

dehyph-exptl*

Experimentelle Trennmuster für die deutsche Sprache

Die deutschsprachige Trennmustermannschaft

4. April 2019

Abstract

This document describes the experimental hyphenation patterns for the German language, covering traditional and reformed orthography for several varieties of Standard German. When using \XeTeX / \LuaTeX , these patterns are used automatically. In other cases, they have to be loaded explicitly. More information can be found in the [Trennmuster-Wiki¹](#) (in German).

Zusammenfassung

Dieses Dokument beschreibt die experimentellen Trennmuster für die deutsche Sprache, die das in Deutschland, Österreich und der Schweiz gebräuchliche Standarddeutsch in der traditionellen und reformierten Rechtschreibung abdecken. Bei Nutzung von \XeTeX / \LuaTeX werden diese Muster automatisch verwendet; in anderen Fällen müssen sie i. d. R. ausdrücklich geladen werden.

Warnung!

Diese Trennmuster befinden sich im experimentellen Status. Sie können auf dem CTAN und in den \TeX -Verteilungen jederzeit durch umbruchinkompatible aktualisierte Versionen ersetzt werden. Für Anwendungen, die einen dauerhaft stabilen Umbruch erfordern, sind sie nur geeignet, falls manuell eine feste Version installiert wird.

*This document describes the dehyph-exptl package vo.6.

¹<http://projekte.dante.de/Trennmuster>

Inhaltsverzeichnis		2.2.5. ConTeXt	9
1.	Einleitung	2	3. Trennregeln und Konventionen
2.	Verwenden der Trennmuster	4	3.1. Traditionelle Rechtschreibung in Deutschland und Österreich
2.1.	Sprachvarietät und Rechtschreibung	5	3.2. Traditionelle Rechtschreibung in der Schweiz
2.2.	Aktivieren der Trennmuster . .	5	3.3. Reformierte Rechtschreibung .
2.2.1.	Plain-TeX	6	4. Trennfehler
2.2.2.	L ^A T _E X mit dem Sprachenpaket Babel	7	A. Dateien und Installation
2.2.3.	X _Y L ^A T _E X/LuaL ^A T _E X mit dem Sprachenpaket Polyglossia	8	B. Fragen & Antworten
2.2.4.	p _T E _X /up _T E _X	9	

1. Einleitung

Der in T_EX implementierte Trennalgorithmus arbeitet musterbasiert [Lia83]. Prinzipiell können mit einem solchen Algorithmus nicht alle möglichen Wörter korrekt getrennt werden. Die Qualität der Worttrennung einer Sprache wird jedoch maßgeblich von der Qualität der Wortliste beeinflusst, aus der die verwendeten Trennmuster berechnet wurden.

Obwohl die herkömmlichen Trennmuster für die deutsche Sprache bei der Worttrennung in gewöhnlichen Texten eine akzeptable Fehlerrate erreichen, enthalten sie doch eine Reihe von Schwächen:²

traditionelle und reformierte Rechtschreibung

- In zusammengesetzten Wörtern treten häufig Trennfehler an Wortfugen auf.
- Fremdwörter mit akzentuierten Buchstaben werden mangelhaft getrennt: »Café«, »Ci-troën«, »Façon«, »voilà«.
- Die Trennmusterdateien enthalten eine Mischung aus T₁- sowie unvollständigen OT₁-kodierte Mustern. Mit Erscheinen von 16-Bit-fähigen T_EX-Varianten werden sauber UTF-8-kodierte Trennmuster nötig [MRo8].

²Diese Liste bezieht sich auf die Trennmusterdateien `dehypht.tex`, Version 3.2a vom 3. 3. 1999, und `dehyphn.tex`, Version 31 vom 7. 5. 2001.

traditionelle Rechtschreibung

- Die herkömmlichen Trennmuster für die traditionelle deutsche Rechtschreibung können mit patgen nicht reproduziert werden, da die zugrundeliegende Wortliste verschollen ist. Die Pflege der Trennmuster ist daher schwierig bis unmöglich. Für freie Software ist dies kein zufriedenstellender Zustand.
- Umfang und Qualität der ursprünglichen Wortliste lassen sich nicht mehr einschätzen. Für die Trennmuster in traditioneller Rechtschreibung existiert jedoch inzwischen eine Ausnahmeliste mit über 3500 korrigierten Trennungen einfacher Wörter [Lemo3, Lemo5].³
- Wird in der traditionellen Rechtschreibung *ß* durch *ss/SS* oder *sz/SZ* ersetzt, so bleibt die Trennung davon unberührt. Die herkömmlichen Trennmuster berücksichtigen diese Regel nicht und trennen häufig den Ersatz: »GRÖS-SE«, »GRÜS-SE«, »MAS-SES«.⁴
- Abweichende Schreibweisen, die in der traditionellen Rechtschreibung in Österreich und der Schweiz verwendet werden, werden mangelhaft getrennt: »Gros-son-kel«, »Kus-shand«, »Mass-nah-me«, »mis-sach-ten« (nur Schweiz) und »Ex-presszug«, »Fit-nes-scen-ter«, »Fit-nesstrai-ner«.

reformierte Rechtschreibung

- Die Trennmuster für die reformierte deutsche Rechtschreibung wurden nicht mit patgen aus einer Wortliste erstellt. Stattdessen wurden die Trennmuster für die traditionelle Rechtschreibung von Hand an die reformierten Regeln angepasst [Sch98]. Aus diesem Grund ist die Worttrennung mit den Trennmustern für die reformierte Rechtschreibung etwas schlechter als mit den Trennmustern für die traditionelle Rechtschreibung.

Das Projekt *Freie Wortlisten und Trennmuster für die deutsche Sprache* hat sich das Ziel gesetzt, neue Trennmuster hoher Qualität für die deutsche Sprache zu erstellen, die die genannten Probleme ausräumen.

³<http://mirrors.ctan.org/language/hyphenation/dehyph/dehyphtex.tex>

⁴Die Trennung der herkömmlichen Muster entspricht den Regeln der deutschen Standardsprache in der Schweiz, obwohl diese Sprachvarietät vom Paket Babel bis 2013 nicht offiziell unterstützt wurde. Babel versucht den Mangel mit Hilfe des Kürzels "S zu kompensieren.

Den experimentellen Trennmustern dieses Pakets liegt eine Wortliste mit den etwa fünfhunderttausend häufigsten Wörtern der deutschen Sprache zugrunde. Vermutlich ist diese Liste erheblich umfangreicher als die ursprüngliche Wortliste, in der Worthäufigkeiten wahrscheinlich überhaupt nicht berücksichtigt wurden. Die verwendete Wortliste deckt das in Deutschland, Österreich und der Schweiz gebräuchliche Standarddeutsch ab.

Mit den vorliegenden Trennmustern sollte für nicht-fachsprachliche Wörter eine sehr gute Trennqualität erreicht werden. Insbesondere sollte sich die Trennung häufig auftretender zusammengesetzter Wörter verbessern.

