

Hyphenation Exception Log

Barbara Beeton

This is the periodic update of the list of words that T_EX fails to hyphenate properly. The list last appeared in *TUGboat* 10, no. 3, starting on page 336. Everything listed there is repeated here. Owing to the length of the list, it has been subdivided into two parts: English words, and names and non-English words that occur in English texts.

This list is specific to the hyphenation patterns that appear in the original `hyphen.tex`, that is, the patterns for U.S. English. If such information for other patterns becomes available, consideration will be given to listing that too. (See below, “Hyphenation for languages other than English”.)

In the list below, the first column gives results from T_EX’s `\showhyphens{...}`; entries in the second column are suitable for inclusion in a `\hyphenation{...}` list.

In most instances, inflected forms are not shown for nouns and verbs; note that all forms must be specified in a `\hyphenation{...}` list if they occur in your document.

Thanks to all who have submitted entries to the list. Since some suggestions demonstrated a lack of familiarity with the rules of the hyphenation algorithm, here is a short reminder of the relevant idiosyncrasies. Hyphens will not be inserted before the number of letters specified by `\lefthyphenmin`, nor after the number of letters specified by `\righthyphenmin`. For U.S. English, `\lefthyphenmin=2` and `\righthyphenmin=3`; thus no word shorter than five letters will be hyphenated. (For the details, see *The T_EXbook*, page 454. For a digression on other views of hyphenation rules, see below under “English hyphenation”.) This particular rule is violated in some of the words listed; however, if a word is hyphenated correctly by T_EX except for “missing” hyphens at the beginning or end, it has not been included here.

Some other permissible hyphens have been omitted for reasons of style or clarity. While this is at least partly a matter of personal taste, an author should think of the reader when deciding whether or not to permit just one more break-point in some obscure or confusing word. There really are times when a bit of rewriting is preferable.

One other warning: Some words can be more than one part of speech, depending on context, and have different hyphenations; for example, ‘analyses’ can be either a verb or a plural noun. If such a word

appears in this list, hyphens are shown only for the portions of the word that would be hyphenated the same regardless of usage. These words are marked with a ‘*’; additional hyphenation points, if needed in your document, should be inserted with discretionary hyphens.

The reference used to check these hyphenations is *Webster’s Third New International Dictionary*, Unabridged.

English hyphenation

It has been pointed out to me that the hyphenation rules of British English are based on the etymology of the words being hyphenated as opposed to the “syllabic” principles used in the U.S. Furthermore, in the U.K., it is considered bad style to hyphenate a word after only two letters. In order to make T_EX defer hyphenation until after three initial letters, set `\lefthyphenmin=3`.

Of course, British hyphenation patterns should be used as well. A set of patterns for UK English has been created by Dominik Wujastyk and Graham Toal, using Frank Liang’s PATGEN and based on a file of 114925 British-hyphenated words generously made available to Dominik by Oxford University Press. (This list of words and the hyphenation break points in the words are copyright to the OUP and may not be redistributed.) The file of hyphenation patterns may be freely distributed; it is stored in the UK T_EX Archive at Aston and may be retrieved by anonymous FTP.

Hyphenation for languages other than English

Patterns now exist for many languages other than English. The Stuttgart archive

`ftp.uni-stuttgart.de`

holds an extensive collection in the directory

`soft/tex/languages/hyphenation`

The List — English words

<code>academy(ies)</code>	<code>acad-e-my(ies)</code>
<code>addable</code>	<code>add-a-ble</code>
<code>ad-di-ble</code>	<code>add-i-ble</code>
<code>adrenaline</code>	<code>adren-a-line</code>
<code>af-terthought</code>	<code>af-ter-thought</code>
<code>agronomist</code>	<code>agron-o-mist</code>
<code>am-phetamine</code>	<code>am-phet-a-mine</code>
<code>anal-yse</code>	<code>an-a-lyse</code>
<code>anal-y-ses</code>	<code>analy-ses*</code>
<code>anomaly(ies)</code>	<code>anom-aly(ies)</code>
<code>an-tideriva-tive</code>	<code>an-ti-deriv-a-tive</code>
<code>anti-nomy(ies)</code>	<code>an-tin-o-my(ies)</code>

