Tg̃TEX: TEX Spells for Typesetting in Tengwar User's Manual Version 1.10*

Ivan A Derzhanski 22 September 2000

1 Introduction

This document describes the use of $T\acute{g}T_EX$ (tengtex.tex and tengtex.sty) for the $T_EX / I^AT_EX / I^AT_EX 2_{\varepsilon}$ typesetting of text in the languages of Middle-earth in Fëanorian tengwar, using my Computer Modern Tengwar family, its variant Computer Roman Tengwar, the fount teng10 by Julian C Bradfield or the revised fount tengwar by Michael P Urban (or any combination of these).

Perhaps it is worth reminding that the name $T_E\!X$ is derived from the Greek word $tekhn\bar{e}$ 'art', which (along with the Latin $tex\bar{o}$ 'weave', textum 'web, fabric, texture', textus 'text') goes back to the Proto-Indo-European root * $tek\bar{s}$ - 'create, frame, carve', no doubt from a Lemberin cognate of the Proto-Eldarin base TEK- 'write' (The Lost Road: 391), whence also Quenya péq 'writes', Sindarin pébî 'write!'.

The input of TgTEX is essentially an Asciisation of the Rómenórean spelling adopted by J R R Tolkien and described in Section 1 of Appendix E to The Lord of the Rings. (One difference is that capital letters are used for increasing the total number of letters available for the input, instead of singling out sentence-initial words and proper names, a device which is not a regular feature of the Fëanorian script²; another is that no accent marks are used, because of the limitations of Ascii.) Its output strives to follow the orthographic conventions of the Third Age, discussed in Section 11 of the same Appendix and illustrated

^{*}Inspiration has been drawn from the package ArabTEX by Klaus Lagally. I give thanks to Emanuele Vicentini, Gernot Katzer, Dirk Thierbach and Michael Urban for their comments and cooperation.

¹This document uses Computer Modern Tengwar by default. To substitute Computer Roman Tengwar, J Bradfield's or M Urban's tengwar, uncomment line 18, 19 or 20 of the source, then recompile it.

²Large ornate initials do occur in some documents, but are used very inconsistently.

in the few surviving samples of text in tengwar³; on those occasions on which the sources disagree, the evidence from the corpus takes precedence over the description in the Appendix. In its current version TgT_FX is intended for processing text in Quenya, Sindarin and their close relatives, and it also provides some support for Westron and the Black Speech, but not for any other languages, although it may work by accident for some post-Third_Age Mannish tongues (such as Irish in Sindarin mode).

TgTEX is not copyrighted, but its use is protected by an oath, the text of which (in English and Quenya) is to be found in the file vanda.tex in this distribution.

$\mathbf{2}$ Generalia

To activate TgTFX, load it by \input tengtex into a TFX source file, include the option tengtex into the document header of a IATEX document or instruct LATEX 2ε to \usepackage{tengtex}.

The fount used is selected by one of the spells \cmteng (Computer Modern Tengwar, the default option), \crteng (Computer Roman Tengwar), \jcb (J Bradfield's teng10) and \mpu (M Urban's tengwar). Tengwar can be magnified by the spell \tengwarmag (with one argument; the default is 0). The chosen magnification is activated when you next select a fount.

Three major modes are available: Quenya, Classical Sindarin and Beleriandic Sindarin. The appropriate mode is selected by the spells \quenya, \sindarin and \beleriand (there is no default mode). There are also certain options within each of the modes, described in the dedicated sections.

The typesetting of text in tengwar, particularly of longer passages, is done by the environment elvish. For the inclusion of short quotations in tengwar into text written in other scripts the macro \quetta is also available. An Elvish environment, but not an Elvish quotation, may consist of two or more paragraphs separated by blank lines.

In an Elvish environment or quotation continuous strings of letters (in the T_EX sense) are recognised as words and written in tengwar in the output, following the orthographic conventions of the mode (and ignoring letters to which a value is not assigned); continuous strings of decimal digits are recognised as numbers, and are inverted in the output.⁵ Numbers may form part of words.

