

%META:TOPICINFO{author="HerbertVoss" date="1503166792" format="1.1" reprev="1.17"  
version="1.17"}% %META:TOPICPARENT{name="FAQIndex"}%

# Silbentrennung, Absatz- und Seitenumbruch

- Wieso bekomme ich falsche Trennungen?
- Wie kann ich die experimentellen Trennmuster laden?
- Wie kann ich LaTeX an die neue deutsche Rechtschreibung anpassen? Gibt es Trennmuster für die neuen Trennregeln?
- Warum trennt (La)TeX Worte mit Umlauten nicht oder nicht korrekt?
- Wie kann ich Trennungsausnahmen mit Umlauten in \hyphenation verwenden?
- Wie kann ich ‚Hurenkinder‘ (engl. widows), also einzelne Zeilen eines Absatzes am Kopf von Buchseiten, bzw. ‚Schusterjungen/Waisenkinder‘ (engl. clubs), einzelne Absatzzeilen auf der Seite unten, verhindern?
- Wie muß in LaTeX die Datei „hyphen.cfg“ aussehen, um die deutschen Trennmuster „dehyphn.tex“ (vormals „ghyph31.tex“) in das LaTeX-Format einzubinden?
- Wieso werden Texte in ‚typewriter‘-Schriften nicht getrennt? Gibt es eine Möglichkeit, dass doch getrennt werden kann?
- <font>Warum trennt </font>LaTeX Wörter mit Bindestrich nicht bzw. nur am Bindestrich, auch wenn die Trennausnahmeliste diese enthält?
- Wie kann man URLs an allen Zeichen trennen?

## <a name="1" title="1"></a>Wie kann ich LaTeX an die neue deutsche Rechtschreibung anpassen? Gibt es Trennmuster für die neuen Trennregeln?

Die „alten“ Trennmuster wurden von Hand so angepasst, dass die neuen Trennregeln beachtet werden, wobei die Trennung weiterhin eher nach Wortherkunft und nicht nach strikter Silbentrennung, die ebenfalls erlaubt ist, erfolgt. Diese Trennmuster sind in der Datei „dehyphn.tex“ (vormals „gnhyph01.tex“) verfügbar, was den Anwender jedoch nicht sonderlich interessieren muss. Um diese Trennmuster nutzen zu können, benötigt man ein aktuelles Paket babel mit der Option [ngerman]. Das Paket gehört standardmäßig zu jeder TeX-Distribution.

Bei der Anwendung von **XeLaTeX** kann man statt babel das Paket polyglossia verwenden und die Sprache mit den folgenden Makros setzen:

```
\setdefaultlanguage  
\setmainlanguage  
\setotherlanguage  
\setotherlanguages
```

Die **aktuelle** Entwicklerversion von `"%%babel%%"` unterstützt alle TeX-Systeme: pdflatex, xelatex und lualatex.

<a name="9" title="9"></a>Möchte man die experimentellen Trennmuster der [Arbeitsgruppe "Trennmuster"](#) verwenden, die bislang auf einer Datenbasis von ca. 600 Tausend korrekt getrennten Wörtern basiert, so ist dies möglich mit:

```
\RequirePackage{hyphsubst}%
```

```
\HyphSubstIfExists{ngerman-x-latest}{\HyphSubstLet{ngerman}{ngerman-x-latest}}{}
\listfiles
\documentclass[ngerman]{article}
\usepackage{babel}
```

[dehyphn.tex](#) | [babel](#) | [polyglossia](#) | [dehyph-exptl](#)

## **<a name="2" title="2"></a>Warum trennt (La)TeX Wörter mit Umlauten nicht oder nicht korrekt?**

Wie TeX Wörter mit Umlauten trennt, hängt davon ab, ob die verwendete Schrift Umlaute als eigene Lettern enthält und ob man auf diese Lettern durch entsprechende Paket-Deklarationen beim Setzen eines Textes auch zugreift.