antin-u-clear	an-ti-nu-clear	cognac	co-gnac
antin-u-cleon	an-ti-nu-cle-on	cog-nacs	co-gnacs
an-tirev-o-lu-tion-ary	an-ti-rev-o-lu-tion-ary	com-parand	com-par-and
apotheeses	apoth-e-o-ses	com-para-nds	com-par-ands
apotheo-sis	apoth-e-o-sis	comptroller	comp-trol-ler
ap-pendix	ap-pen-dix	con-formable	con-form-able
archipelago	arch-i-pel-ago	con-formist	con-form-ist
archety-pal	ar-che-typ-al	con-for-mity	con-form-ity
archetyp-i-cal	ar-che-typ-i-cal	congress	con-gress
arc-t-an-gent	arc-tan-gent	cose-cant	co-se-cant
	(better: arc tangent)	cotan-gent	co-tan-gent
assignable	as-sign-a-ble	crankshaft	crank-shaft
as-sig-nor	as-sign-or	crocodile	croc-o-dile
as-sis-tantship	as-sist-ant-ship	crosshatch(ed)	cross-hatch(ed)
asyp-tomatic	as-yp-to-matic	dachshund	dachs-hund
asyp-totic	as-yp-tot-ic	database	data-base
asyn-chronous	asyn-chro-nous	dat-a-p-ath	data-path
atheroscle-ro-sis	ath-er-o-scle-ro-sis	declarable	de-clar-able
at-mo-sphere	at-mos-phere	defini-tive	de-fin-i-tive
at-tributed	at-trib-uted	delectable	de-lec-ta-ble
at-tributable	at-trib-ut-able	democratism	de-moc-ra-tism
avoidupois	av-oir-du-pois	de-mos	demos
awo-ken	awok-en	deriva-tive	de-riv-a-tive
ban-dleader	band-leader	diffract	dif-fract
bankrupt(cy)	bank-rupt(-cy)	di-rer	direr
ba-ronies	bar-onies	di-re-ness	dire-ness
base-li-neskip	\base-line-skip	dis-parand	dis-par-and
bathymetry	ba-thym-e-try	dis-traugh-tly	dis-traught-ly
bathyscaphe	bathy-scaphe	dis-tribute(d)	dis-trib-ute(d)
bea-nies	bean-ies	dou-blespace(ing)	dou-ble-space(-ing)
be-haviour	be-hav-iour	dol-lish	doll-ish
be-vies	bevies	drif-tage	drift-age
bib-li-ographis-che	bib-li-o-gra-phi-sche	driver(s)	dri-ver(s)
bid-if-fer-en-tial	bi-dif-fer-en-tial	dromedary(ies)	drom-e-dary(ies)
bil-l-able	bill-able	duopolist	du-op-o-list
biomath-e-mat-ics	bio-math-e-mat-ics	duopoly	du-op-oly
biomedicine	bio-med-i-cine	eco-nomics	eco-nom-ics
biorhythms	bio-rhythms	economist	econ-o-mist
bitmap	bit-map	elec-trome-chan-i-cal	electro-mechan-i-cal
blan-der	bland-er	elec-tromechanoa-cous-tic	electro-mechano-acoustic
blan-d-est	bland-est	eli-tist	elit-ist
blin-der	blind-er	en-trepreneur(ial)	en-tre-pre-neur(-ial)
blon-des	blondes	epinephrine	ep-i-neph-rine
blueprint	blue-print	equiv-ari-ant	equi-vari-ant
bornolog-i-cal	bor-no-log-i-cal	ethy-lene	eth-yl-ene
bo-tulism	bot-u-lism	ev-ersible	ever-si-ble
brus-quer	brusquer	ev-ert(s,ed,ing)	evert(s,-ed,-ing)
bus-ier	busier	exquisite	ex-quis-ite
bus-i-est	busiest	ex-traor-di-nary	ex-tra-or-di-nary
buss-ing	bussing	fermions	fermi-ons
but-ted	butted	flag-el-lum(la)	fla-gel-lum(-la)
buz-zword	buzz-word	flammables	flam-ma-bles
ca-caphony(ies)	ca-caph-o-ny(ies)	fledgling	fledg-ling
cam-er-a-men	cam-era-men	flowchart	flow-chart
cartwheel	cart-wheel	formidable(y)	for-mi-da-ble(y)
catar-rhs	ca-tarrhs	forsythia	for-syth-ia
catas-trophic	cat-a-stroph-ic	forthright	forth-right
catas-troph-i-cally	cat-a-stroph-i-cally	freeloader	free-loader
cauliflower	cau-li-flow-er	friendlier	friend-lier
cha-parral	chap-ar-ral	frivolity	fri-ivol-ity
chartreuse	char-treuse	frivolous	fri-vo-lous
cholesteric	cho-les-teric	ga-some-ter	gas-om-e-ter
cigarette	cig-a-rette	geodesic	ge-o-des-ic
cin-que-foil	cinque-foil	geode-tic	ge-o-det-ic