2. Verwenden der Trennmuster

Die experimentellen Trennmuster werden durch zwei verschiedene T_EX-Pakete zur Verfügung gestellt, die beide in einer vollständigen Installation einer aktuellen T_EX-Distribution enthalten sein sollten.

Das Paket `hyph-utf8` enthält Trennmuster für alle von T_EX unterstützten Sprachen. Für die deutsche Sprache sind die experimentellen Muster enthalten. Diese werden standardmäßig nur von X_YT_EX, LuaT_EX und pT_EX⁵ sowie von allen T_EX-Varianten für die Schweizer Varietät der deutschen Sprache in traditioneller Rechtschreibung⁶ benutzt. In den übrigen Fällen wird dagegen auf die herkömmlichen Trennmuster zurückgegriffen.

Das Paket `dehyph-exptl` enthält die experimentellen Trennmuster für die deutsche Sprache. Es hat den Zweck, diese den Nutzern der älteren T_EX-Varianten wie pdfT_EX zugänglich zu machen. Die hierfür nötigen Vorkehrungen werden unten beschrieben.

Siehe [Anhang B](#) für Hinweise, wie die Version der in beiden Paketen enthaltenen Muster ermittelt werden kann.

Zur manuellen Installation einer bestimmten Version der experimentellen Trennmuster siehe [Anhang A](#).

Beachte, in den folgenden Abschnitten ist `<datum>` durch das bei der Installation angegebene Datum in ISO-Notation (JJJJ-MM-TT) oder die Zeichenkette `latest` zu ersetzen.

⁵Eine in Japan populäre T_EX-Variante.

⁶Diese Varietät wird erst seit dem Jahr 2013 unterstützt, sodass keine Abwärtskompatibilität erforderlich ist.

Tabelle 1: Die unterstützten Varietäten und Rechtschreibungen

Sprachvarietät	Trennmusterbezeichner
<i>traditionelle Rechtschreibung</i>	
Deutschland, Österreich	german-x-<datum>
Schweiz	gswiss-x-<datum>
<i>reformierte Rechtschreibung</i>	
Deutschland, Österreich, Schweiz	ngerman-x-<datum>

2.1. Sprachvarietät und Rechtschreibung

Experimentelle Trennmuster stehen für die Worttrennung der deutschen Sprache in der traditionellen und der reformierten Rechtschreibung zur Verfügung. Die Trennmuster unterstützen die drei in Deutschland, Österreich und der Schweiz gebräuchlichen Hauptvarietäten der deutschen Standardsprache. [Tabelle 1](#) zeigt die Trennmuster, die abhängig von gewünschter Varietät und Rechtschreibung in einem Dokument zu aktivieren sind.

Varietäten, die sich nur in der Verwendung und Schreibung einzelner Wörter voneinander unterscheiden, können durch gemeinsame Trennmuster unterstützt werden. Zum Beispiel werden Besonderheiten der österreichischen Standardsprache in den Trennmustern für die Standardsprache Deutschlands berücksichtigt. Österreichische und deutsche Anwender können daher dieselben Trennmuster verwenden (wie das auch schon bei den herkömmlichen Trennmustern der Fall war).

Aufgrund unvereinbarer Trennregeln in der traditionellen Rechtschreibung der Standardsprachen Deutschlands/Österreichs und der Schweiz werden für die letztere Varietät eigene Trennmuster bereitgestellt.

Mit der Rechtschreibreform 1996 wurden die Trennregeln aller drei Standardsprachen so weit angeglichen, dass für die reformierte Rechtschreibung einheitliche Trennmuster für alle drei unterstützten Sprachvarietäten bereitgestellt werden können.

2.2. Aktivieren der Trennmuster

Im Folgenden wird dargestellt, wie die experimentellen Trennmuster bei Verwendung unterschiedlicher $\text{T}_{\text{E}}\text{X}$ -Varianten geladen werden können.

Tabelle 2: Trennvarianten

<i>traditionelle Rechtschreibung</i>		<i>reformierte Rechtschreibung</i>	
herkömmlich	experimentell	herkömmlich	experimentell
<u>lös</u> -te	lö-ste	lös-te	lös-te
Fas-sa-de	Fas-sa-de	F <u>assa</u> -de	Fas-sa-de
mo- <u>d</u> -ern-ste	mo-dern-ste	mo- <u>d</u> -erns-te	mo-derns-te
Abend- <u>s</u> tern	Abend- <u>s</u> tern	Abend <u>s</u> -tern	Abend- <u>s</u> tern
Mor- <u>dop</u> -fer	Mord-op-fer	Mor- <u>dop</u> -fer	Mord-op-fer

Das Ergebnis für die traditionelle und reformierte Rechtschreibung mit herkömmlichen und experimentellen Trennmustern ist in [Tabelle 2](#) zusammengefasst. Ob die experimentellen Trennmuster korrekt aktiviert werden, kann mit dem folgenden kleinen Dokument getestet werden, dessen Ausgabe in der log-Datei erscheint.

```
\begin{document}
\showhyphens{löste Fassade modernste Abendstern Mordopfer}
\end{document}
```

2.2.1. Plain-TeX

Bei Verwendung eines ϵ -TeX-basierten Compilers (etex, pdftex, xetex oder luatex) steht der Befehl `\uselanguage` zum Laden von Trennmustern zur Verfügung. Er akzeptiert neben anderen die Argumente `german`, `ngerman` und `swissgerman` (nicht jedoch `austrian`, `naustrian` und `nswissgerman`).⁷ Um die Silbentrennung gemäß reformierter deutscher Rechtschreibung zu aktivieren, ist also folgender Befehl ausreichend:

```
\uselanguage{ngerman}
```

Die Compiler `xetex` und `luatex` verwenden für alle definierten Varietäten der deutschen Sprache die experimentellen Trennmuster, die mit dem Paket `hyph-utf8` installiert sind; `etex` und `pdftex` tun dies nur für `swissgerman`.

Für die Varietäten `german` und `ngerman` laden `etex` und `pdftex` hingegen aus Gründen der Abwärtskompatibilität die herkömmlichen Trennmuster. Die experimentellen Trennmuster können am einfachsten aktiviert werden, indem etwa `\uselanguage{ngerman}` durch den folgenden Befehl ersetzt wird:

⁷Die zulässigen Sprachbezeichnungen finden sich in der Datei `language.def` der TeX-Installation.

```
\uselanguage{ngerman-x-latest}
```

Auf entsprechende Weise ist es mit allen Compilern möglich, eine alternative, manuell installierte Trennmusterversion zu verwenden:

```
\uselanguage{ngerman-x-<datum>}
```

Vorausgesetzt ist dabei, dass ein entsprechender Eintrag in der Konfigurationsdatei `language.def` oder einer lokalen Version derselben existiert.