Bei den alten CM-Schriften (Computer Modern) muss TeX Umlaute aus einem Buchstaben und einem Akzent zusammenbasteln, da diese keine Lettern für Umlaute enthalten. Diese Akzent-Konstruktion irritiert TeXs Trennalgorithmus, und so trennt TeX nur den Wortteil bis zum ersten Umlaut. Im Paket ‚ngerman‘ **<a name="28" id="28"></a>** und *einigen* Babel-Versionen wird durch einen aus dem TeXbook (Anhang D, S. 394, \allowhyphens) entnommenen Trick erreicht, dass alle Wortteile zwischen den Umlauten einzeln getrennt werden. Bei dieser Vorgehensweise werden nicht alle Trennstellen gefunden oder es treten teilweise falsche Trennungen (Bsp.: “ubert-ragen statt “über-tra-gen) auf.

Abhilfe schaffen nur Schriften, die die Umlaute und andere akzentuierte Zeichen als einzelne Zeichen enthalten. Ein Beispiel hierfür sind alle T1-kodierten Schriften, darunter die EC-Schriftfamilie (siehe [Frage zu Umlauten](#)) und die T1-kodierten PostScript-Schriften. Unter neueren Versionen von LaTeX und Babel lassen sich T1-kodierte Schriften recht einfach durch die Deklaration

```
<pre class='TMLverbatim'>\usepackage[T1]{fontenc} </pre>
```

in der Präambel des Dokuments verwenden. Beispiel:

```
\usepackage[T1]{fontenc}
\usepackage{mathpazo}% Palatino mit mathematischen Zeichen
```

## **<a name="3" title="3"></a>Wie kann ich Trennungsausnahmen mit Umlauten in \hyphenation verwenden?**

Auch hier gilt, dass man in \hyphenation in den Trennungsausnahmen nur Zeichen verwenden kann, für die auch Lettern in der momentan benutzten Schrift existieren. Da mit CM-Schriften die Umlaute aus zwei Lettern zusammengesetzt werden müssen, kann man für Wörter mit Umlauten keine Trennungsausnahmen angeben. Erst bei Verwendung von T1-kodierten Schriften, wie der EC-Schriftfamilie, oder auch von Schriften in einer entsprechenden anderen Kodierung, die eigene Lettern für Umlaute enthalten, kann man Trennungsausnahmen angeben.

Hier ist ein einfaches Beispiel (‚ngerman‘ ist hier nur zum Umschalten auf die deutschen Trennmuster notwendig):

```
\documentclass{article}
\usepackage[T1]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[ngerman]{babel}
\hyphenation{Ver-stärker-aus-gang}
\begin{document}
\showhyphens{Verst\ "arkerausgang}
\showhyphens{Verstärkerausgang}
\end{document}
```

**<a name="4" title="4"></a>Wie kann ich ,Hurenkinder` (engl. widows), also einzelne Zeilen eines Absatzes am Kopf von Buchseiten, bzw. ,Schusterjungen/Waisenkinder` (engl. clubs), einzelne Absatzzeilen auf der Seite unten, verhindern?**

TeX kennt die Parameter `\widowpenalty` und `\clubpenalty`, mit denen man die ,Strafpunkte` festlegen kann, falls ein Hurenkind bzw. ein Schusterjunge beim Seitenumbruch entstehen sollte. Außerdem gibt es noch den Parameter `\displaywidowpenalty` für Hurenkinder nach einer abgesetzten mathematischen Formel. Da TeX den Wert 10000 als unendlich große Strafe interpretiert, kann man mit den Zeilen

```
\clubpenalty=10000
\widowpenalty=10000
\displaywidowpenalty=10000
```

Hurenkinder und Schusterjungen komplett unterdrücken. (Das Paket , [schuster](#)` enthält genau diese Zeilen.)

Man sollte jedoch bedenken, dass man bei einem Satzspiegel mit fester Höhe (`\flushbottom`) und registerhaltigem Satz Hurenkinder und Schusterjungen nur schwerlich verhindern kann - insbesondere wenn man dies bei zweiseitigem Druck für jede Doppelseite beachten muss. Hier ist es oft besser, die einzelne Zeile zuzulassen oder zu versuchen, einen Absatz durch geringfügige Umformulierung eine Zeile länger oder kürzer zu machen.