ge-o-met-ric	geo-met-ric	metropoli-tan	met-ro-pol-i-tan
geotropism	ge-ot-ro-pism	mi-croe-co-nomics	micro-eco-nomics
gnomon	gno-mon	mi-crofiche	mi-cro-fiche
grievance	griev-ance	mil-lage	mill-age
grievous(ly)	griev-ous(-ly)	milliliter	mil-li-liter
hairstyle	hair-style	mimeographed	mimeo-graphed
hairstylist	hair-styl-ist	mimeographs	mimeo-graphs
harbinger	har-bin-ger	mimi-cries	mim-ic-ries
harlequin	har-le-quin	mi-nis	min-is
hatcheries	hatch-eries	min-uter(est)	mi-nut-er(-est)
hemoglobin	he-mo-glo-bin	mis-chievously	mis-chie-vous-ly
hemophilia	he-mo-phil-ia	mis-ers	mi-sers
hep-atic	he-pat-ic	mis-ogamy	mi-sog-a-my
hermaphrodite(ic)	her-maph-ro-dite(-ic)	mod-elling	mod-el-ling
heroes	he-roes	molecule	mol-e-cule
hexadec-i-mal	hexa-dec-i-mal	monar-chs	mon-archs
holon-omy	ho-lo-no-my	mon-eylen-der	money-len-der
ho-mo-th-etic	ho-mo-th-etic	monochrome	mono-chrome
horseradish	horse-rad-ish	mo-noen-er-getic	mono-en-er-getic
hy-potha-la-mus	hy-po-thal-a-mus	monoid	mon-oid
ide-als	ide-als	monopole	mono-pole
ideographs	ideo-graphs	monopoly	mo-nop-oly
id-iosyn-crazy	idio-syn-crazy	monos-pline	mono-spline
ig-niter	ig-nit-er	monos-trofic	mono-strofic
ig-n-i-tor	ig-ni-tor	mono-tonies	mo-not-o-nies
ig-nores-paces	\ignore-spaces	monotonous	mo-not-o-nous
impedances	im-ped-ances	mo-ro-nism	mo-ron-ism
in-finitely	in-fin-ite-ly	mosquito	mos-qui-to
in-finites-i-mal	in-fin-i-tes-i-mal	mu-d-room	mud-room
in-fras-truc-ture	in-fra-struc-ture	mul-ti-faceted	mul-ti-fac-eted
in-ter-dis-ci-plinary	in-ter-dis-ci-pli-nary	mul-ti-pli-ca-ble	mul-ti-plic-able
inu-tile	in-utile	mul-tiuser	multi-user (better with explicit hyphen)
inu-til-ity	in-util-i-ty	ne-ofields	neo-fields
ir-re-vo-ca-ble	ir-rev-o-ca-ble	newslet-ter	news-let-ter
itinerary(ies)	itin-er-ary(ies)	none-mer-gency	non-emer-gency
jeremi-ads	je-re-mi-ads	nonequiv-ari-ance	non-equi-vari-ance
keystroke	key-stroke	noneu-clidean	non-euclid-ean
kil-ning	kiln-ing	non-i-so-mor-phic	non-iso-mor-phic
la-ciest	lac-i-est	nonpseu-do-com-pact	non-pseudo-com-pact
lamentable	lam-en-ta-ble	non-s-mooth	non-smooth
land-sca-per	land-scap-er	nonuni-form(ly)	non-uni-form(ly)
larceny(ist)	lar-ce-ny(-ist)	nore-pinephrine	nor-ep-i-neph-rine
lifes-pan	life-span	nutcracker	nut-crack-er
lightweight	light-weight	oer-st-eds	oer-steds
limousines	lim-ou-sines	oligopolist	oli-gop-o-list
linebacker	line-backer	oligopoly(ies)	oli-gop-oly(ies)
lines-pac-ing	\line-spacing	operand(s)	op-er-and(s)
lithographed	lith-o-graphed	orangutan	orang-utan
lithographs	lith-o-graphs	or-thodon-tist	or-tho-don-tist
lobotomy(ize)	lo-bot-omy(-ize)	or-thok-er-a-tol-ogy	or-tho-ker-a-tol-ogy
lo-ges	loges	or-thoni-tro-toluene	ortho-nitro-toluene (or-tho-ni-tro-tol-u-ene)
macroe-co-nomics	macro-eco-nomics	overview	over-view
malapropism	mal-a-prop-ism	ox-i-dic	ox-id-ic
manuscript	man-u-script	painlessly	pain-less-ly
marginal	mar-gin-al	pal-mate	palmate
mat-tes	mattes	parabola	par-a-bola
med-i-caid	med-ic-aid	parabolic	par-a-bol-ic
mediocre	medi-ocre	paraboloid	pa-rab-o-loid
medi-o-crities	medi-oc-ri-ties	paradigm	par-a-digm
me-galith	mega-lith	parachute	para-chute
metabolic	meta-bol-ic	paradimethyl-ben-zene	para-di-methyl-benzene (para-di-meth-yl-ben-zene)
metabolism	me-tab-o-lism		
met-a-lan-guage	meta-lan-guage		
metropo-lis(es)	me-trop-o-lis(es)		

