”, *vapor*, the stem of the word for “ship”, *ŝip*, and its noun part of speech ending, *o*, in that order. The product, *vaporŝipo*, is a kind of *ŝipo*, not a kind of *vaporo*. Thus *birdokanto* (“birdsong”) is a kind of *kanto* and *kantobirdo* (“songbird”) is a kind of *birdo*. The other word(s) involved in the compound is/are called the *modifier(s)*.

This is exactly the same way as noun compounds are formed in English (if we ignore the complication of the part of speech ending for the moment). The word “steam” is compounded with the word “ship” to form “steamship”, which is a kind of “ship”. Many other languages do the same thing, e.g. German brings together *Dampf* (“steam”) and *Schiff* (“ship”) to form *Dampfschiff* (“steamship”) in exactly the same order as English and Esperanto. Hungarian forms its word for “steamship”, *gőzhajó*, in the same way and in the same order with the words *gőz* (“steam”) and *hajó* (“ship”), and the Mandarin Chinese word for “steamship” is formed by compounding words in a way which is also essentially *headword last*. The Mandarin for “steamship” is *qì chuán*, composed of the words for “gas” and “ship”. This is short for *zhēng-qì chuán*, or “evaporate-gas ship”, *zhēng-qì* meaning “steam”. Another way to express “steamship” in Mandarin is *lún chuán* or “wheel ship” (because of the paddles on some of them). Either way, they are all kinds of *chuán*. (Or *hajó*, or *Schiff*, or *ŝipo*, or “ship”).

However, this is not how things are universally done. The Austronesian languages put the words the other way round (*headword first*), e.g. the Indonesian for “steamship” is *kapal api* (literally “ship” + “fire/boiler”), which is a kind of *kapal* (“ship”), not a kind of *api* (“fire/boiler”), or Hawaiian where “steamship” is *mokumāhu*, *moku* being “ship” and *māhu* “steam”. The Celtic languages, for the most part, operate the same way as the Austronesian languages, except for some terms directly borrowed from English. Other languages, such as Neo-Melanesian, use a mixture of headword last and headword first in compounds: *paiauwut* (“firewood”) but *kotren* (“raincoat”). This isn’t a great problem for Neo-Melanesian speakers, because there isn’t such a thing as we would call “coat rain” (using our headword last order) for it to be confused with.

Still other languages, such as French and other languages derived from Latin, do not allow compounding like this at all. Here you have to provide some sort of way

of linking the two concepts, such as a preposition, as in *bateau à vapeur* (literally “ship at steam”). The virtue of the French way of doing things is that the relationship between the two components of the compound is more clearly defined, but on the down side it is longer and clumsier to express. The closest you can get to compounding in French is to use compounds comprised of a verb and a noun. Thus the French for “a toothpick” (Esperanto *dento-pik-ilo* or “toothpick-instrument”) has to be *un cure-dents*, or “a scrapes-teeth”.

The relationship between the components when they are simply compounded together as they are in Esperanto (or English etc.) is not defined. Because of this vagueness, the question can arise as to which element should be the headword, and sometimes in Esperanto there is more than one form of a compound using different orders of the component parts. The Esperanto word for “millennium” can be either *jarmilo* or *miljaro*. (Is it a kind of “group of 1000”, or is it a kind of “year”? Does it matter?). Similarly, “midnight” can be *meznokto*, which is a kind of “night”, or *noktomezo*, which is a kind of “middle”. Both are acceptable. And while it could be argued at a pinch that a *farbisto-tapetisto* (“painter and decorator”) is a kind of *tapetisto*, *Bosnio-Hercegovino* isn't really a kind of *Hercegovino*.

3. What can be compounded?

We have already seen how we can compound two nouns, as in *vaporo + ŝipo = vaporŝipo*. We could also say *vaporoŝipo*, leaving the part of speech marker (or *desinence*) at the end of *vaporo*. This means exactly the same as *vaporŝipo*, and the usual reason why we might choose to do this is to make the resultant compound easier to pronounce. But *vaporŝipo* is already easy enough to pronounce. There are no fixed criteria upon which to base the judgement as to what constitutes ‘easier to pronounce’, and users are free to decide for themselves. For example *kantobirdo* might be thought of as easier to pronounce than *kantbirdo*, but again, both are perfectly acceptable. We'll come back to the use of desinences later and also discuss the extent to which the stem *vapor-* means the same as *vaporo* or not.

Remaining with using a noun as the headword (second part) for the moment, we could use a verb as its modifier (the first part) rather than a noun. We could combine for example *skribi* “to write” and *maŝino* “machine” to make *skribmaŝino* “typewriter”. Here the stem *skrib-* has been used. Another example: *manĝi* “to eat”, and *horo* “hour, time” gives us *manĝohoro* “eating time, or mealtime”. We could also think of this as being a compound of *manĝo* “meal” and *horo*. The meaning of **manĝihoro* and *manĝohoro* is the same, and when there is a choice which does not affect the meaning, we use the compound which has a noun modifier, here *manĝohoro*, but usually *manĝohoro* is used because it is easy to

pronounce the way it is.

