

## Word Building in Tokcir: Compounding and Derivation

by Stephen DeGrace, stevedegrace@yahoo.ca, 2002/04/06

### 1.0 Introduction

This is the fifth in a series of essays intended for conlangers interested in the NGL Project, both to bring potential newcomers to the project up to speed on the basic state of the project and the language and to serve as reference materials for those already active in the project.

This essay deals with word building in Tokcir. The Tokcir language, at the current phase of the language's development, has a constantly growing vocabulary. Please see the Central Repository at [http://www.geocities.com/ngl\\_repository](http://www.geocities.com/ngl_repository), or the files section of the NGL Yahoo Group web site at <http://groups.yahoo.com/group/NGL> for the General Vocabulary Rules, which outline how this process works, as it is beyond the scope of this essay. However, in accepting new morphemes in the language, a principle we have generally always followed has been to avoid redundancy. Tokcir makes *extensive* use of compounding and derivational affixes to produce new words; project workers can propose new official derivations as conventions for certain meanings in proposing "modules", the basic unit of new material proposal in the NGL Project, and speakers and writers can freely use these official derivations or create their own on the fly. Compounding and derivation are enormously productive in Tokcir, moreso even than in a language like German, although not nearly so much as in a language like Inuktitut. Most derivational affixes are highly, highly productive, and compounding can be engaged in very freely.

Although monomorphemic roots in Tokcir are coined with a default part of speech, powerful derivational affixes, usually suffixes, exist, which can transform a morpheme of practically any part of speech into practically any other. See the General Vocabulary for details. If you download the Shoebox Database from the Central Repository and filter the database for Derivational, you should get a fairly complete list of the Tokcir derivational affixes; you should undertake to familiarise yourself with all the common derivational affixes before attempting composition or word coining.

In general, before you coin a new morpheme, you should search the database firstly to make sure a word with the meaning you want or close to it doesn't already exist, obviously, and secondly to make sure the meaning you want could not be generated through compounding, the employment of derivational affixes to existing morphemes, or both. Although we are purposely a little vague about this, try and make sure any new coining is useful and not redundant, and justify the inclusion of a new morphemic root in some way. At the same time, please don't be *too* shy about making vocabulary proposals. The best way to learn is by doing, and your fellow group members will assist you in getting into the swing of things in this area. Remember, too, that there are always legitimate grounds for argument in this area and that we basically operate by deate and consensus as much as possible.

**Please Note:** Tokcir does **not** permit double letters, neither double consonants nor double

vowels. All tokcir vowels are the same length, or their length is determined by prosody. Vowel length is not phonemic. In coining new words morphemic words, you may not use double consonants or vowels. See section 3.0 for a discussion of what happens when derivation or compounding causes double vowels or consonants.

One thing to consider, too, in attempting to coin new words is that you don't always *have* to get the exact meaning you want. Different languages have radically different idioms some times, and the idiom of Tokcir is basically under construction through the process of making new compositions; compositions are not merely a frivolous or fun exercise in taking a break for some artistic expression, but are a serious part of the exercise of language building as they are instrumental in building the Tokcir idiom. There are numerous different ways things can be phrased or couched, and when you seem to run into a wall, consider rephrasing things and inventing a new idiom or looking for an old one for what you want rather than always trying to make new words.

Tokcir generally has no policy about basing new coinings on words from other languages. New coinings may be inspired by other languages or may be completely a priori. We have a precedent of discouraging *too* much "borrowing", but we also have a precedent of a fair bit of actual "borrowing" taking place. I'd say the basic rule is try to be a bit original; it's okay to take inspiration from other languages by all means but make up some words a priori too. In principle it doesn't matter whether any particular word in Tokcir resembles a word from another language with a similar meaning or not. This aspect simply has little or no meaning or relevance to our project except as something of an aesthetic concern. You can also bring words into the language that are to be officially recognised as borrowings from other languages (an example being {politik})... at this point in the language's development, you tend to have to justify these strongly, though, to get them accepted. They tend to be fairly international words and used for a certain style or specialised purpose. An exception is in the area of scientific terminology, where borrowing of internationalisms is not discouraged or even encouraged. See for example the IUPAC Nomenclature Chemistry Module. Some of this vocabulary is Tokcirised, like the chemical names, and some of it is just swallowed whole, like biological organism names.

That's some general advice; the main thing is to try and do good work but above all to have fun with it.