Soll der `\uselanguage`-Befehl seine gewohnte Form behalten, ist zum Laden alternativer Trennmuster auch folgendes Vorgehen mit Hilfe des Pakets `hyphsubst` möglich:

```
\input hyphsubst.sty
\HyphSubstLet{ngerman}{ngerman-x-<datum>}
\uselanguage{ngerman}
```

Zu beachten ist, dass in Plain- \TeX unabhängig vom verwendeten Compiler Wörter mit Umlauten, β oder Akzentbuchstaben an vielen Trennstellen nicht automatisiert getrennt werden können.

2.2.2. \LaTeX mit dem Sprachenpaket Babel

Das Paket Babel kennt für die traditionelle Rechtschreibung die Sprachbezeichnungen `german`, `austrian` und `swissgerman`, für die reformierte Rechtschreibung `ngerman`, `naustrian` und `nswissgerman`, wobei diese sechs Varietäten durch drei Trennmustersätze abgedeckt werden können (Tabelle 1).

Die Compiler `xelatex` und `lualatex` verwenden für alle definierten Varietäten der deutschen Sprache die experimentellen Trennmuster, die mit dem Paket `hyphutf8` installiert sind; `latex` und `pdflatex` tun dies nur für `swissgerman`.

Für die übrigen fünf Varietäten laden `latex` und `pdflatex` hingegen aus Gründen der Abwärtskompatibilität standardmäßig die herkömmlichen Trennmuster. Um stattdessen die experimentellen Trennmuster zu aktivieren, sind für die reformierte Rechtschreibung folgende Zeilen in die Präambel des Dokuments aufzunehmen. Die Trennmusterbezeichner, die vom Paket Babel verwendet werden, werden dazu mit Hilfe des Pakets `hyphsubst` aus dem Oberdiek-Bündel so geändert, dass sie auf die experimentellen Trennmuster verweisen:

```
\usepackage[T1]{fontenc}
\usepackage[ngerman]{babel}
\usepackage[ngerman=ngerman-x-latest]{hyphsubst}
```

Für die österreichische Varietät:

```
\usepackage[T1]{fontenc}
\usepackage[naustrian]{babel}
\usepackage[naustrian=ngerman-x-latest]{hyphsubst}
```

Auf entsprechende Weise ist es mit allen \LaTeX -Compilern möglich, eine alternative, manuell installierte Trennmusterversion zu aktivieren:

```
\usepackage[ngerman]{babel}
\usepackage[ngerman=ngerman-x-<datum>]{hyphsubst}
```

Das Paket `hyphsubst` erlaubt auch das mehrfache Umschalten der Trennmuster innerhalb des Dokuments. Hinweise hierzu können der Paketanleitung entnommen werden.⁸

2.2.3. X_{\LaTeX} / $\text{Lua}\LaTeX$ mit dem Sprachenpaket Polyglossia

Für X_{\LaTeX} und $\text{Lua}\LaTeX$ steht als Alternative zu Babel das Paket Polyglossia⁹ zur Verfügung. Mit den Zeilen

```
\usepackage{polyglossia}
\setmainlanguage{german}
```

in der Präambel werden die mit dem Paket `hyph-utf8` installierten Trennmuster für die *reformierte* Rechtschreibung geladen. Mit den Optionen `variant=austrian`, `variant=swiss` und `spelling=old` können wahlweise die österreichische Varietät, die Schweizer Varietät sowie die traditionelle Rechtschreibung gewählt werden, beispielsweise:

```
\usepackage{polyglossia}
\setmainlanguage[variant=swiss, spelling=old]{german}
```

Um Trennhilfen wie `"`, `"ff` oder `"ck` verwenden zu können, muss zusätzlich die Option `babelshorthands=true` gesetzt werden.

Mit X_{\LaTeX} ist es außerdem möglich, manuell installierte statt der voreingestellten Trennmuster zu verwenden. Hierzu kommt das Paket `hyphsubst` wie folgt zum Einsatz. Für die reformierte Rechtschreibung:

```
\usepackage{polyglossia}
\setmainlanguage{german} % ggf. variant=austrian/swiss
\usepackage[ngerman=ngerman-x-<datum>]{hyphsubst}
```

⁸<https://ctan.org/pkg/hyphsubst>

⁹<https://ctan.org/pkg/polyglossia>

Für die traditionelle Rechtschreibung in Deutschland und Österreich:

```
\usepackage{polyglossia}
\setmainlanguage[spelling=old]{german} % ggf. variant=austrian
\usepackage[german=german-x-<datum>]{hyphsubst}
```

Für die traditionelle Rechtschreibung in der Schweiz:

```
\usepackage{polyglossia}
\setmainlanguage[variant=swiss, spelling=old]{german}
\usepackage[swissgerman=gswiss-x-<datum>]{hyphsubst}
```

Mit Lua^ATeX in Verbindung mit Polyglossia ist das Ersetzen der Trennmuster auf diese Weise nicht möglich.

2.2.4. pTeX/upTeX

Die in Japan gebräuchlichen TeX-Varianten pTeX und upTeX verwenden für deutschsprachige Texte die experimentellen Trennmuster, die mit dem Paket `hyph-utf8` installiert sind.

2.2.5. ConTeXt

ConTeXt besitzt eigene Trennmusterdateien. Die Trennmuster werden aus dem Paket `hyph-utf8` übernommen, jedoch zeitverzögert, sodass sie in ConTeXt im Allgemeinen in einer älteren Version vorliegen. Die Trennmuster für die traditionelle deutsche Rechtschreibung in der Schweiz fehlen zur Zeit.

Siehe [Anhang B](#) für Hinweise, wie die Version der ConTeXt-Trennmuster bestimmt werden kann.

3. Trennregeln und Konventionen

Die Trennmuster für die traditionelle Rechtschreibung in Deutschland und Österreich orientieren sich an den verbindlichen Regeln des Dudens in der Fassung von 1991 [[Wis91](#)]. Dasselbe gilt für die Trennmuster für die traditionelle Rechtschreibung in der Schweiz, jedoch mit einer unten beschriebenen Abweichung. Die Trennmuster für die reformierte Rechtschreibung orientieren sich an den amtlichen Regeln für die Rechtschreibung der deutschen Sprache in der Fassung von 2006 [[Rato6](#), [Wiso6](#)].

Tabelle 3: Unterschiedlich ermittelte Trennmöglichkeiten.

