As with the ligature scheme, it is the user's responsibility to switch the meanings of control sequences together with the fonts.

There are 32 letters in the Russian alphabet, and only 26 in the English one. For this reason, before the \patterns command can be used in INITEX, it is necessary to extend the \uccode and \lccode tables for the letters ж, й, э, ю, ь, ъ, and я.

I use a short PASCAL program to translate GOSTCII letters into Latin letters and control sequences before it can be fed to TeX. The proposed version 3.0 of TeX will accept 8-bit input, making this preprocessor unnecessary; and its enhanced handling of ligatures and hyphenation may make it unnecessary to enter as many as 10 characters to produce a single Russian letter, when GOSTCII is not used. For Russian \language, these patterns should work with \left and \righthyphenmin=2.

There is a Bitnet mailing list dedicated to the discussion of the Russian TEX project. To subscribe, send the command

SUB RUSTEX-L (your name)

to

Bitnet: LISTSERV@UBVM
To submit an article, mail it to

Bitnet: RUSTEX-LQUBVM

(Note that node UBVM on BITNET is the same as node UBVM.CC.BUFFALO.EDU on Internet.)

Acknowledgements

I would like to thank Barbara Beeton for the truly incredible amount of help, support, and warm encouragement, all rendered via Internet; Alexander Samarin for the voluminous advice and support, again rendered via Internet; Donald Knuth, both for his interest in this project, and for creating TEX in the first place. Last but not least, I would like to thank Frank Liang for his doctoral research; without him, we would still be breaking перес-тройка.

Bibliography

- [1] Яков Грот = Yakov Grot. Русское Правописание: Руководство = Russian Orthography: a Guide.
- [2] International Organization for Standardization, Standard 8859-5: 1989, Information Processing—8-bit Single-Byte Coded Graphic Character Sets—Part 5: Latin/Cyrillic Alphabet.

- [3] В. Ф. Иванова = V. F. Ivanova. Современный русский язык = Modern Russian language. Moscow, Prosveshchenie, 1976.
- [4] Ю. Г. Павлов, В. А. Павлова,
 А. П. Соколов = Yu. G. Pavlov,
 V. А. Pavlova, А. Р. Sokolov. Алгоритм
 автоматизированного переноса русских
 слов = An algorithm for automatic hyphenation of Russian words, Institute of High
 Energy Physics, Preprint #83-72, Serpukhov,
 1983.
- [5] А. И. Кайдалова, И. К. Калинина = A. I. Кайdalova, I. К. Kalinina. Современная Русская Орфография = Modern Russian Orthography, p. 189. Moscow, Vysshaya Shkola, 1973.
- [6] Д. Э. Розентал = D. É. Rozental, ed. Современный Русский Язык = Modern Russian language, v. 5, §100. Moscow, Vysshaya Shkola, 1986.
 - Dimitri Vulis
 Department of Mathematics
 Graduate Center
 City University of New York
 33 West 42nd Street
 New York, NY 10036-8099
 Bitnet: dlv@cunyvms1

Hyphenation Exception Log

Barbara Beeton

This is the annual update of the list of words that TEX fails to hyphenate properly. The list last appeared in Volume 9, No. 3, starting on page 239. 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 American English. In the future, if such information becomes available, exceptions to other patterns will also be listed. See below, "Hyphenation for languages other than English".

In the list below, the first column gives results from TEX'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 second letter, nor after the third-from-last letter of a word; thus no word shorter than five letters will be hyphenated. (For the details, see The TeXbook, 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 TeX 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 TEX defer hyphenation until after three initial letters, some new patterns can be added, as communicated to me my Donald Knuth:

To suppress hyphenation after two letters, you need new patterns of the form .ab6 for all pairs of letters ab that begin words of English. I think the number of such pairs is well under 200.

Running PATGEN on a British, rather than a U.S., dictionary would probably result in a useful, but smaller, set of patterns, as more ambiguities might be expected in an etymologically-segmented word base. This is just a guess; I would be interested in a report on actual results, if anyone has tried it.

Hyphenation for languages other than English

Patterns now exist for many languages other than English. Michael Ferguson, the father of Multilingual TÊX (also known as MLTÊX), has volunteered to collect these hyphenation patterns for TUG, and I have volunteered him to collect exceptions to those patterns as well. If you have any patterns or exceptions to contribute, please send them to

mike@inrs-telecom.uquebec.ca
via the Bitnet.

At the top of pattern files, please include full documentation, including the name of the author and an address at which the author or other knowledgeable contact can be reached. For exceptions, please specify the patterns on which they fail, and for each word the "bad" hyphenation and the correct hyphenation; a transcript of a TEX session with \showhyphens is also useful.