## 2.0 The Parts of Speech

Tokcir recognises the normal parts of speech, those being nouns, verbs, adjectives, prepositions, adverbs and other familiar categories. Tokcir uses prepositions as opposed to postpositions in constructing modifying phrases.

## 3.0 Derivation and Compounding

See section 1.0 for a discussion of derivational affixes. Tokcir has a rich variety of derivational

affixes, and the conlanger and language user alike needs to be well familiar with the common affixes to work in Tokcir. Tokcir has suffixes and prefixes, and even some interfixes (affixes that go between morphemic elements in a word) and a small number of infixes (affixes that are actually inserted into a morphemic element), although infixes tend not to be very productive and their usability has to be specifically coined with the coining of the morpheme they are to be applied to. The most common category is suffixes, followed by prefixes. Prefixes are always syllabic. Suffixes and interfixes can be syllabic or phonemic, and infixes are all phonemic. In coining derivational affixes, some thought always has to be given to the way they will be modified when and if their application causes violations of the Tokcir laws of syllable pattern. The typical default devices are the insertion of 'e's to break up illegal consonant clusters and the appending or insertion of glottal stops to legalise or break up illegal vowel clusters. The first preference is usually to legalise the cluster (as happens when a word is caused to begin or end with two vowels, for example) by placing a glottal stop (') before or after it as the case may be to close the syllable, the second preference is to break up illegal clusters by inserting a glottal stop.

Some prefixes are unstressed, that is, they never take stress themselves but instead shift stress to the immediately following syllable. This causes a variation from the normal stress pattern of the language (where the first syllable of a word is always stressed no matter what). Unstressed, or stress-shifting prefixes are marked as {pfx-`}... so, for example, the inversion prefix (a *very* important feature of the Tokcir language which is dealt with in more detail in section 3.2) is written {in-`}. Example of usage:

kelan N            -            mercy#cruelty

{kelan} = "mercy", {inkèlan} = "cruelty"

In actual fact, what the {-`} means is more precisely that the {in} is unstressed, rather than to say that it moves stress to the second syllable. If stress on the word is already irregularly marked, then the stress remains on that syllable rather than moving to the second syllable. The main thing is that {in} be unstressed. A detailed discussion of the rules of stress is beyond the scope of this essay but may be taken up in a future essay. In the meantime, a Google Groups search will turn up group discussions on this topic which would shed light on it, should you run into any complex situations. An example of applying a {-`} prefix to an irregularly stressed word would be if we imagine a fictional inverse to the word {edè}, meaning "to spend time". Its inverse would be {inedè}, not {inède}.

Tokcir, as has been stated, does not permit double vowels or double consonants. If, in the process of applying derivational affixes or in the process of compounding you bring two of the same consonant or vowel together, what happens is that they combine into a single consonant or vowel.

Examples:

tiwe V            -            to drain

-em	-	NVS third person singular verb agreement ending
tiwem	-	he, she or it drains
duin N	-	a human being
-ne	-	particle forming question word from a noun
duine	-	"who?"

There may be times when it might be desirable to evade this assimilation rule, however, it is generally applied very strictly. In coining affixes it is possible to specify unique, irregular behaviours that apply instead of assimilation of vowels or consonants, and speakers may insert glottal stops on the fly to disambiguate certain vowel assimilations, or else use de facto consonant gemination (doubling) to disambiguate certain consonant assimilations. But in general it is assumed that native speakers prefer to assimilate and work out intended meaning from context.

Tokcir permits complex and elaborate compounding of any number of words, although in practise most compounds are two or three morphemic words long. Tendency towards compounding or isolation in tokcir is something which is currently in flux, with some writers liking rather long compounds and some avoiding them. It is sufficient to know that you *can* make rather large and elaborate compounds, but you don't have to. Tokcir has a somewhat Germanic penchant for using compounds, and clever compounding, like in German, would tend to be an area where speakers would enjoy exercising their creativity.

Tokcir is a head first language, unlike English and German which are head last languages. This means that the most principal element in a Tokcir compound is always the first element. Where English makes compounds like "snowman" and "paperboy", Tokcir would make compounds like "mansnow" and "boypaper". Generally, if you make compound words backwards of what your English instincts tell you, you will tend to make compounds which those with Tokcir instincts would find acceptable. Both head first and head last languages exist in nature. I am given to understand that head last languages tend to have prepositions (this is the case in English) and head first languages tend to have postpositions. However, Tokcir as a language what is head first but uses prepositions is not, I am given to understand, unprecedented in nature.