Quelle	Trennmöglichkeiten
pdf \LaTeX mit Babel 3.8	Me-tall(-l)egierung
pdf \LaTeX mit Babel 3.9	Me-tall(-l)e-gie-rung
erwünscht (vgl. Regel T6)	Me-tall(-l)egie-rung

Die Regeln lassen gewisse Freiheiten bei der Schreibung und Trennung von Wörtern zu. Da sich solche Freiheiten nicht ohne weiteres auf die maschinelle Worttrennung übertragen lassen, wurden die im folgenden beschriebenen Konventionen getroffen. Hauptsächlich betreffen diese die reformierte Rechtschreibung, die zusätzliche Freiheiten eingeführt hat.¹⁰ Beziehen sich die Konventionen für die reformierte Rechtschreibung auf die traditionelle Rechtschreibung, so werden die entsprechenden Regeln etwas ausführlicher dargestellt. Die folgenden Abschnitte enthalten jedoch keine vollständige Aufstellung der Silbentrennregeln. Diese sind den entsprechenden Regelwerken zu entnehmen. Es folgen zunächst einige allgemeine Hinweise:

- In Liangs Trennalgorithmus werden Groß- und Kleinschreibung nicht unterschieden [[Lia83](#)]. Die Schreibweisen *Nachtritt* und *nachtritt* werden aus Sicht des Trennalgorithmus gleich behandelt (siehe auch [Regel T3](#) und [Regel R5](#)).
- Die von einem Programm aus diesen Mustern abgeleiteten möglichen Trennstellen können (u. a. durch Programmfehler) durchaus von denen der zugrundeliegenden Wortliste abweichen. So führt zum Beispiel die Eingabe `Meta"llegierung` (Dreikonsonantenregel in der traditionellen Rechtschreibung) mit dem Paket Babel zu den in [Tabelle 3](#) gezeigten Trennmöglichkeiten.
- Die von \TeX gewählte Trennung kann in Einzelfällen mit den \TeX - und Babel-Kürzeln `\-` und `"-` geändert werden. Für dokumentweite Änderungen der Trennung eignet sich das Kommando `\hyphenation`.
- Die Datei `CHANGES` beschreibt bekannte, systematische Fehler der Trennmuster.

¹⁰Im Ergebnis weicht in reformierter Rechtschreibung die Trennung zum Beispiel des Dudens (nach Sprechsilben) von der Trennung mit diesen Trennmustern (bevorzugt etymologisch) ab, siehe auch [Regel R2](#) und [Regel R3](#) sowie [Anhang B](#).

- In den Beispielen zeigt die linke (grüne) Spalte jeweils die Trennung mit den experimentellen Trennmustern, die rechten (roten) Spalten zeigen alternative oder unerwünschte Trennungen.

3.1. Traditionelle Rechtschreibung in Deutschland und Österreich

T1 Die minimale unterstützte Silbenlänge am Wortanfang und -ende beträgt zwei Buchstaben [[Wis91](#), R 178].

Siehe [Anhang B](#) für Hinweise, wie sich diese Länge abweichend einstellen lässt. Wird die Mindestlänge auf weniger als zwei Buchstaben verringert, so können fehlerhafte Trennungen auftreten.

T2 Sinnentstellende und irreführende Trennungen werden möglichst vermieden [[Wis91](#), R 181] (siehe auch [Regel T6](#)):

An-alpha-bet	Anal-phabet
Kaf-ka-kenner	Kafkaken-ner
Tal-entwäs-se-rung	Talent-wässerung

Beachte, dass derzeit die Unterdrückung von solchen Trennstellen bis zu einem gewissen Grade willkürlich und subjektiv ist. Außerdem ist die Erfassung weit davon entfernt, vollständig zu sein. Um diesen Problemen abzuhelpfen, ist für zukünftige Versionen der Trennmuster ein anderer, automatisierter Ansatz geplant, welcher die Anzahl manuell zu erfassender Fälle sehr stark reduzieren wird.

T3 In mehrdeutigen Wörtern werden Trennungen nur an übereinstimmenden Trennstellen zugelassen.

nachtritt	nach-tritt	Nacht-ritt
Wachstu-be	Wach-stube	Wachs-tube
Druckerzeug-nis	Druck-erzeugnis	Drucker-zeugnis
Mu-sikerle-ben	Musik-erleben	Musi-ker-leben
Fuß-balleh-re	Fußball-ehre	Fußball-lehre

Beachte, die Trennstellen »Drucker-zeugnis« und »Musiker-leben« sind in den Interpretationen *Druck-Erzeugnis* und *Musik-Erleben* irreführend. Sie entfallen nach [Regel T2](#) und sind nicht als übereinstimmende Trennstellen anzusehen. Zur Spezialtrennung »Fußball-lehre« siehe auch [Regel T6](#).

Für diese Regel gelten die folgenden Einschränkungen:

T5 In Ableitungen von Namen auf *-ow* wird die Nottrennung der Ableitungssilben *-er*, *-ern*, *-ers* unterdrückt [Wis91, R 180]:

Tel-tower	Teltow-er
Trep-towern	Treptow-ern
Pan-kowers	Pankow-ers

T6 Spezialtrennungen (*engl.*: non-standard hyphenation), die nach Regeln erfolgen, die über das bloße Einfügen eines Trennstrichs hinausgehen, wie die *ck-* oder die Dreikonsonantenregel, kann T_EX nicht automatisch behandeln. Aus diesem Grund sind solche Trennstellen in diesen Trennmustern nicht berücksichtigt.

drucken	drucken	druk-ken
Zuckerbäcker	Zucker-bäcker	Zuk-kerbäk-ker
Brennessel	Brennes-sel	Brenn-nessel
Stoffetzen	Stoffet-zen	Stoff-fetzen

Die Dreikonsonantenregel birgt aufgrund des ausgefallenen Konsonanten die Gefahr irreführender und sinnentstellender Trennungen (siehe auch [Regel T2](#)). Trennstellen, die in einem Abstand von zwei Lauten auf eine Wortfuge mit Anwendung der Dreikonsonantenregel folgen, werden daher grundsätzlich unterdrückt.

Metalllegierung	Me-tallegie-rung	Metall-legierung	Metalle-gierung
schnellebige	schnellebi-ge	schnell-lebige	schnelle-bige
Stilleben	Stilleben	Still-leben	Stille-ben

auch:

Abfallager	Ab-fallager	Abfall-lager	Abfalla-ger
Zellstoffabrik	Zell-stoffabrik	Zellstoff-fabrik	Zellstoffa-brik

Das Paket Babel stellt verschiedene Kürzel zur Verfügung, u. a. "ck¹¹ oder "ff *etc.*, mit denen Spezialtrennungen im Quelldokument ausgezeichnet werden können (siehe auch [Tabelle 3](#)). Das Alternativpaket Polyglossia stellt diese Kürzel ebenfalls zur Verfügung, falls beim Laden der deutschen Sprache (in der Regel mit dem Befehl `\setmainlanguage`) die Option `babelshorthands=true` angegeben wird.

¹¹Spezialtrennungen werden in T_EX mit Hilfe des Kommandos `\discretionary` kodiert. So wird zum Beispiel das Babel-Kürzel "ck in der Eingabe während der Kompilation durch `\discretionary{k-}{k}{ck}` ersetzt, wodurch *k-k*-Trennungen möglich werden.

LuaTeX¹² soll in einer zukünftigen Version Mechanismen zur automatischen Behandlung von Spezialtrennungen bereitstellen. Eine physische Auszeichnung im Quelltext ist dann nicht mehr erforderlich. Die entsprechenden Spezialtrennmuster für die deutsche Sprache werden ebenfalls im Rahmen dieses Projekts erstellt.