Most often compounds with a verb as the modifier do not have the verbal *-i* used in the compound, but they might if adding it would make the meaning clearer, as it does in *pagipovo* (“ability to pay”). If all we want to do is make the verb-noun compound easier to pronounce and it doesn’t otherwise actually matter to the meaning if we use a noun or a verb as the first part, then we use a noun. Thus *servi* “to serve” and *preto* “readiness” gives us *servopreto*. Pronouncing **servipreto* would have been slightly difficult, and *servopreto* means much the same as **servipreto* would have done.

In exactly the same way, we can make compounds of adjectives and nouns, e.g. *ora* “golden” + *horloĝo* “clock” = *orhorloĝo* “gold clock”, or nouns and adjectives: *akvo* “water” and *plena* “full” gives *akvoplena* “waterlogged”, or adjectives and adjectives: *hela* “light-coloured” and *blua* “blue” gives *helblua* “light-blue”, and so on. Note *dubasenca* “of dubious meaning” (adjective + adjective) as compared with *dubosenco* “meaning of doubt” (noun + noun), because the part of speech of the modifier would alter the meaning here.

We can do compounding with adverbs: *matenmanĝi* “to breakfast” is *matene manĝi*, “to eat in the morning”. *Fingromontri* “to point” is *fingre montri* “to show with a finger”, but as *iu fingro montras ĝin* “some finger shows it” is much the same as *iu fingre montras ĝin* “somebody shows it with a finger”, we have a choice of **fingremontri* or *fingromontri*, so we usually choose *fingromontri*. Saying **fingrmontri* wouldn’t have been possible (try it).

The word *finfari* “to complete” is from *ĝisfine fari* “to do to the end” rather than from *fine fari* “finally to do”. We see here that despite (or maybe because of) the essential vagueness of the relationship between the two components, you might sometimes create a concept that is slightly different in meaning from that which you might get by spelling out the two components separately.

Other components involved in compounding are prepositions, which are among the stand-alone roots of Esperanto which do not require a desinence (part of speech ending): *sen* “without”, *forta* “strong” gives *senforta* “weak”. There are quite a few other stand-alone roots including a whole lot ending in *-aŭ*, an ending we’ll talk about later. Some prefixes can be stand-alone, like *ek* and *fi*. Numbers (below *miliono*) are another class of word that is used in a stand-alone way, without needing a desinence.

The prefix *mal-* which means “opposite” is not a preposition and normally requires a desinence, e.g. *male* “to the contrary” or something else to be attached

to as in *malantaŭ* “behind”. This affix would give another word for “weak”: *malforta*, and the difference between this and *senforta* is a very subtle nuance indeed. But the difference between a word beginning with *sen-* and a similar word with *mal-* instead might be very different depending on the words involved, e.g. *senamika* “without friends” and *malamika* “pertaining to the enemy”.

We can make compounds with pronouns: *onidiro* “rumour” from *oni diras* “one says”. One can do it with interjections: *ve-kanto* “song of woe”, comprised of *ve!* “oh dear!” and *kanto*. Notice the hyphen in *ve-kanto*. I have put it in here because we have a problem deciding how the word *vekanto* is composed. It might be *vek-anto* “someone who wakes someone else up”. This type of problematic compound which has more than one way of being analysed is called an *isomer* and it is a source of humour (of sorts) in Esperanto. Another example is *informa* - is this *inform-a* (“informative”) or *in-forma* (“female-shaped”)?

If the meaning of a compound is an isomer that might be misanalysed, or if it simply makes the word easier to read and understand, a hyphen is inserted. This option exists in English, but often whether a hyphen is used or not is specific to the pair of components in the compound. Another option in English is to write the two words side by side, e.g. “steam ship”. This is never an option in Esperanto, except with proper names. For some reason, probably the influence of French during Esperanto’s early history, these usually take the form of headword first if written as two separate words, and the normal headword last order if hyphenated: *la Turo Eiffel*, but *la Eiffel-Turo*. There are infinite compounds that can be made with the particle *ĉi* and because of its shortness and hence lack of distinctiveness, these compounds tend always to be hyphenated; thus *ĉi-nokte* but *ĉi tiun nokton* (because we are not surprised to find *ĉi* together with the correlative *tiun*).

We’ve explained when we do and don’t include desinences in the compounds, and with nouns the plural or accusative markers are not usually included in word-formation. But there are exceptions. In the article by John Francis elsewhere in this number of *Esperanto en Skotlando* you will see the word *ovojnmetilo* (“ovipositor”). This makes it clear that the organ in question is for burying “eggs” (plural) in the tree bark. You can even make words out of whole sentences: *neforĝesumino* (*ne-forĝesu-min-o*) “forget-me-not” (the flower *Myosotis*), also known as *miozoto* in Esperanto. All of this is proof that in Esperanto, it is *words* that are compounded, not *roots*.