Components of compounds tend to be monomorphemes, although elements other than the last can have suffixes that apply only to that element, and components other than the first can have prefixes that apply only to that element. Prefixes applied to the first element of the compound tend to apply to the compound as a whole, and likewise suffixes applied to the last element of a compound tend to apply to the compound as a whole. I believe that what part of speech the unadorned compound functions as is in practise somewhat flexible and dependent on context and convention, however, the part of speech of the head element strongly influences the part of speech of the compound as a whole and tends to dictate it.

Some cases may be ambiguous as to what to use for the head; use your best judgement. There are a number of derivations recorded in the Database which may not have the correct head, as we have had to learn over time how to use the compounding rules correctly; if you see

something in the database that you think may be in error don't hesitate to point it out, and equally, while you can often use forms listed in the database as a guide to help you in determining what precedents exist, always take it with a little grain of salt, too, at least until we have had more time to do a spring cleaning.

A couple possible compounds that have attestation in the Tokcir corpus:

tiug N	-	authority
demos N	-	the cizizen body
tiugdemos N	-	repulic or democracy
kas N	-	breath
mog N	-	shadow
kasmog N	-	spirit

### 3.1 Clitic Suffixes

A special derivational category midway between compounding and the use of derivational affixes is the use of clitic suffixes. Tokcir clitics can always function as independent words, but also have an ability, built into the word when it is coined or else conferred upon the word in some specialised meaning through a module, to suffix to the end of other words, creating new words with predictable meanings and parts of speech. Tokcir clitics are always suffixes. Either nouns or verbs may be clitic, and they always derive the same part of speech as the clitic itself. In cases where the ability to function as a clitic is defined into the word when it is coined, the word is fully capable of standing alone but tends to prefer to "lean" on some other word, which it modifies, as a clitic when at all possible. The host of a Tokcir clitic is always modified by its clitic guest, and Tokcir clitics cannot lean on words they do not modify.

The most noticeable effect of clitic suffixes is that they produce forms which tend to be the reverse of normal Tokcir compounds, as the clitic itself in a normal compound would often serve as the head and therefore be the first element of the compound. Take the clitic endings {-set} and {-tok}. {Set} means "land" and {tok} means "to speak". However, as part of the Countries module, whose conventions were generally accepted, these two words have specialised meanings as clitic suffixes. {-set} derives the name of a country from the name of an ethnic group with which it is associated, and {-tok} derives the name of a language (and by decree derives nouns rather than verbs - irregular exceptions such as this are possible but need to be explicitly proposed and accepted). So, for example:

Suèd	-	a Swedish person
Suèdset	-	Sweden
Suèdtok	-	Swedish language

If these words were produced from normal compounding, they would be {Setsuèd} and {Tokciarsuèd}.

Another quasi-clitic is discussed in a bit of detail in the fourth essay of this series (dealing with Tokcir possessives). That is the ephemeral possessive ending {-dasi}, which is composed of the verb {das}, "to hold", plus the adjective-deriving suffix {-i}.

The most common and important clitics are the verbs {je} and {fe}, which in the corpus are attested extensively functioning both independently and as the clitics {-je} and {-fe}. In nearly all cases, when it is possible for {je} and {fe} to function as clitics they do so - they prefer to attach clitically rather than function independently, but it is never ungrammatical to use them independently, just not normal style.

je V	-	to be, to have
fe V	-	to make, to do

{-je} and {-fe} can make almost any other word (but typically nouns and adjectives) into verbs meaning to be/have the substrate or to make/do the substrate. The most typical host for {-je} is an adjective, independent uses of {je} are more likely when the possible host is a noun.

Examples:

hael Adj.	-	blue
haelje V	-	to be blue (e.g. {q kuaj haelje}, "the car is blue")
xuxu Adj.	-	cute
xuxuje V	-	to be cute (e.g. {visom ke xuxujem}, "I think it is cute")
siliv N (proposed)	-	dirt, grime
silivje V	-	to have dirt, to be dirty (e.g. {q zuen mouk silivje}, "the child is very dirty")
dus N	-	agreement
dusfe V	-	to make an agreement, to reach and agreement
eklät N	-	a dull clatter
eklätfe V	-	to make a clatter
pis N	-	urine
pisfe V	-	to urinate

### 3.2 The Tokcir Inverse with {in-`}

A major feature of the Tokcir language is the device of word "inversion" to expand the vocabulary. Inversion is a favoured route in Tokcir to produce new words, and one thing you should always be looking for when you are considering a new coining is whether the meaning you want might instead be produced by defining an inverse for an existing word.