3.2. Traditionelle Rechtschreibung in der Schweiz

Die Trennmuster für die traditionelle Rechtschreibung in der Schweiz folgen weitgehend den Konventionen für die traditionelle Rechtschreibung in Deutschland (siehe [Unterabschnitt 3.1](#)). Die folgende Liste enthält daher nur Fälle, in denen davon abgewichen wird oder deren Beschreibung aus anderen Gründen sinnvoll erscheint.

TS1 Wörter mit β werden gemäß den Regeln für die traditionelle Rechtschreibung in Deutschland getrennt.

TS2 Abweichend von [Regel T4](#) wird *ss/SS* immer als Doppelkonsonant behandelt und gegebenenfalls getrennt:

grös-ste

Grüs-se

Mas-ses

Beachte, wird β jedoch mit `\MakeUppercase` durch *SS* ersetzt, so bleibt *SS* stets ungetrennt. Die Trennung richtet sich dann nach der Schreibweise mit β im Quelldokument (siehe [Regel TS1](#)).

3.3. Reformierte Rechtschreibung

R1 Die minimale unterstützte Silbenlänge am Wortanfang und -ende beträgt zwei Buchstaben [[Rato6](#), [Wiso6](#), § 107].

Siehe [Anhang B](#) für Hinweise, wie sich diese Länge abweichend einstellen lässt. Wird die Mindestlänge auf weniger als zwei Buchstaben verringert, so können fehlerhafte Trennungen auftreten.

R2 Falls die Trennung nach Sprechsilben und die etymologische (sprachgeschichtliche) Trennung kollidieren, wird weitgehend die etymologische Trennung gewählt [[Rato6](#), [Wiso6](#), § 113]:

¹²<http://www.luatex.org/>

in-ter-es-sant inte-ressant
 Lin-ole-um Li-noleum Lino-leum
 Päd-ago-ge Pä-dagoge Päda-goge

R3 In Fremdwörtern bleiben die Buchstabengruppen *bl, pl, fl, gl, cl, kl, phl; br, pr, dr, tr, fr, vr, gr, cr, kr, phr, thr; chth; gn, kn* im allgemeinen ungetrennt, nicht jedoch *str* [Rato6, Wiso6, § 112] i. V. m. [Wis91, R 179]:

Ar-thri-tis Arth-ritis
 Di-plom Dip-lom
 igno-rie-re ig-noriere
 In-te-gral Integ-ral

aber:

In-dus-trie Indu-strie Indust-rie
 de-struk-tiv
 sub-lim

R4 Sinnentstellende und irreführende Trennungen werden möglichst vermieden [Rato6, Wiso6, § 107]:

An-alpha-bet Anal-phabet
 Kaf-ka-kenner Kafkaken-ner
 Tal-entwäs-se-rung Talent-wässerung

R5 In mehrdeutigen Wörtern werden Trennungen nur an übereinstimmenden Trennstellen zugelassen:

Druckerzeug-nis Dru-ckerzeugnis Druck-erzeugnis Drucker-zeugnis
 Mu-sikerle-ben Musi-kerleben Musik-erleben Musiker-leben
 nachtritt nach-tritt Nacht-ritt
 Wachstu-be Wach-stube Wachs-tube

Beachte, die Trennstellen »Drucker-zeugnis« und »Musiker-leben« sind in den Interpretationen *Druck-Erzeugnis* und *Musik-Erleben* irreführend. Sie entfallen nach Regel R4 und sind nicht als übereinstimmende Trennstellen anzusehen.

Für diese Regel gilt die folgende Einschränkung:

- Bei mehrdeutigen Wörtern endend auf *-ende, -enden, -endes* wird stets die Trennung der Partizipform des Verbs verwendet.

fu-ßen-de Fuß-ende fußende
 spie-len-de Spiel-ende spielende

R6 Wird der Buchstabe β durch *ss/SS* ersetzt, so wird *s-s* getrennt [Rato6, Wiso6, §§ 25 E3, 110]:

GRÖS-STE

GRÜS-SE

MAS-SES

Wird β mit `\MakeUppercase` oder in Kapitalchen in *SS* gewandelt, so bleibt *SS* ungetrennt. Dies ist kein Fehler in den Trennmustern, sondern im L^AT_EX-Kern fest implementiert.

4. Trennfehler

Mit den vorliegenden Trennmustern können sämtliche Wörter der zugrundeliegenden Wortliste fehlerfrei getrennt werden. Technisch gesprochen endet der letzte `patgen`-Lauf mit der Meldung

```
1266082 good, 0 bad, 0 missed
100.00 %, 0.00 %, 0.00 %
```

(der Wert vor `good` ist vom Listenumfang abhängig). Trotz des großen Umfangs der Wortliste lassen sich Trennfehler in Wörtern, die nicht in der Liste enthalten sind, nicht vermeiden. Der Umfang der Wortliste kann allerdings nicht beliebig erweitert werden.¹³ In den folgenden Fällen sollten fehlerhafte Trennungen der Trennmuster jedoch gemeldet werden:

A. Das Wort ist bereits in der Wortliste enthalten. Der Eintrag ist jedoch fehlerhaft.

Falls das Wort nicht in der Wortliste enthalten ist, bestehen sehr gute Chancen, dass es aufgenommen wird, wenn eines der folgenden Kriterien erfüllt ist:

B. Das betreffende Wort wird mit den *herkömmlichen* Trennmustern für die traditionelle oder reformierte Rechtschreibung korrekt getrennt. Korrekt bedeutet hier: Nicht alle möglichen Trennstellen müssen erkannt werden; es werden jedoch in keinem Fall falsche Trennstellen ermittelt. Zum Testen kann in T_EX der folgende Aufruf verwendet werden (die Ausgabe erfolgt in der `log`-Datei):

¹³Liangs Schema sieht nur einen begrenzten Bereich für die Trennstellenbewertungen vor (0–9). Die derzeitigen Trennmuster verwenden bereits Bewertungen der Höhe 8.

`\showhyphens{durch Leerzeichen getrennte Wörter}`

- C. Es handelt es sich um eine orthographisch richtige, aber sinnentstellende oder irreführende Trennung. Berücksichtigt werden allerdings nur Wörter, die aus höchstens zwei (gegebenenfalls prä- und suffigierten) Wörtern zusammengesetzt sind, zum Beispiel »Talent-wässerung«. Nicht berücksichtigt wird hingegen die »Talent-wässerungsanlage«.

Einige bekannte Fehler in den Trennmustern sind in der Datei `CHANGES` verzeichnet. Noch nicht bekannte falsche, fehlende und unerwünschte Worttrennungen können an die E-Mail-Adresse `trennmuster@dante.de` gerichtet werden.

Trennfehler, die in den Trennmustern nicht korrigiert werden können, können mit Hilfe einer privaten Ausnahmeliste behandelt werden:

```
\hyphenation{Tal-entwäs-se-rungs-an-la-ge Kaf-ka-kenner-klub}
```

Die aktuelle und ältere Ausgaben der Trennmuster sind im Dateibereich des Trennmuster-Wikis erhältlich.¹⁴ Im Entwicklerrepositorium¹⁵ befindet sich ein Makefile, mit dem jederzeit neue Trennmuster erzeugt werden können.