paraflu-o-ro-toluene	para-fluoro-toluene (para-flu-o-ro-tol-u-ene)	quasiequiv-a-lence	qua-si-equiv-a-lence or quasi-
para-g-ra-pher	para-graph-er	quasi-hy-ponor-mal	qua-si-hy-po-nor-mal
par-ale-gal	para-le-gal	quasir-ad-i-cal	qua-si-rad-i-cal
par-al-lelism	par-al-lel-ism	quasiresid-ual	qua-si-resid-ual
para-m-ag-netism	para-mag-net-ism	qua-sis-mooth	qua-si-smooth
paramedic	para-medic	qua-sis-ta-tion-ary	qua-si-sta-tion-ary
param-ethy-lanisol	para-methyl-anisol (para-meth-yl-an-is-ole)	qu-a-sito-pos	qua-si-topos
parametrize	pa-ram-e-trize	qu-a-si-tri-an-gu-lar	qua-si-tri-an-gu-lar
paramil-i-tary	para-mil-i-tary	quintessence	quin-tes-sence
paramount	para-mount	quintessen-tial	quin-tes-sen-tial
pathogenic	path-o-gen-ic	rab-bi-try	rab-bit-ry
pee-vish(ness)	peev-ish(-ness)	ra-dio-g-ra-phy	ra-di-og-ra-phy
pen-ta-gon	pen-ta-gon	raf-f-ish(-ly)	raff-ish(-ly)
petroleum	pe-tro-le-um	ramshackle	ram-shackle
phe-nomenon	phe-nom-e-non	ravenous	rav-en-ous
philatelist	phi-lat-e-list	re-ar-range-ment	re-arrange-ment
phos-pho-ric	phos-phor-ic	re-ciproc-i-ties	rec-i-proc-i-ties
pi-cador	pic-a-dor	reci-procity	rec-i-proc-i-ty
pi-ran-has	pi-ra-nhas	rect-an-gle	rec-tan-gle
pla-ca-ble	placa-ble	ree-cho	re-echo
plea-sance	pleas-ance	restorable	re-stor-able
poltergeist	pol-ter-geist	ret-ri-bu-tion(s)	ret-ri-bu-tion(s)
polyene	poly-ene	retrofit(ted)	retro-fit(-ted)
polyethy-lene	poly-eth-yl-ene	rhinoceros	rhi-noc-er-os
polygamist(s)	po-lyg-a-mist(s)	righ-teous(ness)	right-eous(-ness)
poly-go-niza-tion	polyg-on-i-za-tion	ringleader	ring-leader
polyphonous	po-lyph-o-nous	robot	ro-bot
polystyrene	poly-styrene	robotics	ro-bot-ics
pomegranate	pome-gran-ate	roundtable	round-table
poroe-las-tic	poro-elas-tic	salesclerk	sales-clerk
postam-ble	post-am-ble	salescle-rks	sales-clerks
postscript	post-script	saleswoman(en)	sales-woman(en)
pos-tu-ral	pos-tur-al	salmonella	sal-mo-nel-la
pream-ble	pre-am-ble	sarsaparilla	sar-sa-par-il-la
preloaded	pre-loaded	sauerkraut	sauer-kraut
prepar-ing	pre-par-ing	sca-to-log-i-cal	scat-o-log-i-cal
pre-pro-ces-sor	pre-proces-sor	schedul-ing	sched-ul-ing
pre-s-plit-ting	\pre-split-ting	schizophrenic	schiz-o-phrenic
priestesses	priest-esses	schnauzer	schnau-zer
pro-ce-du-ral	pro-ce-dur-al	schoolchild(ren)	school-child(-ren)
process	process	schoolteacher	school-teacher
procu-rance	pro-cur-ance	scy-thing	scy-thing
pro-ge-nies	prog-e-nies	semaphore	sem-a-phore
progeny	prog-e-ny	semester	se-mes-ter
pro-hib-itive(ly)	pro-hib-i-tive(-ly)	semidef-i-nite	semi-def-i-nite
prosci-utto	pro-sciut-to	semi-ho-mo-th-etic	semi-ho-mo-thet-ic
protector(s)	pro-test-er(s)	semir-ing	semi-ring
protector(s)	pro-tes-tor(s)	semiskilled	semi-skilled
pro-to-ty-pal	pro-to-typ-al	seroepi-demi-o-log-i-cal	sero-epi-de-mi-o-log-i-cal
pseu-dod-if-fer-en-tial	pseu-do-dif-fer-en-tial	ser-vomech-a-nism	ser-vo-mech-anism
pseud-ofi-nite	pseu-do-fi-nite	setup	set-up
pseud-ofinitely	pseu-do-fi-nite-ly	severely	se-vere-ly
pseud-o-forces	pseu-do-forces	sha-peable	shape-able
pseudonym	pseu-do-nym	shoestring	shoe-string
pseu-doword	pseu-do-word	sidestep	side-step
psychedelic	psy-che-del-ic	sideswipe	side-swipe
psy-chs	psychs	skyscraper	sky-scraper
pubescence	pu-bes-cence	smokestack	smoke-stack
quadrat-ics	qua-drat-ics	snorke-l-ing	snor-kel-ing
quadra-ture	quad-ra-ture	solenoid	so-le-noid
quadriplegic	quad-ri-pleg-ic	so-lute(s)	solute(s)
quain-ter(est)	quaint-er(est)	sovereign	sov-er-eign
		spaces	spa-ces