Sauron ataltane 3019sse. božm ipzpm 🔁 ເວແເອ໌: [Quenya] 'Sauron fell in 3019.'

³One of these sources, Aragorn's letter to Master Samwise, is in fact an early Fourth Age document, but it can be taken as representative of Third Age Sindarin writing, since it was during the Third Age that the author received his Elvish education.—Some archaic (pre-Third_Age) conventions, in particular regarding Quenya, are also supported.

⁴ All the macro does is send its parameter to the environment for processing; it thus uses more of T_EX's memory, but it is shorter to type.

⁵Numbers are not explicitly marked as such with a dot or line above them, since the numerals in the founts are already sufficiently distinct from tengwar, in any case more so than a dot would be from an overdot tehta.

Any other characters (such as punctuation marks), blank space and T_EX, I^AT_EX or Tg̃T_EX parameterless commands are let through unchanged, and it is the user's responsibility to ensure that their visible effect, if any, makes sense in the chosen fount (for example, because of the way the character code tables are organised, a hyphen yields a tengwa number 11 in J Bradfield's fount and a wavy dash in M Urban's and in Computer Modern Tengwar). Empty pairs of curly brackets can be used to break ligatures, as they make Tg̃T_EX think that the word ends there and a new one begins.

3 Computer Modern Tengwar

The Computer Modern Tengwar family (included in this distribution) is to Computer Modern Roman as the tengwar themselves are to the mediæval Anglo-Saxon and Celtic calligraphic Roman hands that were their Primary World prototype. Just as Computer Modern Roman is TeX's default fount family, the Computer Modern Tengwar are now TgTeX's native accent.

I made this family with several purposes. First, I expect it to help us see the tengwar with the eyes of a European of old (such as Ottor Wæfre the mariner, named Eriol in Tol Eressëa), to whom they would not have appeared archaic, only different in the use, but not the style, of eminently familiar graphic elements.

Second, chances are that those who wish to use the tengwar for large-scale writing will need a wider variety of sizes and styles. Within this family counterparts of all Computer Modern and Concrete Roman typefaces can be generated and used. (Support for their selection has not been included in the package at this stage, but it can be added easily enough if the idea proves popular. In the meantime users are encouraged to experiment.)

Third, it blends better with surrounding text in Roman (so well, in fact, that I have found it necessary to underline all inclusions in tengwar in this document), which may be an advantage in some situations.

And fourth, there now is a fount that I can make $T \hat{g} T_E X$'s default, without having to give priority to either J Bradfield's or M Urban's.

4 Computer Roman Tengwar

The Computer Roman Tengwar family endeavours an even more radical move in the same direction as Computer Modern Tengwar: on top of the modernisation of the entire character set it performs a further degree of romanisation of the primary tengwar (the secondary ones are not altered). Needless to say, it blends even better with Roman text.

Although it claims to be a separate family, in reality it uses the same typefaces as Computer Modern Tengwar, only different parts of them.

5 Quenya

5.1 Vowels

In Quenya mode you have the option of choosing between heavy (default) and light vocalisation, using the words \heavyvocalise and \lightvocalise for switching back and forth.

Short vowels are written as tehtar over the preceding consonant or (when syllable-initial) over a short carrier. When light vocalisation (a glimpse of which is to be seen in J R R Tolkien, Life and Legend (Catalogue of the Bodleian Library Exhibition, Oxford, 1992), Plate 218) is selected, the tehta for a is not written, but all tengwar standing for syllable-final consonants receive a subscript dot tehta (falmar 'waves', for example, is set as $\underline{b_z \underline{w}}$ if vocalised lightly and $\underline{b_z}$ $\underline{\hat{w}}$ if vocalised heavily).

Long vowels can be either doubled in the input, and then the tehta is also doubled (yeen <u>and</u> 'year', mool <u>not</u> 'slave, thrall', uur <u>not</u> 'fire'; note that this only works with e, o and u), or capitalised, and then the single tehta is written over a long carrier (yAr gĵp 'blood').