I will request a report on the collection for the next issue, and exceptions will be published annually along with the (American) English list.

The List — English words

acad-e-my(ies) academy(ies) add-a-ble addable add-i-ble ad-di-ble adren-a-line adrenaline af-terthought af-ter-thought agronomist agron-o-mist am-phetamine am-phet-a-mine an-a-lyse anal-yse analy-ses* anal-y-ses anomaly(ies) anom-aly(ies) an-tideriva-tive an-ti-deriv-a-tive an-tin-o-my(ies) anti-nomy(ies) antin-u-clear an-ti-nu-clear antin-u-cleon an-ti-nu-cle-on an-tirev-o-lu-tion-ary an-ti-rev-o-lu-tion-ary apotheoses apotheo-sis ap-pendix archipelago archety-pal archetyp-i-cal assignable as-sig-nor as-sis-tantship asymp-tomatic asymp-totic asyn-chronous atheroscle-ro-sis at-mo-sphere at-tributed at-tributable avoirdupois awo-ken ban-dleader bankrupt(cy) ba-ronies base-li-neskip bathymetry bathyscaphe bea-nies be-haviour be-vies bib-li-ographis-che bid-if-fer-en-tial bil-1-able biomath-e-mat-ics biomedicine biorhythms blan-der blan-d-est blin-der blon-des blueprint bornolog-i-cal bo-tulism brus-quer bus-ier bus-i-est buss-ing but-ted buz-zword ca-caphony(ies) cam-er-a-men cartwheel catar-rhs catas-trophic

catas-troph-i-cally

cauliflower

cha-parral

chartreuse

cigarette

cognac

cog-nacs

congress

crankshaft

crocodile

cholesteric

cin-que-foil

comptroller

apoth-e-o-ses apoth-e-o-sis ap-pen-dix arch-i-pel-ago ar-che-typ-al ar-che-typ-i-cal as-sign-a-ble as-sign-or as-sist-ant-ship asymp-to-matic as-ymp-tot-ic asyn-chro-nous ath-er-o-scle-ro-sis at-mos-phere at-trib-uted at-trib-ut-able av-oir-du-pois awok-en band-leader bank-rupt(-cy) bar-onies \base-line-skip ba-thym-e-try bathy-scaphe bean-ies be-hay-iour bevies bib-li-o-gra-phi-sche bi-dif-fer-en-tial hill-able bio-math-e-mat-ics bio-med-i-cine bio-rhythms bland-er bland-est blind-er blondes blue-print bor-no-log-i-cal bot-u-lism brusquer busier busiest bussing butted buzz-word ca-caph-o-ny(ies) cam-era-men cart-wheel ca-tarrhs cat-a-stroph-ic cat-a-stroph-i-cally cau-li-flow-er chap-ar-ral char-treuse cho-les-teric cig-a-rette cinque-foil co-gnac co-gnacs comp-trol-ler con-gress crank-shaft

croc-o-dile

crosshatch(ed) dachshund database dat-a-p-ath declarable defini-tive delectable democratism de-mos deriva-tive diffract di-rer di-re-ness dis-traugh-tly dis-tribute dol-lish drif-tage driver(s) dromedary(ies) duopolist duopoly eco-nomics economist elec-trome-chan-i-cal elec-tromechanoa-cous-tic eli-tist en-trepreneur(ial) epinephrine equiv-ari-ant ethy-lene ev-ersible ev-ert(s,ed,ing) exquisite ex-traor-di-nary fermions flag-el-lum(la) flammables fledgling flowchart formidable(y) forsythia forthright freeloader friendlier frivolous ga-some-ter geodesic geode-tic ge-o-met-ric geotropism gnomon grievance grievous(ly) hairstyle hairstvlist harbinger harlequin hatcheries hemoglobin hemophilia hep-atic hermaphrodite(ic)

heroes

hex-adec-i-mal

cross-hatch(ed) dachs-hund data-hase data-path de-clar-able de-fin-i-tive de-lec-ta-ble de-moc-ra-tism demos de-riv-a-tive dif-fract direr dire-ness dis-traught-ly dis-trib-ute doll-ish drift-age dri-ver(s) drom-e-dary(ies) du-op-o-list du-op-oly eco-nom-ics econ-o-mist electro-mechan-i-cal electro-mechano-acoustic elit-ist en-tre-pre-neur(-ial) ep-i-neph-rine equi-vari-ant eth-yl-ene ever-si-ble evert(s,-ed,-ing) ex-quis-ite ex-tra-or-di-nary fermi-ons fla-gel-lum(-la) flam-ma-bles fledg-ling flow-chart for-mi-da-ble(y) for-syth-ia forth-right free-loader friend-lier friv-o-lous gas-om-e-ter ge-o-des-ic ge-o-det-ic geo-met-ric ge-ot-ro-pism gno-mon griev-ance griev-ous(-ly) hair-style hair-styl-ist har-bin-ger har-le-quin hatch-eries he-mo-glo-bin he-mo-phil-ia he-pat-ic her-maph-ro-dite(-ic) he-roes hexa-dec-i-mal