Weiterhin gibt es die Möglichkeit, Absätze, die schon sehr voll sind, mit `\looseness=1` um eine Zeile zu dehnen oder Absätze, die recht leer sind, mit `\looseness=-1` um eine Zeile zu kürzen. Manchmal führt dies jedoch zu merklichen Änderungen des Grauwerts des Absatzes. (Achtung: Der Wert von `\looseness` gilt außerhalb von Listen-Umgebungen nur für den aktuellen Absatz. Innerhalb einer Listen-Umgebung muss die Zuweisung zusammen mit `\par` durch eine Gruppe lokal gehalten werden `{\looseness=1 \par}`, wenn die Änderung nur für einen Absatz gelten soll.)

**<a name="5" title="5"></a>Wie muss in LaTeX die Datei „hyphen.cfg`` aussehen, um die deutschen Trennmuster „dehypht.tex`` (vormals „ghyph31.tex``) in das LaTeX-Format einzubinden?**

Eine Anleitung zur Installation der deutschen Trennmuster enthält „gerdoc``, die Kurzbeschreibung des German-Styles vom Mai 1997. Diese Anleitung beschreibt auch die Installation der Trennmuster in Plain-TeX und LaTeX 2.09. Beispiel von „hyphen.cfg`` für LaTeX.

## Grundsätzlich braucht man sich jedoch darum nicht zu kümmern, denn die gängigen TeX-Distributionen erledigen dies automatisch.

*Hinweis:* Einige Distributionen benutzen stillschweigend die Datei „hyphen.cfg“ aus dem Babel-Paket. Außerdem wurde meist „hyphen.tex“ in „ushyph1.tex“ umbenannt, um unter dem Namen „hyphen.tex“ die Datei „hyphen.cfg“ aus Babel auch für Plain-TeX einbinden zu können. In diesem Fall führt ein „\input hyphen“, wie im obigen Beispiel beschrieben, zu einer Endlosschleife!

[gerdoc.tex](#)

### **Wieso werden Texte in ,typewriter`-Schriften nicht getrennt? Gibt es eine Möglichkeit, dass doch getrennt werden kann?**

TeX trennt nur, wenn die aktuell verwendete Schrift einen Divis (Trennungsstrich) besitzt. Dazu ist jeder Schrift ein Divis durch den Parameter `\hyphenchar<font>` zugeordnet, der bei den meisten Schriften auf den Wert `\-`, also auf das Zeichen „-“, gesetzt ist. Bei den ,typewriter`-Schriften ist dieser Wert aber auf `-1` gesetzt, d.h. auf ein nicht vorhandenes Zeichen. Damit ist die Trennung für diese Schriften unterdrückt, was in den meisten Fällen auch sinnvoll ist.

Will man dennoch, dass TeX einen in ,typewriter` gesetzten Textteil automatisch trennt, so muss man den `\hyphenchar` für diese Schriften auf `\-` setzen. Dies geht in LaTeX am einfachsten, wenn man die Anweisung `\ttfamily` umdefiniert:

```
\newcommand\origttfamily{}% sollte noch nicht definiert sein!
\let\origttfamily=\ttfamily % alte Definition von \ttfamily sichern
\renewcommand\ttfamily{\origttfamily \hyphenchar\font=\-}
```

Man könnte alternativ dazu auch die Dateien „**cm**tt.fd“ ändern und die Anweisung `\hyphenchar\font\<m@ne>` in `\DeclareFontFamily` entfernen. Dies ist jedoch nicht ratsam, wenn man Texte an andere weitergeben will, da man dann nicht vergessen darf, die so geänderten „\*.fd“-Dateien mitzugeben.

Für LaTeX 2.09 (oder analog auch für Plain-TeX) muss das Makro `\tt` umdefiniert werden:

```
\newcommand\origtt{}
\let\origtt=\tt
\renewcommand\tt{\origtt \hyphenchar\font=\-}
```

Wichtiger Hinweis: TeX kann eine Schrift unter unterschiedlichen Namen laden, wobei TeX alle diese Schriften nur ein einziges Mal lädt und für die anderen Schriften nur „Aliases“ anlegt. Daher gibt es `\hyphenchar` nur einmal pro ,echten` Font und `\hyphenchar` ist wie auch `\fontdimen` ein besonderer Parameter: Alle Zuweisungen an `\hyphenchar` werden *immer* global ausgeführt. Außerdem sieht sich TeX diesen Parameter erst am Absatzende beim Umbruch dieses Absatzes an, so dass mit obiger Änderung nicht möglich ist, die Trennung für einen in ,typewriter` gesetzten Textteil zu ermöglichen, für einen anderen im selben Absatz zu verbieten.