Here are another couple of examples of the role of the desinence: we can say *ĉokoladskatolo* or *ĉokoladoskatolo*, but if we want a compound ending in *-ujo*, we only ever say *ĉokoladujo*. We don’t say **ĉokoladujo*. This is probably because the

affixes are easy to add on to roots without having an adverse effect on the pronunciation, so the question of having to insert the desinence never arises. Sometimes we need to insert a desinence because of a clash with a word that already exists. **Post** “after” + **-ulo** “person” would give us **postulo** which already means “demand”. So we say **posteulo** “successor”, using **poste** “afterwards” instead. And of course there is a distinction between **unuco** “unity” and **unuaeco** “the quality or fact of having been first”.

There are many prefixes and suffixes which can be used in word formation in Esperanto, and there is no need to go into all of them again here. The interesting point about them is that they can be used independently – they don’t need to be used together with another word. Here are some of them: **eta**, **ega**, **emo**, **ejo**, **ero**, **aro**, **ilo**, **ono**, **ulo**, **ano**, **ismo**, **eco**, **inda**, **enda**, **estro**, **ujo**. Of course, we could talk about an “ism” and an “ology” in English, but in Esperanto this is a facility in word-formation that can be taken advantage of across the board. Prefixes don’t need to be combined with nouns or verbs or adjectives only, they can also be combined with a part of speech ending alone e.g. **praa** “ancient”. Affixes can also be combined with one another, e.g. **anaro** “team”, **fiulo** “scoundrel”.

4. Derivation

The **headword last** rule doesn’t necessarily apply with some of the **affixes** we use in word formation in Esperanto. While **manĝilo** “eat-tool”, i.e. “item of cutlery” is a kind of **ilo**, **manĝeto** “eat-diminutive”, i.e. “snack” isn’t a kind of **eto**, it’s a kind of **manĝo**. Because of this we have to think of some of these compounds as being **derived** from their **headword** rather than as being compounded from more than one component, at least as far as their meaning is concerned. Other affixes in addition to **-et-** which could be regarded as producing derivations when added at the end of a word are **-eg-** and **-aĉ-**. **Domego** (“mansion”) is a kind of “house”, and so is **domaĉo** (“hovel”). Esperanto grammarians describe these three affixes as **diafanaj** or **diaphanous** (“see-through”), because they don’t affect what the headword is. Almost all other suffixes do. They become the headword.

However, we could use **-et-** as a prefix instead of a suffix, with a slightly different meaning. If we compare Esperanto **industrieto** “industry on a small-scale” and **etindustrio** “light industry” we can see that the first word, using the suffix, **re-describes** the original word it is derived from, while the second one, with the prefix, **re-defines** it, and **industrio** remains the headword in both cases. This is not a technique that is universally applicable, but it is there to take advantage of should a suitable opportunity present itself. Sometimes the affixes in Esperanto are presented as if they can be used with any word we like, but they can not. They make as much or as little sense as compounding any chosen combination of

two words would.

Most languages conventionally make a distinction between **compounding** which is what we have just been talking about, which involves compounding free-standing meaningful elements, and **derivation**, which involves the use of endings which cannot be used in a free-standing way to create different forms derived from a word. **Derivation** in some languages can also involve using beginnings, or sticking things in in the middle, or otherwise changing a word. There are also a number of other methods of word formation apart from compounding and derivation, as we'll see later. In Esperanto there is not a rigid dividing line between compounding and derivation.

The major kind of **derivation** in Esperanto is changing the part of speech (e.g. noun, adjective, verb) by changing the part of speech ending. This is much easier, much more regular, and much more comprehensively available than in many other languages. Here are some examples: **varmo** “warmth/heat” from **varma** “warm/hot”; **demando** “question” from **demandi** “to ask”; **manĝo** “food/meal” from **manĝi** “to eat”; **kanto** “song” from **kanti** “to sing” and so on. Notice how often this is a process that in English is impossible to do without borrowing from Latin or using all sorts of different affixes with different words in order to perform the same function, e.g. **luno** “moon”, **luna** “lunar”; **vidi** “to see”, **vido** “view”.

Usually this transformation is not problematic and the meaning is obvious to everybody whatever their native language background. However, if **domo** is “house” and then we derive **doma** as “domestic”, what would **domi** mean? Speakers of English might think the translation would reflect the vagaries of English semantics and mean “to house”. But it doesn't. Similarly, speakers of Spanish might think it means “to marry” (compare Spanish **casa** “house” and Spanish **casar** or **casarse** “to marry”). Speakers of German might think of their word **hausieren** and think that **domi** means “to sell from door to door” (or house to house even), or “to hawk”. It isn't that either. In fact **domi** is hardly ever used as a word in Esperanto and if it were used, it would mean something much closer to the meaning of the root word, such as “to be a house”. So how would we say “to house”? Well, we could say **loĝ-igi** (“to cause to reside”). The German **hausieren** has an Esperanto equivalent in **kolporti**. (And of course, the verb “to hawk” has nothing to do with “hawk” as in the bird of genus **Accipitridae**, in Esperanto **akcipitro**.)