Happy T_EXing!

Die deutschsprachige Trennmustermannschaft

Literatur

- [Lem03] Lemberg, Werner: *Hyphenation Exception Log für deutsche Trennmuster*. Die T_EXnische Komödie, 15(2):28–31, Mai 2003.
- [Lem05] Lemberg, Werner: *Hyphenation Exception Log für deutsche Trennmuster, Version 1*. Die T_EXnische Komödie, 17(2):24–51, Mai 2005.
- [Lia83] Liang, Franklin Mark: *Word Hy-phen-a-tion by Com-put-er*. Dissertation, Stanford University, 1983. <http://www.tug.org/docs/liang/>.
- [MRo8] Miklavec, Mojca und Arthur Reutenauer: *Putting the Cork back into the bottle : Improving Unicode support in T_EX*. TUGboat, 29(3):1001–1004, 2008. <http://www.tug.org/tex-hyphen/pdf/tb93miklavec.pdf>.

¹⁴<http://projekte.dante.de/Trennmuster>

¹⁵siehe <http://projekte.dante.de/Trennmuster/Entwickler>

Tabelle 4: Trennmuster- und Manteldateien

Rechtschreibung	Trennmusterdatei	Manteldatei
traditionell	dehypht-x-<datum>.pat	dehypht-x-<datum>.tex
traditionell (Schweiz)	dehyphts-x-<datum>.pat	dehyphts-x-<datum>.tex
reformiert	dehyphn-x-<datum>.pat	dehyphn-x-<datum>.tex

- [Rato6] Rat für deutsche Rechtschreibung: *Deutsche Rechtschreibung*. <http://rechtschreibrat.ids-mannheim.de/download/regeln2006.pdf>, München, 2006.
- [Sch98] Schmidt, Walter: *T_EX und die neue deutsche Rechtschreibung*. Die T_EXnische Komödie, 10(2):35–37, September 1998.
- [Wis91] Wissenschaftlicher Rat der Dudenredaktion (Herausgeber): *Duden : Rechtschreibung der deutschen Sprache*, Band 1 der Reihe *Der Duden in 12 Bänden*. Dudenverlag, Mannheim, 20. Auflage, 1991.
- [Wiso6] Wissenschaftlicher Rat der Dudenredaktion (Herausgeber): *Duden : Die deutsche Rechtschreibung auf der Grundlage der neuen amtlichen Rechtschreibregeln*, Band 1 der Reihe *Der Duden in 12 Bänden*, Seiten 1161–1216. Dudenverlag, Mannheim, 24. Auflage, 2006.

A. Dateien und Installation

Die eigentlichen Trennmusterdateien liegen in UTF-8-Kodierung vor (siehe [Tabelle 4](#), Endung pat). Sie werden von T_EX nicht direkt geladen, sondern durch Manteldateien, die ebenfalls Teil des Pakets sind (Endung tex). Wird eine 8-Bit-fähige T_EX-Variante erkannt, übernehmen diese Manteldateien die Konvertierung der Trennmuster in die T₁-Kodierung.

Bei der Installation werden die Manteldateien an die in [Tabelle 1](#) gezeigten Trennmusterbezeichner gebunden. Diese Schritte werden für verschiedene T_EX-Verteilungen in der Datei INSTALL beschrieben. Nach der Installation können die experimentellen Trennmuster wie in [Unterabschnitt 2.2](#) gezeigt verwendet werden.

B. Fragen & Antworten

I. Verwenden der Trennmuster

1. *Wie kann für die experimentellen Trennmuster aus dem Paket `dehyph-exptl` das Datum ermittelt werden, das Teil des Trennmusterbezeichners ist, wie in Abschnitt 2 erwähnt?*
2. *Wie kann die Version der experimentellen Muster ermittelt werden, die im Paket `hyph-utf8` enthalten sind?*
3. *Wie kann die Version der experimentellen Muster ermittelt werden, die von `ConTEXt` verwendet werden?*
4. *Ich habe neue Trennmuster erstellt. Wie kann ich diese installieren?*
5. *Um in bestimmten Dokumenten einen dauerhaft stabilen Umbruch zu erreichen, will ich verhindern, dass die zur Zeit vorhandenen Trennmuster bei einer Aktualisierung der `TEX`-Installation verloren gehen. Wie gehe ich vor?*

II. Rechtschreibung

6. *Verlag, Prüfer o. ä. bemängeln die Trennung der Trennmuster für die reformierte Rechtschreibung. Zum Beispiel wird »In-dus-trie« getrennt, der Duden trennt jedoch »In-dust-rie«.*
7. *Gibt es dudenkonforme Trennmuster für die reformierte Rechtschreibung?*
8. *Weshalb werden noch Trennmuster für die traditionelle Rechtschreibung bereitgestellt?*
9. *In der Voreinstellung für die deutsche Sprache beträgt die Mindestlänge abgetrennter Silben zwei Buchstaben. Wie kann diese Mindestlänge verändert werden?*
10. *Können Trennungen mit einer Silbenlänge von nur einem Buchstaben ermöglicht werden, zum Beispiel für den Satz in schmalen Spalten?*

III. Mitarbeit

11. *Wie kann ich mich über dieses Projekt informieren?*
12. *Wie kann ich helfen?*
13. *Sollten Trennfehler einzeln oder gesammelt eingereicht werden?*

I. Verwenden der Trennmuster

1. *Wie kann für die experimentellen Trennmuster aus dem Paket `dehyph-exptl` das Datum ermittelt werden, das Teil des Trennmusterbezeichners ist, wie in [Abschnitt 2](#) erwähnt?*

Die Trennmusterbezeichner werden in der Datei `language.dat` definiert, wo auch die Verbindung zu den Manteldateien hergestellt wird. Da ein Teil der Trennmusterbezeichner bereits bekannt ist, `german-x`, vgl. [Tabelle 1](#), kann die Datei `language.dat` danach durchsucht werden. Zunächst muss der Ort der Datei mit

Hilfe des Kommandos `kpsewhich` ermittelt werden. Es folgen die vollständigen Kommandos für unixähnliche Shells und die Windows-Kommandozeile `cmd.exe` (einzugeben ohne Zeilenumbruch). Achtung, bei den einfachen Anführungszeichen handelt es sich um Gravis (»Backquotes«).

Listing 1: Shell

```
grep -i german-x `kpsewhich language.dat`
```

Listing 2: cmd.exe

```
for /F "usebackq" %f in (`kpsewhich language.dat`)  
do find /i "german-x" "%f"
```

Die Ausgabe dieser Kommandos sieht etwa wie folgt aus (das Datum kann abweichen):

```
german-x-2009-06-19 dehyph-t-x-2009-06-19.tex  
=german-x-latest  
ngerman-x-2009-06-19 dehyphn-x-2009-06-19.tex  
=ngerman-x-latest
```

Die gesuchten Trennmusterbezeichner befinden sich in der ersten Spalte und lauten in diesem Beispiel `german-x-2009-06-19` und `ngerman-x-2009-06-19`. In der zweiten Spalte kann man die Namen der Manteldateien erkennen (vgl. [Tabelle 4](#)). Die mit einem Gleichheitszeichen beginnenden Zeilen definieren ein Synonym für den Trennmusterbezeichner der unmittelbar vorangehenden Zeile in der Datei `language.dat`.