Diphthongs with i as second element are written as yanta bearing the tehta for the first element (oire $\hat{\Lambda}\hat{g}$ 'everlasting age'). Diphthongs with u as second element are written as úre bearing the tehta for the first element (leuke \underline{z} ó¢ 'snake').

5.2 Consonants

The following consonants and consonant clusters are available:

c, k	q	f	b	G	cı	h	λ	hl	lح
hr	α l	ht	dp	hw	d	Н	d	1	7
lb	722	ld	ζ	m	122	mb	p2	mp	b_2
n	133	nc, nk	ccl	nd	bo	ng	cq	ngw	æ
nq, nqu	α d	nt	\mathbf{p}			N	ccı	Nw	व्य
p	p	q, qu	q			r	\mathbf{a}	rd	8
s	6	ss	દ			S	b	t	\mathbf{p}
v	12	W	a	x	Я	У	ä	z	ધ

Double consonants are written with an underbar tehta (quetta $\underline{\acute{q}p}^6$ 'word', tyelle $\underline{\acute{p}}\acute{\underline{\acute{e}}}$ 'grade'). The only exception is ss, which is written as esse (Gassa $\alpha\acute{e}$ 'hole, opening, mouth').

G. H. S.

N, Nw yield tengwar which had lost their original phonetic values by the Third Age. They are meant for writers concerned with etymology (Galda ας 'tree', HOn djm 'heart', Noola αζ 'wise, learned', Sinde bm 'grey').

⁶The examples in this section are vocalised lightly.

⁷The first notation is suggested by the fact that Proto-Eldarin g, preserved in Sindarin, became zero in Quenya, presumably going through a stage where it was an approximant, such

- q may be used without the following u, so both quetta and qetta yield άp.
- r is written as rómen when followed by a vowel or by y (r0men ½jíźp 'east', hrIve lgjź 'winter', kirya ġɣ 'ship') and as óre elsewhere (kirkar ġpqp 'sickles').
- R is always written as óre (ORe ĵń 'heart, inner mind').
- s is written as an underhook tehta when following another consonant (otso ip 'seven', aksa iq 'ravine', norsa mp 'giant').
- y is written as an under-diæresis tehta when following another consonant (Nwalya az 'torment').

When silme and esse bear a tehta for a following vowel, their nuquerna shapes are automatically substituted (silme ½½½ 'starlight', esse íá 'name').

5.3 An illustration

Galadriel's lament in Lórien

(From The Road Goes Ever On.)

ဂ် ၃၀ င်္ဂ ဂြိသာ ဂို ဂြို့သား (ထိုက် ကြော် ကြော် ကြော် ကြော် ကို ဂြို့သည်။ ထိုင်းကို ထိုင်းကို ကြော် ကြောက် ကြော် ကြော် ကြောက် က

îm 6j plâz \hat{n} ê \hat{n} î \hat{n} ê bâpa \hat{n} î pi \hat{n} î pi iz pi i i pi i pi \hat{n} î \hat{n} ê \hat{n} ê \hat{n} ê pî î pâpa ji \hat{n} ê pî î pâpa ji î pâpa pi î pâpa î pâpa

 $\hat{m}_{2}\hat{j}\hat{j}\hat{j}$ $\hat{m}\hat{n}$ \hat{n} \hat

6 Sindarin

6.1 Vowels

Short vowels are written as tehtar over the following consonant or (when none follows) over a short carrier. By default o is written as a curl open to the right and u as a curl open to the left, but their rôles can be swapped by specifying \oleft (Black Speech uruk ½q 'great soldier-orc, Uruk-hai') and restored by specifying \oright.

Long vowels can be either doubled in the input, and then the tehta is also doubled (eel $\underline{\underline{r}}$ 'star', too $\underline{\underline{p}}$ 'thither', uur $\underline{\underline{n}}$ 'fire'; note that this only works with e, o and u), or capitalised, and then the single tehta is written over a long carrier (Ar $\hat{\underline{j}}\underline{\underline{n}}$ 'king, chieftain', gwI $\underline{\underline{\pi}}\underline{\underline{j}}$ 'net, web', rhYn $\underline{\underline{\nu}}\underline{\underline{m}}$ 'hound of chase', Black Speech shaRkU $\underline{\underline{d}}\underline{\hat{n}}\underline{\underline{q}}$ 'old man'). No distinction is made between long and overlong vowels.

as might have been denoted by anna. The remaining four tengwar are pronounced as the ones input by the corresponding lowercase Roman letters.