Another example: **fiŝo** “fish”; **fiŝa** “pertaining to fish”; **fiŝi** ...? That's right, it doesn't mean “to fish”, it means “to be a fish”, or “to act like a fish” etc. The English “to fish” is **fiŝkapti** (“to catch fish”). But to translate “fisher” you don't

have to say *fiŝkaptisto*, simply *fiŝisto* will do, because like all *istoj*, they are professionally associated with fish. Here's another one: *blanka* "white (adj.)"; *blanko* "white (noun), whiteness"; *blanki* "to be white". How do we say "to whiten"? Well, this has two meanings in English, and in Esperanto we have a choice of *blankiĝi* "to become white" and *blankigi* "to make white":

Li blankigis la muron per farbo "he whitened the wall with paint"
La muro blankiĝis pro la suno "the wall whitened because of the sun"

This brings us on to another issue in Esperanto, the use of transitive (with an object) and intransitive (without an object) verbs. We shall only touch on this briefly because it is quite a large subject in itself. Many verbs in Esperanto, if they are not derived from another part of speech, are in some cases, conventionally thought of as being intrinsically transitive, and in others, intrinsically intransitive. With some verbs, such as *komenci* or *fermi*, their category (in this case transitivity) is widely acknowledged and observed. If we want to use these verbs intransitively (without an object) we have to use the affix *-iĝ-* to derive an intransitive verb as in e.g. *la pordo fermiĝis* (as opposed to *iu fermis la pordon*).

Similarly with intrinsically intransitive verbs like *boli*, if we want to use these transitively, we use the affix *-ig-* to derive a transitive verb for use in sentences like *mi boligis la akvon* (as opposed to *la akvo bolas*). Some verbs can be used transitively or intransitively, e.g. *manĝi* (we can say *mi manĝis* (full stop) as well as *mi manĝis pomon*), and in some cases the category that verbs are allocated to is regarded as arbitrary by many speakers of the language. For example the Plena Ilustrita Vortaro gives *bani* as transitive, but sentences like *mi banis en la rivero* are widely heard without anyone batting an eyelid. Presumably the speakers meant *mi banis min en la rivero* and they missed out the *min*. Not a great problem, surely? Some verbs don't easily fall into any category. While *dormi* is normally thought of as intransitive, there is nothing wrong with saying *mi dormis bonan dormon*. The same applies to *vivi* when the object we use it with is *vivo*.

5. Digression: What's a word?

A "word" is, in written language, thought of as any form that is conventionally separated from another form by a space. It is the smallest form that can stand alone in a meaningful way. Similarly, we could think of the spoken "word" as something which could meaningfully stand on its own, separated from other words by silence on either side. In turn, each word can consist of individual elements, or "morphemes", each of which may or may not be able to stand alone, but which contribute an identifiable meaning or grammatical role to the word.

Sometimes there is not an easy way of deciding what is a word and what is a component part of a word. For example, in Urdu, there is a class of words called postpositions. These are just like prepositions, words (like “to”, “for”, “with”, “at”; *al, por, kun, ĉe*) in English and Esperanto, except they come after the word they govern rather than before it. In Urdu, postpositions are written separately from the nouns and pronouns they relate to. The Hindi language is so closely related to Urdu that it may be thought of as the same language, except for the different writing system, and some different vocabulary items (only slightly more differences than there are, say, between British English and American English). However, in Hindi, these postpositions are written as **part of** the words they govern, as if they were simply endings that get added. And yet, they are in no way different in meaning or grammatical role from when the same elements are used in Urdu.

There is a similar issue in Esperanto. Esperanto is often described as an **agglutinative** language, like Turkish, Japanese, Swahili and many other languages outside of Europe. This means that the endings have a fixed meaning, and are added essentially without change and without affecting what they are added to, and usually with one ending having one meaning or function. (Unlike **flectional** languages like Latin where the behaviour of such endings is altogether more complicated).

On the other hand, just as all free-standing words (like e.g. *vaporo* and *ŝipo*) can take part in **compounding**, some of the **affixes** in Esperanto don't have to be “bound” to some other word, they could stand alone. Take the affix *ek-*, for example. This is usually added to verbs to give an aspect of suddenness or beginning, e.g. *tondris* “it was thundering” versus *ektondris* “it thundered suddenly” or “there was a clap of thunder”. We can use *ek-* on its own, e.g. *Ekl* “Let's go!”. *Pra* is used as an affix to add the meaning of “ancient” to another word, e.g. *avo* “grandfather” compared with *praavo* “great-grandfather”. Reto Rossetti used *pra* on its own in a poem in the phrase *pra ses mil jaroj*, which is much more evocative than had he simply said *antaŭ ses mil jaroj*.