holon-omy ho-mo-th-etic horseradish hy-potha-la-mus ide-als ideographs id-i-o-syn-crasy ig-niter ig-n-i-tor ig-nores-paces impedances in-finitely in-finites-i-mal in-fras-truc-ture in-ter-dis-ci-plinary inu-tile inu-til-ity ir-re-vo-ca-ble itinerary(ies) jeremi-ads keystroke kil-ning la-ciest lamentable land-sca-per larceny(ist) lightweight limousines linebacker lines-pac-ing lithographed lithographs lobotomy(ize) lo-ges macroe-co-nomics malapropism manuscript marginal mat-tes med-i-caid medi-o-crities me-galith metabolic metabolism met-a-lan-guage metropo-lis(es) metropoli-tan mi-croe-co-nomics mi-crofiche mil-lage milliliter mimeographed mimeographs mimi-cries mi-nis min-uter(est) mis-chievously mis-ers mis-ogamy mod-elling molecule monar-chs mon-eylen-der monochrome

ho-lo-no-my ho-mo-thetic horse-rad-ish hy-po-thal-a-mus ideals ideo-graphs idio-syn-crasy ig-nit-er ig-ni-tor \ignore-spaces im-ped-ances in-fin-ite-ly in-fin-i-tes-i-mal in-fra-struc-ture in-ter-dis-ci-pli-nary in-utile in-util-i-ty ir-rev-o-ca-ble itin-er-ary(ies) je-re-mi-ads key-stroke kiln-ing lac-i-est lam-en-ta-ble land-scap-er lar-ce-ny(-ist) light-weight lim-ou-sines line-backer \line-spacing lith-o-graphed lith-o-graphs lo-bot-omy(-ize) loges macro-eco-nomics mal-a-prop-ism man-u-script mar-gin-al mattes med-ic-aid medi-oc-ri-ties mega-lith meta-bol-ic me-tab-o-lism meta-lan-guage me-trop-o-lis(es) met-ro-pol-i-tan micro-eco-nomics mi-cro-fiche mill-age mil-li-liter mimeo-graphed mimeo-graphs mim-ic-ries min-is mi-nut-er(-est) mis-chie-vous-ly mi-sers mi-sog-a-my mod-el-ling mol-e-cule

mon-archs

money-len-der

mono-chrome

mo-noen-er-getic monopole monopoly monos-pline monos-trofic mono-tonies monotonous mo-ro-nism mosquito mu-d-room mul-ti-faceted mul-ti-pli-ca-ble mul-tiuser ne-ofields none-mer-gency nonequiv-ari-ance noneu-clidean non-i-so-mor-phic nonpseu-do-com-pact non-s-mooth nore-pinephrine nutcrackeroer-st-eds oligopolist oligopoly(ies) orangutan or-thodon-tist or-thok-er-a-tol-ogy or-thoni-tro-toluene ox-i-dic painlessly pal-mate paradigm parabolic paraboloid parachute para-g-ra-pher par-ale-gal paramedic parametrize

paradimethyl-ben-zene paraflu-o-ro-toluene

para-m-ag-netism param-ethy-lanisole

paramil-i-tary paramount pathogenic pee-vish(ness) pen-tagon petroleum phe-nomenon philatelist phos-pho-ric pi-cador pi-ran-has pla-ca-ble plea-sance poltergeist polyene

mono-en-er-getic mono-pole mo-nop-oly mono-spline mono-strofic mo-not-o-nies mo-not-o-nous mo-ron-ism mos-qui-to mud-room mul-ti-fac-eted mul-ti-plic-able multi-user (better with explicit hyphen) neo-fields

non-emer-gency

non-equi-vari-ance non-euclid-ean non-iso-mor-phic non-pseudo-com-pact non-smooth nor-ep-i-neph-rine nut-crack-er oer-steds oli-gop-o-list oli-gop-oly(ies) orang-utan or-tho-don-tist or-tho-ker-a-tol-ogy ortho-nitro-toluene (or-tho-ni-tro-tol-u-ene)

ox-id-ic pain-less-ly palmate par-a-digm par-a-bol-ic pa-rab-o-loid para-chute para-di-methyl-benzene