2. *Wie kann die Version der experimentellen Muster ermittelt werden, die im Paket `hyph-utf8` enthalten sind?*

Das Vorgehen ähnelt dem der vorherigen Antwort. Zunächst wird der Ort einer bestimmten Datei ermittelt. Diese wird dann nach einer hilfreichen Zeichenkette durchsucht.

Listing 3: Shell

```
grep dehyph `kpsewhich hyph-de-1901.lic.txt`
```

Listing 4: cmd.exe

```
for /F "usebackq" %f in (`kpsewhich hyph-de-1901.lic.txt`)  
do find "dehyph" "%f"
```

Die Ausgabe dieser Kommandos sieht etwa wie folgt aus (das Datum kann abweichen):

```
% dehyph-t-2011-06-01.pat
\message{German Hyphenation Patterns (Traditional Orthography)
'dehyph-t' 2011-06-01 (WL)}
```

3. *Wie kann die Version der experimentellen Muster ermittelt werden, die von ConT_EX verwendet werden?*

Das Vorgehen ähnelt dem der vorherigen Antwort. Zunächst wird der Ort einer bestimmten Datei ermittelt. Diese wird dann nach einer hilfreichen Zeichenkette durchsucht.

Listing 5: Shell

```
grep dehyph `kpsewhich lang-de.rme`
```

Listing 6: cmd.exe

```
for /F "usebackq" %f in (`kpsewhich lang-de.rme`)
do find "dehyph" "%f"
```

Die Ausgabe dieser Kommandos sieht etwa wie folgt aus (das Datum kann abweichen):

```
% dehyphn-x-2014-05-21.pat
\message{German Hyphenation Patterns (Reformed Orthography, 2006)
'dehyphn-x' 2014-05-21 (WL)}
```

4. *Ich habe neue Trennmuster erstellt. Wie kann ich diese installieren?*

Die folgende Antwort richtet sich an Nutzer von T_EX Live, die auf ihrem System über Administrationsrechte verfügen.

- i. Überprüfen Sie, ob auf Ihrem System das Verzeichnis `TEXMFLOCAL/tex/generic/config` existiert.¹⁶ Falls nicht, legen Sie es an.
- ii. Überprüfen Sie, ob im Verzeichnis `TEXMFLOCAL/tex/generic/config` die Dateien `language-local.dat` und `language-local.def` existieren. Falls nicht, legen Sie sie an.

¹⁶Bei `TEXMFLOCAL` handelt es sich um den lokalen Baum der T_EX-Installation, unter Unix in der Regel `/usr/local/texlive/texmf-local`, unter Windows `C:\texlive\texmf-local`.

iii. Ergänzen Sie in der Datei `language-local.def` die folgende Zeile:¹⁷

```
\addlanguage{<Name>}{<Datei>}{<Ausnahmedatei>}{<lmin>}{<rmin>}
```

Dabei ist `<Name>` eine selbstgewählte Bezeichnung für die Trennmuster, `<Datei>` der Name der Trennmusterdatei bzw. der Manteldatei, wenn es eine solche gibt (vgl. [Anhang A](#)), `<Ausnahmedatei>` ggf. eine Datei mit Trennausnahmen (ansonsten leer), `<lmin>` die Mindestlänge einer am Wortanfang und `<rmin>` die Mindestlänge einer Wortende abgetrennten Silbe. Für deutsche Trennmuster setzt man die beiden letztgenannten Werte typischerweise auf 2. Ergänzen Sie entsprechend in der Datei `language-local.dat` die folgende Zeile:

```
<Name> <Datei>
```

iv. Verschieben Sie Ihre Trennmusterdateien in das Verzeichnis `TEXMFLOCAL/tex/generic`.

v. Führen Sie *als Administrator* die Befehle

```
mktexlsr
```

und

```
tlmgr generate --rebuild-sys language
```

aus.

Die so installierten Trennmuster können nun den Hinweisen in [Unterabschnitt 2.2](#) entsprechend aktiviert werden.

5. *Um in bestimmten Dokumenten einen dauerhaft stabilen Umbruch zu erreichen, will ich verhindern, dass die zur Zeit vorhandenen Trennmuster bei einer Aktualisierung der T_EX-Installation verloren gehen. Wie gehe ich vor?*

Suchen Sie auf Ihrem System nach der benötigten Trennmusterdatei und der zugehörigen Manteldatei gemäß [Tabelle 4](#) und erstellen Sie Kopien dieser Dateien. Folgen Sie dann den Anweisungen der Antwort auf die vorige Frage.

¹⁷Als Vorbild für die erforderlichen Angaben können die vorhandenen Eintragungen in den Konfigurationsdateien `language.def` und `language.dat` der T_EX-Installation dienen.

II. Rechtschreibung

6. Verlag, Prüfer o. ä. bemängeln die Trennung der Trennmuster für die reformierte Rechtschreibung. Zum Beispiel wird »In-dus-trie« getrennt, der Duden trennt jedoch »In-dust-rie«.

Die amtlichen Regeln für die Rechtschreibung der deutschen Sprache lassen für viele Wörter mehrere Trennvarianten zu. Die Trennmuster und ebenso Wörterverzeichnisse legen sich aus praktischen Gründen auf eine Trennvariante fest. Sie können daher unterschiedliche Trennungen verwenden, ohne dass eine von beiden falsch ist. Aus diesem Grund sind Wörterverzeichnisse nicht geeignet, eine bestimmte Trennung auf Richtigkeit zu prüfen. Verbindlich sind einzig die amtlichen Regeln für die Rechtschreibung der deutschen Sprache [Rato6]. Häufig sind diese im Anhang eines Wörterbuchs abgedruckt. Die von den Trennmustern befolgten Konventionen können [Abschnitt 3](#) entnommen werden.

Wenn Unsicherheit darüber herrscht, wie die Rechtschreibung geprüft wird, sollte dies frühzeitig geklärt werden. Nicht jedem Redakteur oder Prüfer ist bewusst, dass der Duden seine normative Stellung mit der Rechtschreibreform 1996 eingebüßt hat.

7. Gibt es dudenkonforme Trennmuster für die reformierte Rechtschreibung?

Zur Zeit nicht, es ist auch nicht geplant. Dieses Projekt ist jedoch offen für Vorschläge und Mitarbeit.

8. Weshalb werden noch Trennmuster für die traditionelle Rechtschreibung bereitgestellt?

Die amtlichen Regeln für die Rechtschreibung der deutschen Sprache in der Fassung von 2006 sind nur für öffentliche Einrichtungen und Behörden verbindlich. Im privaten Schriftverkehr kann man wahlweise die traditionelle oder die reformierte Rechtschreibung verwenden.¹⁸ Daher erfreut sich die traditionelle Rechtschreibung weiterhin großer Beliebtheit.