Desinences are very rarely used in word-formation except in their conventional role, and in grammatical terms such as **u-formo** (to refer to the imperative or volitive mood in Esperanto verbs). But even desinences have been used in a free-standing way in Esperanto, as in Grabowski's poetic use of the verbal ending **-as** in *al mi junul' kuraĝa, mond' as malfermita*. (There's a nuance hidden here, which we'll talk about later). Examples like *pra* and *ek* and *as* are few and far between, but there are many prepositions and other words that routinely appear on their own without being compounded with any other morphemes.

Because of this, and the quasi-independent role of most affixes, Claude Piron has argued that the nature of Esperanto is not *agglutinative* like Turkish etc., or *flectional* like Latin, but *isolating*, like Chinese, and to a lesser extent, English. However, affixes in Esperanto usually require the support of at least one desinence per word, and the stems of ordinary nouns and verbs do not usually stand alone either. Although we can say *anaro* (*an-ar-o*, “members”), we cannot usually say **an de la klubo*, we have to say *ano*; although we can say *tio estas farinda* (“worth doing”), we cannot say **mi far tion hieraŭ*, we have to say *faris*.

But there are exceptions to this too. If instead of writing **an de la klubo*, I wrote *an' de la klubo* (the spoken form is identical between the two, of course), that would be o.k. because we would recognise it as a poetic version of *ano*, with the final vowel elided because I maybe felt like it. Some parts of speech, e.g. prepositions, don't have part of speech endings in Esperanto and their root stem is the same as their root word. So, by the same token, the process of deriving words could include not adding an ending on to a root stem, just as we saw with *ek!* above: *traduko de Shakespeare farita de William Auld* “a translation of Shakespeare done by William Auld” can be said as: *traduko de Shakespeare far William Auld*. Here to avoid the possible ambiguity (by a listener potentially unaware that Shakespeare died about 300 years before William Auld was born) in the phrase *traduko de Shakespeare de William Auld* a preposition *far*, meaning “done by” has been derived from the verb *fari* “to do”.

The whole thrust of Esperanto word-formation is to treat the words or morphemes as unchangeable (although we will look at some circumstances in which they actually can change later on). In the original design proposal for the basics of Esperanto, Zamenhof's main concern was to avoid the kind of irregular change that we have already mentioned as being typified by languages such as Latin. He said:

“Mi enkondukis plenan dismembriĝon de la elementaj konceptoj, kiuj estas esprimataj per memstaraj vortoj. La tuta lingvo konsistas ne el gramatike ŝanĝataj formoj de diversaj vortoj sed el vortoj neŝanĝeblaj”.

We have already seen that, contrary to the explanations often given to learners, word-formation is a matter of combining words rather than roots. However, the concept of a distinction between the root and the words derived from it is still a useful one, and this is a matter which we shall discuss next.

6. Parts of speech and the “morpheme effect”

All Esperanto words are either root words themselves or are derived from these

root words. Although the root stem itself does not have a part of speech ending, this does not mean that the concept inherent in the root stem is any different from the part of speech implied in the root word that root is derived from. Then other affixes and part of speech endings are added on to turn it into another part of speech, but the meaning is based on the meaning the root word had, including the part of speech it was. An example may make more sense: Let us take the root word **haki** to chop. All other words related to **haki** are derived from it. Removing the part of speech ending **-i** gives the root stem **hak-**. Then from this we can derive:

haka	“pertaining to chopping”
hako	“the concept of chopping, the action of chopping”

If we wanted a word for the tool that does the chopping, we have to say **hakilo** using the affix **-il-** because the idea of the tool is not contained in the root concept.

Contrast this with another word where the root word is not a verb, but a noun. **Martelo** is “a hammer”. We can add endings on the root stem **martel-** as follows:

martela	“relating to hammers” (NOT “relating to hammering”)
marteli	“to hammer (i.e. acting like, or being, a hammer)”

If you then want a word for the noun “hammering”, you need to use the affix **-ad-** which is used to derive a word relating to a continuous action to get **martelado**.

Here’s another example. Suppose you have the word **homo** “person, human (noun)”, and you derive the word **homa** “person, human (adj.)” from this. How do you then derive an abstract noun from **homa**? After all, **homa** is an abstract quality, so for making a noun we could just change it to **-o**. But we can’t do that otherwise we’d just end up with **homo**. Answer: you use the affix **-eco** and form **homeco** “humanity”. This behaviour we see in words like **martelado** and **homeco** in Esperanto is called the **morpheme effect**. We are trying to add one desinence on to another, which happens not to be possible:

martel-i + o	= martel-ad-o
hom-a + o	= hom-ec-o

If we didn’t do this we would end up back with **martelo** and **homo**, which is not what we intended. There’s only one difficulty, and that’s remembering what the original root part of speech is, although quite often, as with transitivity and intransitivity, it either doesn’t matter, or people don’t know or care.