(para-di-meth-yl-ben-zene) para-fluoro-toluene

(para-flu-o-ro-tol-u-ene) para-graph-er para-le-gal para-mag-net-ism para-medic para-methyl-anisole

(para-meth-yl-an-is-ole) pa-ram-e-trize para-mil-i-tary para-mount path-o-gen-ic peev-ish(-ness) pen-ta-gon pe-tro-le-um phe-nom-e-non phi-lat-e-list phos-phor-ic pic-a-dor pi-ra-nhas placa-ble pleas-ance pol-ter-geist poly-ene

polyethy-lene polygamist(s) poly-go-niza-tion polyphonous polystyrene pomegranate poroe-las-tic postam-ble postscript pos-tu-ral pream-ble preloaded pre-pro-ces-sor pre-s-plit-ting priestesses pro-ce-du-ral pro-cess procu-rance pro-ge-nies progeny pro-hibitive(ly) prosci-utto protestor(s)

protestor(s)
pro-to-ty-pal
pseu-dod-if-fer-en-tial
pseud-ofi-nite
pseud-ofinitely
pseud-o-forces
pseudonym
pseu-doword
psychedelic
psy-chs
pubescence
quadrat-ics
quadra-ture
quadriplegic
quain-ter(est)

quasiequiv-a-lence

quasi-hy-ponor-mal

qua-sis-ta-tion-ary

quasir-ad-i-cal

quasiresid-ual

qua-sis-mooth

qu-a-si-tri-an-gu-lar quintessence quintessen-tial rab-bi-try ra-dio-g-ra-phy raf-f-ish(ly) ramshackle ravenous re-ar-range-ment re-ciproc-i-ties reci-procity rect-an-gle ree-cho restorable ret-ri-bu-tion(s) retrofit(ted) rhinoceros righ-teous(ness)

ringleader

robot

poly-eth-yl-ene po-lyg-a-mist(s) polyg-on-i-za-tion po-lyph-o-nous poly-styrene pome-gran-ate poro-elas-tic post-am-ble post-script pos-tur-al pre-am-ble pre-loaded pre-proces-sor \pre-split-ting priest-esses pro-ce-dur-al process pro-cur-ance prog-e-nies

prog-e-ny pro-hib-i-tive(-ly) pro-sciut-to pro-test-er(s) pro-tes-tor(s) pro-to-typ-al pseu-do-dif-fer-en-tial pseu-do-fi-nite

pseu-do-fi-nite-ly pseu-do-forces pseu-do-nym pseu-do-word psy-che-del-ic psychs pu-bes-cence qua-drat-ics quad-ra-ture quad-ri-pleg-ic quaint-er(est) qua-si-equiv-a-lence qua-si-hy-po-nor-mal qua-si-rad-i-cal qua-si-resid-ual qua-si-smooth qua-si-sta-tion-ary qua-si-tri-an-gu-lar quin-tes-sence quin-tes-sen-tial rab-bit-ry ra-di-og-ra-phy raff-ish(-ly) ram-shackle

ra-di-og-ra-phy
raff-ish(-ly)
ram-shackle
rav-en-ous
re-arrange-ment
rec-i-proc-i-ties
rec-i-proc-i-ty
rec-tan-gle
re-echo
re-stor-able
ret-ri-bu-tion(s)
retro-fit(-ted)
rhi-noc-er-os
right-eous(-ness)
ring-leader
ro-bot

robotics roundtable salesclerk salescle-rks saleswoman(en) salmonella sarsaparilla sauerkraut sca-to-log-i-cal schedul-ing schizophrenic schnauzer schoolchild(ren) schoolteacher scy-thing semaphore semester semidef-i-nite semi-ho-mo-th-etic semiskilled

seroepi-demi-o-log-i-cal ser-vomech-a-nism

setup severely sha-peable shoestring sidestep sideswipe skyscraper smokestack snorke-1-ing solenoid so-lute(s) sovereign specious spelunker spendthrift spheroid(al) sph-inges spi-cily spinors

spokeswoman(en) sportscast sportively sportswear sportswriter sprightlier squeamish stan-dalone startling(ly) statis-tics stealthily steeplechase stochas-tic straight-est strangeness stratagem stretchier stronghold stupi-der(est) summable su-perego su-pere-gos

supremacist

ro-bot-ics round-table sales-clerk sales-clerks sales-woman(en) sal-mo-nel-la sar-sa-par-il-la sauer-kraut scat-o-log-i-cal sched-ul-ing schiz-o-phrenic schnau-zer school-child(-ren) school-teacher scyth-ing sem-a-phore se-mes-ter semi-def-i-nite semi-ho-mo-thet-ic semi-skilled