Min-nesota	Min-ne-sota
Ni-jmegen	Nij-me-gen
Noethe-rian	Noe-ther-ian
No-ord-wi-jk-er-hout	Noord-wijker-hout
Novem-ber	No-vem-ber
Poincare	Poin-care
Po-ten-tial-gle-ichung	Po-ten-tial-glei-chung
rathskeller	raths-kel-ler
Rie-man-nian	Rie-mann-ian
Ry-d-berg	Ryd-berg
schot-tis-che	schot-tische
Schrodinger	Schro-ding-er
Schwabacher	Schwa-ba-cher
Schwarzschild	Schwarz-schild
Septem-ber	Sep-tem-ber
Stokess-che	Stokes-sche
Susque-hanna	Sus-que-han-na
tech-nis-che	tech-ni-sche
Ten-nessee	Ten-nes-see
ve-r-all-ge-mein-erte	ver-all-ge-mein-erte
Verteilun-gen	Ver-tei-lun-gen
Wahrschein-lichkeit-s-the-o-rie	Wahr-schein-lich-keits-the-o-rie
Werthe-rian	Wer-ther-ian
Winch-ester	Win-ches-ter
Ying-yong Shuxue Jisuan Zeitschrift	Ying-yong Shu-xue Ji-suan Zeit-schrift