7. Limits to word-formation

About a century ago some people, with a rather clockwork sense of logic, tried to say that this inability of Esperanto to be “reversible”, if a root was taken through a series of changes of part of speech back to the original starting one, was a defect in Esperanto. In their proposed reform of Esperanto, called Ido, they made it difficult to change the part of speech of a word except by the use of affixes. So while in Esperanto we can easily form constructions like *ĝardena pordo* “garden gate”, *nokte vestita* “dressed in night attire”, *instrua libro* “an instructive book”, *li bonsanas* “he is in good health”, and so on, Ido speakers had to say things like *gardenala pordo*, *noktale vestizita*, *instruktiva libro* and *il bonsanesas*. (Ido grammar also prohibited the independent use of affixes, and although Ido uses the Esperanto affix *-ido*, the independent word *ido* does not exist in the language, and its users have to say *decendantanto* instead!) Such unnecessary difficulties probably ensured the failure of their project, despite the fact that about a third of the leadership of the Esperanto movement at the time were in favour of the reform.

The split with the Idists and their consequent departure (many later drifted back) also meant that a related issue in Esperanto word-formation was resolved. From early on in Esperanto’s history, a leading figure in the Esperanto movement, Louis de Beaufront (pseudonym of Louis Chevreux, 1855-1935) had practised and advocated a hair-splitting style where every nuance of a word was expressed by putting in every affix that could be added to form the word concerned, even when unnecessary. This conflicted with the usage of good writers in the language such as Kazimierz Bein (a.k.a. Kabe, 1872-1959), and Zamenhof himself. With de Beaufront gone over to Ido, leading Esperanto speaker René de Saussure (1868-1943) decided to resolve the question by the following formulation, generally adopted in 1912, known as the principle of necessity and sufficiency:

“Principo de neceso: En konstruon de vorto oni devas enkonduki ĉiujn vortradikojn, sufiksojn, prefiksojn kaj finiĝojn necesajn por elvoki klare kaj plene la ideon reprezentotan de tiu vorto. Principo de sufiĉo: Se la ideo tiamaniere esprimita estas jam bone kaj nekonfuzeble komprenata el la kunteksto sen iu el la sufiksoj uzataj, tiu ĉi sufikso povas esti elĵetata kiel neutila kaj nenecesa”.

Thus if you want to translate “Do you want a chocolate?”, while offering someone an open box of chocolates, you don’t have to say *Ĉu vi volas ĉokoladaĵon?*; it will suffice to say *Ĉu vi volas ĉokoladon?* and *ĉokoladaĵo* would only be needed if you really needed to specify that you meant “piece of confectionery made out of chocolate” as in *estas ĉokoladaĵo sur la supro de la kuko* as opposed to *estas ĉokolado sur la supro de la kuko*. Similarly in Esperanto “Maria is my friend” is *Maria estas mia amiko*. We don’t need to call Maria an *amikino*. It will be obvious from the

word *Maria* that she is female and even if it isn't obvious, it may not be relevant. Speakers of English tend to want to translate verb forms literally into Esperanto, e.g. "I have done that" as *mi estas farinta tion*. This is again unnecessary; *mi faris tion* will do. Even if it is absolutely vital to specify that the action is completed, you could say *mi jam faris tion*.

Sometimes there are alternatives for words that are constructed by word-formation techniques. Sometimes one is a compound and the other is a single root, perhaps a piece of international scientific vocabulary, for example *sunfloro* as opposed to *helianto*. Both mean "sunflower" but one is a layperson's term, a compound of *suno* and *floro*, and the other is a scientific term (based on the Greek for "sun" and "flower"). Often single terms are neater and more precise than compounded terms, but of course speakers of Esperanto may be less familiar with them, or at least they may not be as able to work out the rough meaning of the word.

As well as scientists, Esperanto poets have given us a large number of alternative one word terms for which compounds may already exist. These include, for example, *zefiro* instead of *venteto*, or *aperta* instead of *malferma*, and many other alternatives for compounds of *mal-*, and from within the resources of the language itself, rather than by borrowing, *eta* instead of *malgranda*. In the case of the alternatives of the *mal-* words some have found wide acceptance outside of poetry, and others have not, but often there is a different nuance to the meaning. *Aperta* is more open than *malferma* and *eta* is smaller than *malgranda*, perhaps because in one the meaning is concentrated into one component and in the other it is spread over two. Similarly if we compare the sentences *la ĉielo estas blua* and the more poetic *la ĉielo bluas*, the colour feels more intense in the latter. And in Grabowski's *mond' as malfermita* we have openness as some intrinsic quality of the world rather than as an accidental feature.