sero-epi-de-mi-o-log-i-cal ser-vo-mech-anism

set-up se-vere-ly shape-able shoe-string side-step side-swipe sky-scraper smoke-stack snor-kel-ing so-le-noid solute(s) sov-er-eign spe-cious spe-lunk-er spend-thrift spher-oid(al) sphin-ges spic-i-ly spin-ors

spokes-woman(en) sports-cast spor-tive-ly sports-wear sports-writer spright-lier squea-mish stand-alone star-tling(ly) sta-tis-tics stealth-ily steeple-chase sto-chas-tic straight-est strange-ness strat-a-gem stretch-i-er strong-hold stu-pid-er(est) sum-ma-ble super-ego super-egos su-prema-cist

surveil-lance sur-veil-lance swim-m-ingly swim-ming-ly symp-tomatic symp-to-matic syn-chro-mesh syn-chromesh syn-chronous syn-chro-nous syn-chrotron syn-chro-tron talkative talk-a-tive tapestry(ies) ta-pes-try(ies) tarpaulin tar-pau-lin tele-g-ra-pher te-leg-ra-pher telekinetic tele-ki-net-ic teler-obotics tele-ro-bot-ics testbed test-bed tha-la-mus thal-a-mus ther-moe-las-tic ther-mo-elas-tic times-tamp time-stamp to-po-graph-i-cal topo-graph-i-cal to-ques toques traitorous trai-tor-ous transceiver trans-ceiver transgress trans-gress transvestite trans-ves-tite traversable tra-vers-a-ble traver-sal(s) tra-ver-sal(s) treacheries treach-eries treach-ery treach-ery trou-ba-dour troubadour turnaround turn-around ty-pal typ-al unattached un-at-tached unerringly un-err-ing-ly un-friendly(ier) un-friend-ly(i-er) va-guer vaguer vaudeville vaude-ville vi-cars vic-ars vil-lai-ness vil-lain-ess viviparous vi-vip-a-rous voiceprint voice-print vs-pace \vspace wallflower wall-flower wastew-a-ter waste-water waveg-uide wave-guide web-like we-b-like weeknight week-night wheelchair wheel-chair whitesided white-sided whites-pace white-space widespread wide-spread wingspread wing-spread witchcraft witch-craft workhorse work-horse wraparound wrap-around wretched(ly) wretch-ed(-ly) yesteryear yes-ter-year

Names and non-English words used in English text

al-ge-brais-che au-toma-tisierter Be-di-enung Brow-n-ian Columbia Czechoslo-vakia al-ge-brai-sche auto-mati-sier-ter Be-die-nung Brown-ian Co-lum-bia Czecho-slo-va-kia

Dijk-stra Di-jk-stra dy-namis-che dy-na-mi-sche En-glish Eng-lish Euler-ian Eu-le-rian Febru-ary Feb-ru-ary Forschungsin-sti-tut For-schungs-in-sti-tut funk-tsional funk-t-sional Gaus-sian Gauss-ian Greifs-wald Greif-swald Grothendieck Grothen-dieck Grund-leh-ren Grundlehren Hamil-ton-ian Hamil-to-nian Her-mi-tian Her-mit-ian Jan-u-ary Jan-uary Japanese Japan-ese Kad-om-tsev Kadomt-sev Karl-sruhe Karls-ruhe Ko-rteweg Kor-te-weg Leg-en-dre Le-gendre Le-ices-ter Leices-ter Lip-s-chitz(ian) Lip-schitz(-ian) Manch-ester Man-ches-ter Marko-vian Mar-kov-ian Mas-sachusetts Mass-a-chu-setts Ni~jmegen Nij-me-gen Noethe-rian Noe-ther-ian Noord-wijker-hout No-ord-wi-jk-er-hout No-vem-ber Novem-ber Poincare Poin-care Po-ten-tial-gle-ichung Po-ten-tial-glei-chung rathskeller raths-kel-ler Rie-man-nian Rie-mann-ian schot-tis-che schot-tische Schrodinger Schro-ding-er Schwarzschild Schwarz-schild Septem-ber Sep-tem-ber

Wahrschein-lichkeit-s-the-o-rie

Stokess-che

tech-nis-che

Verteilun-gen

ve-r-all-ge-mein-erte

Wahr-schein-lich-keits-the-o-rie

ver-all-ge-mein-erte

Stokes-sche

tech-ni-sche

Ver-tei-lun-gen

Winch-ester Win-ches-ter

Yingy-ong Shuxue Jisuan Ying-yong Shu-xue Ji-suan