Eine weitere Methode, Trennungen für ,typewriter`-Schriften zu erlauben, bietet das Paket ,hyphenat`:

```
\usepackage[htt]{hyphenat}
```

Darüber hinaus kann es auch benutzt werden, um Trennung selektiv zu erlauben oder zu verbieten.

[hyphenat](#)

**<a name="7" title="7"></a><font><font><font><font><font> Warum trennt LaTeX Wörter mit Bindestrich nicht bzw. nur am Bindestrich, auch wenn die Trennausnahmeliste diese enthält?</font></font></font> </font></font>**

TeX fügt nach dem aktuellen Trennzeichen implizit (fast) immer einen sogenannten `\discretionary` ein, der wie `\-` eine explizite Trennstelle angibt. Da sowohl `\-` als auch `\discretionary` dazu führen, dass die Trennung in den übrigen Wortteilen unterdrückt wird, werden solche Wörter außer am Trennstrich nicht mehr getrennt. Dieses eigenwillige Verhalten von TeX ist für Sprachen, in denen wenige Bindestriche verwendet werden, günstig und sinnvoll, da zusätzliche Trennungen in solchen Wörtern meist die Lesbarkeit erschweren.

Um Wörter mit Bindestrichen dennoch zu trennen, existieren zwei Möglichkeiten: Man gibt im Dokumenttext bei diesen Wörtern, insbesondere wo schlecht getrennt werden kann, mit `\-` oder `"`- weitere mögliche Trennstellen vor oder man verwendet `"=` anstelle des Bindestrichs. Näheres ist hierzu der Dokumentation des Paketes `,babel'` zu entnehmen.

Die andere Möglichkeit, die bislang nur von Experten umgesetzt werden kann, besteht bei Schriften, die eine zweite Letter für einen Bindestrich besitzen. Für diese Schriften kann man diese zweite Letter als Trennstrich deklarieren. Zusammen mit den zusätzlichen Trennmustern in `„hypht1.tex“` ist damit eine automatische Trennung von Wörtern mit Bindestrichen und deren Aufnahme in die Trennausnahmeliste möglich.

[babel](#) | [hypht1.tex](#)

**<a name="8" title="8"></a><font> Wieso bekomme ich falsche Trennungen?</font>**

Die häufigsten Probleme sind:

- Es wurde die falsche Sprache gewählt, beispielsweise beim Einfügen eines fremdsprachlichen Zitats ohne Wechsel der Sprache.
- Der Trennalgorithmus arbeitet wahrscheinlichkeitsbasiert und hat daher Trennfehler, die zwar weit unter 1% liegen, aber vorhanden sind. Derartige Trennungen sind dann global in der Präambel durch `\hyphenation{Trenn-re-ge\}` vorzugeben oder lokal durch `Trenn\ - re\ - ge\`

**<a name="10" title="10"></a>Wie kann man URLs bei allen Zeichen trennen?**

Normalerweise werden mit dem Paket `url` die Internetadressen nur bei bestimmten Zeichen getrennt. Man kann diese Trennung jedoch auch für alle Buchstaben ermöglichen. Dazu schreibt man in die Präambel:

```
\usepackage{url}
\def\UrlBreaks{\do\/%
\do\a\do\b\do\c\do\d\do\e\do\f\do\g\do\h\do\i\do\j\do\k\do\l\do\m\do\n\do\o\
\do\p\do\q\do\r\do\s\do\t\do\u\do\v\do\w\do\x\do\y\do\z%
\do\A\do\B\do\C\do\D\do\E\do\F\do\G\do\H\do\I\do\J\do\K\do\L\do\M\do\N\do\O\
\do\P\do\Q\do\R\do\S\do\T\do\U\do\V\do\W\do\X\do\Y\do\Z%
\do0\do1\do2\do3\do4\do5\do6\do7\do8\do9\do=\do/\do.\do: }
```

Alternativ kann man auch das Paket [xurl.sty](#) nutzen.

<font>- Main.HerbertVoss - 03 Mar 2009 </font>

From:

<https://wiki.dante.de/> - **DanteWiki**

Permanent link:

<https://wiki.dante.de/doku.php?id=dantefaq:silbentrennung>

Last update: **2023/02/28 10:39**