Are there limits to compounding in Esperanto? Well, in fact there are. Some languages, like Inuit, have words that are equivalent to what would be sentences in English or Esperanto. In the corpus of material studied by Chris Gledhill (2000), he discovered in one source the term *bonfartigilmalaĉetloko*. I suppose this could be translated as "place for refunding the things that make you fare better", but it without doubt was meant by the author of this unusual compound to be a ridiculous extreme. In practice the maximum length of a word in Esperanto, excluding the desinence and any plural or accusative markers is about 3-4 morphemes.

8. A (mostly) rejected method of word-formation

The process of word-formation is nowhere as organised or as easy as this in any other language. However, Esperanto borrows heavily from Latin, and there are traces of Latin's problems that infect Esperanto in some cases. For example we have the distinctions between *procedo*, *proceduro*, *proceso*, *procesio* and *procezo* (too lengthy to go into here), or the irregularity of the derivational relationship between *reagi* and *reakcio*, or between *redakti* and *redakcio*.

Sometimes in Esperanto there are pairs of words that are distinguished by a slight difference in consonant, such as: *pezi* "to weigh, i.e. to be heavy" versus *pesi* "to weigh, i.e. to determine the weight of", *agordi* "to tune" versus *akordi* "to be in tune", *ŝargi* "to load (a gun)" versus *ŝarĝi* "to load (a lorry etc.)" and *kurzo* "rate of exchange, share price" versus *kurso* "course". It is hard to believe this now, but these pairs are all either remnants or analogues of a system that was originally deliberately intended for much more general use in Esperanto. In Zamenhof's notes in 1881-1882 about a possible future proposal for an international auxiliary language, *Lingvo Universala*, the following pairs of verbs are to be found (modern Esperanto equivalents in brackets): *vidi/vitti* (*vidi/rigardi*), *odi/otti* (*aŭdi/aŭskulti*), *trovi/trofi* (*trovi/serĉi*), and *levi/lefi* (*levi/leviĝi*). Happily for us this arrangement never made it to the final version.

9. Unsung methods that are used productively

We said earlier that there are limits to word-formation. However, this relates principally to the number of components and not to the ingenuity with which items are meaningfully combined. In addition to the standard methods outlined above, there are a number of other relatively unacknowledged and unsung methods of word formation, and Esperanto takes advantage of all of them.

Abbreviation: One can combine the initial letters, or the initial parts of words, e.g. *SAT*, *FIFA*, *kolĥozo*. If the new combination is pronounceable without using the names of the letters, an **acronym** results. If the abbreviation is pronounced using letter names, e.g. *LKK*, *EAB*, we call this an **initialism**. And the names of the letters can also be spelled out in full, e.g. *aboco*, *Usono*.

Clipping: The new word is formed by abbreviating the original word. Two affixes in Esperanto require the word they are attached to to be clipped in this way, *-njo* and *-ĉjo*, giving us words like *panjo*, *Manjo*, *oĉjo* and *Vilĉjo*. But the technique has a wider application, e.g. nowadays the word *aŭtomobilo* is often clipped to *aŭto* and this in turn can form words like *aŭtisto*. Often the word *Eŭropo* is clipped as in *eŭroskeptikulo*. It might be thought that the technique is relatively recent in

Esperanto (except with *-njo/-ĉjo*), but it actually has an ancient pedigree and many long standing words in Esperanto actually began life as clippings, e.g. *nepre* “without fail” is originally a clipping of the Russian *nepremenno*.

Blending: The new word is formed by compounding parts of two words, e.g. *flua* “fluent” + *fuŝa* “incompetent” = *fluŝa*, “talking quickly but not very carefully”. Sometimes this technique is used for stylistic or amusing effect, as for example by inserting the root *fuŝ-* when referring to the headquarters of UEA as *La Centra Ofuŝejo* (instead of *La Centra Oficejo*). There are other examples for more serious purposes using parts of correlative words, e.g. *alies* “somebody else’s”, or *estiel* “as or like something, but actually being it as well”. A similar technique with personal pronouns is evident in *ŝli* (more often spelled *ŝ/li* by analogy with English “s/he”), but *ivi* “you (plural)”, formed by analogy with *ili*, has proved less popular. Someone recently wanted elegantly to distinguish the exact position in problematic sentences like *la viro staras malantaŭ la ĉevalo*. (Is the man “behind” from the observer’s point of view, who is perhaps looking at the side of the horse, or is he “behind” from the horse’s point of view?). And so, he proposed *pugaŭ*, a combination of *pugo* “arse” and the ending *-aŭ*, which appears in a number of words, some of which are prepositions like *antaŭ* or *malantaŭ*. Whether this will catch on or not remains to be seen. The status of *-aŭ* is quite strange: there are identifiable categories of words that use it, yet it is not normally considered to be free to behave as our friend proposing *pugaŭ* is doing. The ending *-aŭ* actually was an affix in Zamenhof’s *praesperanto* of 1881-1882 (but it showed the feminine then).