Diphthongs with e as second element are written as yanta bearing the tehta for the first element (oer $\underline{\lambda}\underline{p}$ 'sea'). Diphthongs with i as second element are written as anna bearing the tehta for the first element (uir $\underline{\dot{q}}\underline{p}$ 'eternity'). Diphthongs with u as second element are written as ure bearing the tehta for the first element (naug \underline{m} \hat{q} 'dwarf').

6.2 Consonants

In either of the two Sindarin modes you have the option of choosing whether the calmatéma or the quessetéma is to be used for velar consonants (and, in the latter case, whether the calmatéma is to be used for palatal consonants) and whether the óretyelle or the númetyelle is to be used for single nasal consonants. The selection of the Classical Sindarin mode automatically allocates the velars in the quessetéma, though \vaswestron makes c available for calma (English cek qq´ 'cheque') and \vincalma transfers the rest of the velars into the calmatéma. (The shortcut \blackspeech is equivalent to \sindarin \oleft \vaswestron.) Also, the the nasal consonants are allocated in the númetyelle, though this setting can be altered by specifying \ninOre, and restored by \ninnUmen. The following consonants and consonant clusters are available in the default setting:

$\overline{}$			_			_					
b	p 2	bh	b2	В	þ22			c, k	q	ch, kh	d
d	po	dh	bo	D	bo	f	b2	g	~q	gh	α d
h	λ	hw	d			j	ccl	1	7	lh	ζ
m	122	mh	123	mm	<u>122</u>	n	\mathbf{m}	nn	$\widetilde{\mathbf{a}}$	N	व्य
p	р	ph	b					r	\mathbf{a}	rh	ક્ર
s	6	ss	ધ	sh	d			t	$_{\rm p}$	th	b
v	b_2			w	a			z	દ	zh	ccl

Double consonants are written with an underbar tehta (tellen pź m/ 'footprint'). The only exceptions are ss, which is written as esse, and mm and nn, which are written as malta and numen with or without an overbar or overtilde tehta.

- bh is provided for the sake of completeness, as an alternative to v.
- f is written as formen in initial and medial position (farn <u>bŷm</u> 'enough', tofn pɛm 'deep') and as ampa in final position (tif pɛ 'flute').
- gh, j,
- sh, z, zh are provided in order to make possible the typesetting of text in Westron (zIr &jp 'wise') or the Black Speech (ghAsh zdĵd 'fire').
 - H is always written as hyarmen, without forming a digraph with a preceding consonant letter.
 - i is written as yanta when it appears as a glide in initial position (iAr Δĵp 'blood', iolf Δζ̂b 'brand', iuith Δἀb 'use', iUl ΔĴζ or iuul Δζ̈ 'embers').
 - m is written as an overbar or overtilde tehta when preceding a labial consonant (hamp $\lambda\hat{\hat{p}}$ 'garment', ammarth $\hat{\hat{z}}\hat{z}\hat{b}$ 'doom, fate').

- n is written as an overbar or overtilde tehta when preceding a dental or velar consonant (rhink $y\dot{\tilde{q}}$ 'sudden move', lhann $z\hat{\tilde{m}}$ 'wide').
- ph is written as formen in initial and final position (alph $\underline{\hat{z}b}$ 'swan') and as formen with an underbar tehta in medial position (ephel $\underline{\underline{b}}\underline{\hat{z}}$ 'outer fence').
- r is written as óre when it occurs word-finally (bar pzp̂ 'home') and as rómen elsewhere (rAd zĵpp 'path, track', arth <u>ŷb</u> 'region, realm'). It is, however, always written as rómen if n is óre.
- R is written as ore if n is númen, and as romen otherwise.
- w is written as a squiggle tehta when following another consonant other than h (gwelw 遠茂 'air'). A tengwa can bear no more than one consonant superscript tehta, however, so w following a consonant preceded by a homorganic nasal is written separately as a wilya (angwedh 遠本放 'chain', glentweth 國政政政 'thinness'), unless the etymology suggests a different spelling (as in the case of thurin{}gwethil b汝成政政 'secret shadow woman').