Für Texte in gebrochener Schrift ist die traditionelle Rechtschreibung sogar vorzuziehen.

9. In der Voreinstellung für die deutsche Sprache beträgt die Mindestlänge abgetrennter Silben zwei Buchstaben. Wie kann diese Mindestlänge verändert werden?

Nehmen wir an, am Wortanfang sollen nicht weniger als drei und am Wortende nicht weniger als vier Buchstaben abgetrennt werden. Dann sind folgende Befehle nötig:

- Mit Plain-TeX:

¹⁸Oder auch keine von beiden.

```
\uselanguage{<sprache>}
\lefthyphenmin=3
\righthyphenmin=4
```

Ersetze dabei <sprache> durch eine der in [Unterunterabschnitt 2.2.1](#) erwähnten Trennmusterbezeichnungen, beispielsweise ngerman.

- Mit L^AT_EX und Babel:

```
\usepackage[<sprache>]{babel}
\renewcommand*{\<sprache>hyphenmins}{34}
```

Ersetze dabei <sprache> durch eine der in [Unterunterabschnitt 2.2.2](#) erwähnten Babel-Sprachbezeichnungen, beispielsweise ngerman.

- Mit X_YL^AT_EX/LuaL^AT_EX und Polyglossia:

```
\usepackage{polyglossia}
\setmainlanguage{german}
\renewcommand*{\germanhyphenmins}{34}
```

10. Können Trennungen mit einer Silbenlänge von nur einem Buchstaben ermöglicht werden, zum Beispiel für den Satz in schmalen Kolumnen?

Mit diesen Trennmustern ist das nicht möglich, da die zugrundeliegende Wortliste solche Trennungen nicht enthält. Die Abtrennung einzelner Vokale, zum Beispiel »A-bend«, war nur kurzzeitig zulässig; die entsprechende Regelung von 1996 wurde mit der Rechtschreibreform 2006 wieder zurückgenommen. Wird die Silbenmindestlänge für die Worttrennung auf eins verringert, können falsche Trennungen auftreten.

III. Mitarbeit

11. Wie kann ich mich über dieses Projekt informieren?

Zentrale Anlaufstelle ist das Trennmuster-Wiki, welches sich momentan jedoch noch im Aufbau befindet.¹⁹ Für Fragen und Hinweise kann nach Anmeldung die Mailingliste,²⁰ oder alternativ die Usenetgruppe `de.comp.text.tex` genutzt werden. Die folgenden Dateien enthalten weitergehende Informationen:

¹⁹<http://projekte.dante.de/Trennmuster>

²⁰trennmuster@dante.de

- Neben der Dokumentation zu den Trennmustern enthält dieses Paket eine Projektbeschreibung (`projektbeschreibung.pdf`).
- Die Datei `CHANGES` enthält bekannte, systematische Fehler der Trennmuster.
- Das Entwicklerrepositorium enthält in der Datei `README.wortliste` eine Beschreibung des Formats der Wortliste.

12. *Wie kann ich helfen?*

Die deutschsprachige Trennmustermannschaft ist eine offene Gruppe und benötigt dringend weitere Mithilfe. Interessenten sind daher hoch willkommen! Zum Mitmachen gibt es mehrere Möglichkeiten:

- Die einfachste ist, die experimentellen Trennmuster ausgiebig zu testen und Fehler zu melden (siehe [Abschnitt 4](#)).
- Besonders hilfreich wäre Mitarbeit am Projekt. Als Einstieg kann die (grobe) Aufgabenliste in der Projektbeschreibung dienen.

Obwohl schon experimentelle Trennmuster veröffentlicht wurden, steht dieses Projekt noch ziemlich am Anfang. Ziel ist, die verwendete Wortliste von möglichst vielen Menschen in verteilter Arbeit auf Richtigkeit zu prüfen. Bis dahin bleibt jedoch noch viel zu tun.

- Mittelfristig – nach Fertigstellung der Prüfmaske – kann auch durch systematische Durchsicht eines Teils der Wortliste geholfen werden.
- Fragen, Hinweise und Ideen sind auf der Mailingliste immer willkommen!

13. *Sollten Trennfehler einzeln oder gesammelt eingereicht werden?*

Das ist egal. Es sollten allerdings die folgenden Hinweise beim Einreichen von Fehlern beachtet werden:

Aktualität Wenn ältere Trennmusterdateien verwendet werden, bitte zuerst prüfen, ob der Fehler auch mit aktuellen Trennmustern auftritt. Die aktuellen Trennmuster sind im CTAN im Paket `dehyph-exptl` oder im Dateibereich unter der Projekt-URL erhältlich. Im Repositorium befindet sich auch ein Makefile, mit dem jederzeit neue Trennmuster aus der aktuellen Wortliste erzeugt werden können.

Informationen In einem Fehlerbericht sollte für das betreffende Wort die richtige und bei einzelnen Worteinreichungen auch die falsche Trennung angegeben werden. Außerdem die genaue Version der verwendeten Trennmuster (traditionelle oder reformierte Rechtschreibung, Datum der Trennmusterdatei).

Listenformat Es erleichtert die Korrektur, wenn Berichtigungsvorschläge in Form einer Liste eingereicht werden, die automatisch mit Skripten bearbeitet werden kann. Die folgenden Konventionen sollten dabei eingehalten werden:

- Die Spalten werden mit einem Semikolon ; getrennt.
- Die erste Spalte enthält das betreffende Wort in ungetrennter Form.
- Die zweite Spalte enthält das Wort in der Trennvariante nach traditioneller Rechtschreibung.
- Falls die Trennung nach reformierter Rechtschreibung davon abweicht, steht diese in der dritten Spalte.
- Ab der vierten Spalte *können* weitere Trennvarianten folgen, etwa die falsche Trennung. Eine kurze Erklärung sollte dann darauf eingehen.
- Unerwünschte Trennungen werden mit einem Punkt . markiert.
- Anfang und Ende der Liste sollten klar erkennbar sein.

Eine Liste könnte beispielhaft so aussehen:

```
sonnendurchfluteten;son-nen-durch-flu-te-ten  
Talentwässerung;Tal-ent.wäs-se-rung  
Fensterplatz;Fen-ster-platz;Fens-ter-platz
```

Dieses Format ist auch für einzeln eingereichte Korrekturen sinnvoll.

Betreff Der Betreff einer Fehlermeldung sollte aussagekräftig sein. Daher sollten mehrere Trennfehler, sofern möglich, thematisch zusammengefasst werden. Außerdem kann zur einfacheren Zuordnung das entsprechende Kriterium aus [Abschnitt 4](#) dem Betreff der E-Mail in eckigen Klammern vorangestellt werden, zum Beispiel:

```
Betreff: [A] sonnendurchfluteten
```

für einen Trennfehler im Wortes *sonnendurchfluteten*, der durch einen Fehler in der Wortliste hervorgerufen wird.