Inverses can be constructed on the fly, but by and large official inverses should be defined. The inverse of a word is its "linguistic opposite". This is not the same as a negative or a "logical" opposite, which means the sort of logic used to figure out inverse meanings is rather human and flexible, and maybe a little quirky sometimes. It is for this reason that inverses typically need to be defined along with the regular meaning of a word; there might be more than one possible

inverse interpretation, but every word should have only one official inverse meaning. If, in coining a word, you see a obvious inverse, or actually want to use it in its inverse form, be sure to define an inverse as well as regular meaning; see section 4.0 for more information about the style and mechanics of making such proposals. If you want to assign an inverse meaning to a word which does not have a listed inverse you can as I say, invent one on the fly, but you *should* also propose a modification to the morpheme to make that meaning the official inverse, as this will help in the expansion and codification of the language and therefore is a productive exercise.

Inverses are made by applying the prefix {in-` } to a word. Any part of speech but prepositions and some special morphemes may take an inverse, at least in theory, as long as it is capable of acting as a stand-alone word. For many words it is pointless, confusing or cumbersome to consider their inverses, and for others, while there is a degree of binarism and inverses *could* be employed (for example, we could say that male is the inverse of female or female is the inverse of male) we instead use separate words to avoid assigning primacy and derivativeness to any options. Inverse meanings *are* clearly derivative and less primary than normal meanings, and so when picking the normal and inverse definitions of a coining, consideration should be given to which meaning you want to be primary and also which meaning you see as being used more and make sure that that meaning is the normal meaning of the word.

As stated in section 3.0, the notation {in-` } means that to make the inverse you prefix the word with {in-} and place a grave accent over the syllable immediately following the {in-}. In actual fact this means that the {in} is unstressed; see section 3.0 for a more in depth discussion of what that entails.

A search through some of the compositions of the Tokcir corpus would uncover any number of inverses in use, and a search through the official database will turn up many official defined inverses, which may be a help in getting a further idea of existing conventions in this area.

#### 4.0 Vocabulary Conventions in NGL Project Proposals and Compositions

With every module proposed and every piece composed, the Tokcir language grows in its vocabulary and idiom. We can set the basics down in documents like these but the real life of the language is in its *use*, and as a young conlang, its idiom and conventions are very much in flux - this is one of the most exciting areas of the language's development and one of the areas in which it needs the most work. You can pretty much make it up as you go along, which is what makes it fun. At the moment, while there are some idioms worth mentioning, the Tokcir idiom is too unstable to be set down in any kind of definitive fashion in a piece like this. On account of this fact, when you propose a new module or write a composition, you will likely have to invent new vocabulary - either new derivations being used idiomatically in specific ways, or else new General Vocabulary coinings.

It is common practise in all proposals and compositions that deal with vocabulary creation or which employ idioms which may not be completely transparent or which may be interesting for

their own sake, to include a vocabulary list with your post to the group. This list typically contains all potentially problematic or interesting derivations plus any new coinings which you propose to add to the General Vocabulary. It is generally good form to collect important terminology into lists, and to at some place in the document collect up all General Vocabulary proposals, either to add morphemes, change morphemes in sound or meaning, or delete entries in the General Vocabulary, and put them in one place. This makes researching easier, and makes it easier to deal with the General Vocabulary proposals you are making.

In listing vocabulary, you should generally note the part of speech of any word; this is typically done with abbreviations like N, V, Adj, etc following the word. Part of speech is very important in Tokcir as it tells you how to apply derivations.

When defining the inverse meaning of a new coining, or defining an inverse as a modification to an existing morpheme that lacks a defined inverse, the hash symbol (#) is typically used to separate the normal from the inverse meanings, inverse following normal. So, for example, if I am listing the word {kelan}, I would do it:

kelan N - mercy#crueently

This means that {kelan} means "mercy" and {inkèlan} means "cruelty".

Finally, new coinings which are General Vocabulary proposals are typically marked with a preceding \*. So, let's say I want to coin the new word {pirem} to mean "monkey". This is how I might list the proposal:

\*pirem N - monkey

None of these conventons are written in stone or rigidly enforced. But it *is* helpful to know them so that others can comfortably deal with your proposals and compositions.