Literate Programming

**Errata: Literate Programming,
A Practitioner's View**
TUGboat 13, no. 3, pp. 261–268

Bart Childs

The address `lyman.pppl.princeton.edu` should have read `lyman.pppl.gov` (a careless error on my part). The address `csseq.cs.tamu.edu` no longer permits anonymous ftp. Due to some network breakins from different places, extensive local network changes are being done and relevant sources will be placed on `ftp.cs.tamu.edu` no later than January 1993. The author will e-mail any desired sources in the meantime.

- ◊ Bart Childs
Texas A&M University
Department of Computer Science
College Station, TX 77843-3112
`bart@cs.tamu.edu`

Philology

Hyphenation patterns for ancient Greek and Latin

Yannis Haralambous

The amount of compound words in ancient Greek makes its hyphenation by computer a quite difficult task; it is impossible to predict all combinations of words. To be efficient, a set of patterns must be accessible to the final user; a scholar must be able to add patterns, according to new words he/she encounters. Use of T_EX's `\hyphenation` primitive is not appropriate since most Greek words are declinable: for each word one would have to add a dozen hyphenation exceptions.

After a short introduction to the concept of hyphenation by T_EX the author presents a method for hyphenation of ancient Greek. Using this method, he compiled a list of patterns out of a dictionary [Bai] of 50,000 words. These patterns are presented in a comprehensible format, in a way that scholars can easily determine the patterns that have to be added, to solve specific hyphenation problems.

The same approach is applied to Latin. A list of patterns has been compiled out of a dictionary [B–C]. The size of this list is very small compared to the one of ancient Greek patterns, although Latin also uses compound words.

Finally examples of hyphenated classical texts are given.

1 What do I have to know about hyphenation?

When T_EX creates a format like Plain or LaTeX it also reads information from a file called `hyphen.tex` (or `USHyphen.tex`, or `FRHyphen.tex` and so forth) which contains the *hyphenation patterns* for a specific language. These are clusters consisting of letters separated by digits, like `x1y2z`. The idea is the following:

- if your set of patterns is empty, there is no hyphenation at all.
- if you have a pattern `x1y` then on every occurrence of the cluster “xy”, hyphenation “x-y” will be possible. If the pattern is `x1yzw`, then the pair of letters “xy” will be hyphenated only when followed by “zw”.