Back-formation: This method of word-formation involves leaving out a part of a word which resembles an affix, but actually isn’t one. For example, *uskleco* (“case, i.e. whether printed in upper or lower case”), derived from *majuskleco* and *minuskleco*, which (erroneously) appear to consist of an affix *maj-* or *min-* and a common element *uskleco*. There are also examples where an apparent affix is dropped (often *mal-*, which of course really is an affix in other words, but isn’t in the particular cases involved). Again this is often for stylistic or amusing effect rather than serious use. Examples include *ica* “well intentioned” (as opposed to *malica* “malicious”) and *graŭ* “because of” (as opposed to *malgraŭ* “in spite of”). And of course we have pairs like *turisto* / *turismo* and *feministo* / *feminismo*. *Tur-o* already exists in Esperanto, so it would be a bit difficult to separate it off to mean “tour” (this still has to be something like *rondi-ro*), and *turistoj* are not just interested in towers, but the latest edition of PIV now recognises *femino*, derived from *feministo* / *feminismo*, as a word in its own right as an alternative for *virino* and *ino*.

The first part of *turisto* / *turismo* and *feministo* / *feminismo* is, or used to be, what

linguists refer to as a **cranberry morpheme** i.e. having no meaning except as part of the compound, just like English “cran-” in “cranberry”. Historically, “cranberry” is from the mediaeval German *Kranbeere* where the *Kran-* part of the word did have an independent meaning (“crane bird”). The normal Esperanto word for “cranberry” is *oksikoko* (from the Greek for “sour berry”), but the word *grubero* (from *gruo* “crane bird”) is not unknown. Perhaps this is due to Swedish influence: their *tranbär* is also from *trana* “crane bird”. Modern Germans don’t say *Kranbeere*, they say *Preiselbeere*, which actually does have a **cranberry morpheme**, because **Preisel* isn’t a word in German (it’s from the Old Slavonic word *brusiti* “to strip”, because the berries are easy to strip off, and was borrowed originally into dialects of German spoken in close proximity to speakers of Slavonic languages first of all, but the word later became more widespread).

Reanalysis/misanalysis : In the final method the new word is formed by dividing the word up in a different way from the one represented by the etymology of the word, e.g. the word *hamburgero* is actually derived from *Hamburg*, and the further derivation *burgero* is based on a reanalysis or misanalysis of the original word as *ham+burgero*. Sometimes this technique is used for amusing effect as in “*Kiumotive vi venis al Londono? – Lokomotive!*”. Or with words that are genuine isomers: “*Dom-aĉeto? Domaĉ-eto!*”.

10. And finally

To illustrate the richness of some of the techniques that have been discussed here, I have given here some examples of spontaneous word-formation by Esperanto speakers which Claude Piron lists in his highly recommendable book *La Bona Lingvo*. See if you can work out what they mean, and think of your own expressive creations that have maybe never seen the light of day before.

Ajnista; malstultiĝi; profesori; oĉji; denaskismo; katolikige; malprotesti; disbraki; eldomigi; pria; pera; intera; pratimo; lingvisteco; muzeema; bluegi; distavoliĝo; vortelekti; vortkompari; eki; alprelegoto; buŝpleni; longvorte; elstarigi; transkulturiĝebla; deki; tezi; ekdisponebla; gami; perbuĝete; perprestige; samŝancige.

Ed Robertson

Bibliography:

- Atanas D. Atanasov, *La Vortfarado en Esperanto*, in *La Lingva Esenco de Esperanto*, Rotterdam: UEA, 1983.
- Christopher Gledhill, *The Grammar of Esperanto: A Corpus-Based Description*, München: Lincom Europa, 2000.
- Kálmán Kalocsay & Gaston Waringhien, *Plena Analiza Gramatiko*, Rotterdam: UEA, 1985.

- Christer O. Kiselman, *Transitivaj kaj Netransitivaj Verboj en Esperanto*, in *La Stato kaj Estonteco de la Internacia Lingvo Esperanto*, (ed.) Petr Chrdle, Praha: Kava-Pech, 1995.
- Aleksandro S. Melnikov, *Specifaj Kulturaj Scioj de la Esperantista Kvazaŭetno kaj ilia Respeguliĝo en la Koncerna Lingv(aj)o*, Rostov-na-Donu, 1992.
- Claude Piron, *La Bona Lingvo*, Wien: IEM, 1997 (2nd edition).
- Reto Rossetti, *La Esperanta Vortludo*, in *El la Maniko*, La Laguna: Stafeto, 1955.
- Balázs Wacha, *Antaŭ-Sosurumo: Fragmentoj el verkata studo pri René de Saussure*, in *Studoj pri Interlingvistiko: Festlibro por Detlev Blanke*, (eds.) Sabine Fiedler & Liu Haitao, Praha: Kava-Pech, 2001.
- Gaston Waringhien, *Lingvo kaj Vivo: Esperantologiaj Eseoj*, Rotterdam: UEA, 1989.