When silme and esse bear a tehta for a preceding vowel (which is, in fact, always the case with esse), their nuquerna shapes are automatically substituted (kost qɔ̂p 'quarrel', bess pɔɜ́ 'woman').

6.3 Illustrations

6.3.1 The King's letter to Samwise Gamgee

(From Sauron Defeated.)

î bý λ âz $\hat{\mathbf{n}}$ $\hat{\mathbf{m}}$ $\hat{\mathbf{m}}$ ýż $\hat{\mathbf{m}}$ ćaz $\hat{\mathbf{m}}$ ćaz $\hat{\mathbf{m}}$ $\hat{\mathbf{m}$ $\hat{\mathbf{m}}$ $\hat{\mathbf{m$

6.3.2 The Ring inscription

(From The Lord of the Rings.)

d mộz phốp phá d mộz ap phá d d mộz by ap phá d phá by ap phá d d mộz by ap phá d phá by ap phá d d mộz by ap phá d phá by ap phá d d mộz by ap phá by ap phá by ap phá d d mộz by ap phá by ap phá

7 Beleriandic

7.1 Vowels

Short vowels are written as tengwar. Long vowels, written double (if they could be doubled in the Classical Sindarin mode) or represented as capitals in the

input, are written as the corresponding tengwar with an andaith (eel $\underline{\Delta z}$ 'star', too $\underline{p}\underline{\alpha}$ 'thither', uur $\underline{\delta y}$ 'fire'; Ar $\underline{\delta y}$ 'king, chieftain', gwI $\underline{\underline{\alpha}}\underline{\underline{n}}$ 'net, web', rhYn $\underline{\underline{s}}\underline{\underline{n}}$ 'hound of chase').

Diphthongs with e as second element are written in full (oer <u>uny</u> 'sea'). Diphthongs with i as second element are written as the tengwa for the first element bearing a diæresis tehta (uir <u>öy</u> 'eternity'). Diphthongs with u as second element are written as the tengwa for the first element bearing a squiggle tehta (naug pēq 'dwarf').

7.2 Consonants

The selection of the Beleriandic mode automatically allocates the velars in the calmatéma and the single nasal consonants in the óretyelle, though this setting can be altered by specifying \vinquesse or \vincalma for the former and/or \ninnUmen or \ninnOre for the latter. The following consonants and consonant clusters are available in the default setting:

b	p2	bh	<u>b</u> 2	В	<u>þ</u> 22			c, k	q	ch, kh	d
d	po	dh	bo	D	bo	f	b2	g	cq	gh	ccl
h	λ	hw	d			j	ccl	1	7	lh	ζ
m	12	mh	123	mm	122	n	\mathbf{a}	nn	\mathbf{m}	N	ccı
p	p	ph	b					r	ક	rh	જ
s	6	ss	ધ	sh	d			t	р	th	\mathbf{b}
v	<u>b</u> 2			W	a			z	ધ	zh	ccl

The glide i in initial position is written as a long carrier (iAr <u>jéy</u> 'blood'). The remaining input conventions are the same as in the Classical Sindarin mode.⁸

Since there are no vowel tehtar in this mode, silme and esse always appear in their upright (not nuquerna) shapes.

7.3 Illustrations

7.3.1 The hymn of the Rivendell elves

(From The Road Goes Ever On.)

7.3.2 The inscription on the Moria Gate

(From The Lord of the Rings. Observe the spelling mel{}lon <u>μλζζαρ</u> instead of mellon <u>μλζαρ</u>.)

vide posti cáce paste. By da syváta pipce i píz yide posti cáce paste. By da vyváta pipce i píz yide vyváta pipce i píz yide

⁸An error in the definition of the Beleriandic mode used to interfere with the proper interpretation of word-internal ph. In Version 1.10 this has been corrected.