

Proseminar: Wissenschaftliches Arbeiten mit \LaTeX

Mehrsprachige Texte

Erik Flick, Maurice Ossenbrink

30. Juni 2005

Gliederung

- 1 Das Babel-Package
 - Babel einsetzen
 - Was Babel bietet
- 2 Typographische Eigenarten
 - Absätze
 - French Spacing
 - Anführungszeichen
- 3 Unicode
 - Zeichensatzprobleme bei Mehrsprachigkeit
 - Repräsentation von Unicode-Zeichen
- 4 Ostasiatische Sprachen
 - Allgemeines
 - Das CJK Package

1 Das Babel-Package

- Babel einsetzen
- Was Babel bietet

2 Typographische Eigenarten

- Absätze
- French Spacing
- Anführungszeichen

3 Unicode

- Zeichensatzprobleme bei Mehrsprachigkeit
- Repräsentation von Unicode-Zeichen

4 Ostasiatische Sprachen

- Allgemeines
- Das CJK Package

Motivation

Andere Sprachen (mit Lateinschrift und verwandten Alphabeten) für \LaTeX verfügbar machen. Sonderzeichen und diacritics sind schon \sö erzeugbar, aber Babel bietet mehr.

Überblick: Was Babel bietet

- Ziemlich viele Sprachen:
- Übersetzungen der Standard-Überschriften
- Silbentrennung und Typographie
- Mehrere Sprachen in einem Text (Umgebungen, Wechsel)
- Mit allen Dok.-Klassen, inputenc, fontenc kompatibel

Das Paket einbinden

```
\usepackage[francais,german]{babel}
```

```
\documentclass[german]{article}
```

```
\usepackage{babel}
```

Sprachen, für die Babel-Module existieren

afrikaans, bahasa, breton, catalan, croatian, czech, danish, dutch, english, american, british, esperanto, estonian, finnish, francais, galician, german, magyar, irish, italian, lowersorbian, norsk, nynorsk, polish, portuguese, brazilian, romanian, russian, scottish, spanish, slovak, slovene, swedish, turkish, uppersorbian, welsh

L^AT_EX-Standardüberschriften

2. Juli 2005

- 2 juillet 2005
- July 2, 2005
- 2 a viz Gorffennaf 2005

Sprache aussuchen im Text

- `\selectlanguage{sprache}`
- Ruft mehrere Makros auf.
- Achtet nicht auf Umgebungen!

Sprache aussuchen im Text

- `\begin{otherlanguage}{sprache}`
...
`\end{otherlanguage}`
- Umgebung: Wechselt die Sprache

Sprache aussuchen im Text

- `\foreignlanguage{sprache}{blabla ...}`
- Textabschnitt in anderer Sprache gesetzt.

Umstellungsmakros einzeln

- `\otherlanguage*`
-
-

- 1 Das Babel-Package
 - Babel einsetzen
 - Was Babel bietet
- 2 **Typographische Eigenarten**
 - Absätze
 - French Spacing
 - Anführungszeichen
- 3 Unicode
 - Zeichensatzprobleme bei Mehrsprachigkeit
 - Repräsentation von Unicode-Zeichen
- 4 Ostasiatische Sprachen
 - Allgemeines
 - Das CJK Package

Absätze

Standard sind eingezogene Absätze ohne eine freie Zeile dazwischen. Wie man das hier auch schön sehen kann.

Nicht wahr?

Ja, an diesem Beispiel kann man sehr gut sehen.

Absätze

In deutschen Dokumenten teilweise eher üblich: Leerzeilen.

Für das gesamte Dokument:

```
\usepackage{parskip}
```

oder

```
\setlength{\parindent}{0pt}
```

```
\setlength{\parskip}{\medskipamount}
```

Das ganze sieht dann aus wie hier. Im Beamer Package emuliert durch die quote Umgebung.

Diese benutzt nämlich im Gegensatz zur quotation Umgebung keine Einzüge.

Absätze

Weitere Befehle für Absätze:

Unterdrücken und Erzeugen von Einzügen:

`\indent`

`\noindent`

*Manchmal wollen wir einen Einzug und manchmal
wollen wir auch
Überhaupt keinen.*

*In diesem Beispiel sieht das allerdings ein bisschen
komisch aus*

Absätze

Weitere Befehle für Absätze:

Verschiedene Abstände zwischen Absätzen:

`\smallskip`

`\medskip`

`\bigskip`

Manchmal ist es ganz gut, wenn man kleine Abstände lässt, die einem sagen:

„Hier beginnt etwas Neues!“

Aber auch ganz

neue Ideen Verlangen manchmal nach noch

Grösseren Abständen.

French Spacing

Im Englischen: Zusätzlicher Abstand nach Satzzeichen

French Spacing schaltet dies ab (bei german automatisch aktiviert)

Manuell mit:

```
\frenchspacing
```

```
\nonfrenchspacing
```

Hier ist Frenchspacing. Hier ist Frenchspacing.

*Hier ist kein Frenchspacing. Hier ist kein
Frenchspacing.*

Anführungszeichen

Normale Schreibmaschinen-Anführungszeichen werden eigentlich nicht benutzt, können aber erzeugt werden mit

`\dq`

Guck dir mal die "Gänsefüßchen" an.

Anführungszeichen

Normalerweise werden im Deutschen die deutschen Anführungszeichen verwendet.

" ' und " '

oder

`\glqq` und `\grqq`

Für einfache Anführungszeichen:

`\glq` und `\grq`

Er erzählte: „Und dann rief sie ‚Hilfe!‘ und ich rettete sie.“

Anführungszeichen

In Romanen etc. werden auch französische Anführungszeichen verwendet.

Allerdings werden diese im Französischen genau andersherum benutzt

"> und "<

oder

\frqq und \flqq

Für einfache Anführungszeichen:

\frq und \flq

Er erzählte: »Und dann rief sie ›Hilfe!‹ und ich rettete sie.«

Anführungszeichen

Im Englischen werden folgende Anführungszeichen benutzt.

“ und ”

Für einfache Anführungszeichen:

‘ und ’

Außerdem: Im englischen Anführungszeichen nach jedem Absatz.

Anführungszeichen

'The Babel fish,' said The Hitchhiker's Guide to the Galaxy quietly, 'is small yellow and leechlike, and probably the oddest thing in the universe

' ...

'The argument goes something like this: "I refuse to prove that I exist," said God, "for proof denies faith, and without faith I am nothing."

' "But", says Man, "the Babel fish is a dead giveaway isn't it? It could not have evolved by chance. It proves you exist, and so therefore, by your own arguments, you don't. QED."

' "Oh Dear," says God, "I hadn't thought of that," and promptly vanishes in a puff of logic. ... '

- 1 Das Babel-Package
 - Babel einsetzen
 - Was Babel bietet
- 2 Typographische Eigenarten
 - Absätze
 - French Spacing
 - Anführungszeichen
- 3 Unicode**
 - Zeichensatzprobleme bei Mehrsprachigkeit
 - Repräsentation von Unicode-Zeichen
- 4 Ostasiatische Sprachen
 - Allgemeines
 - Das CJK Package

Mehrsprachige Texte ohne Unicode

- Zueinander inkompatible Zeichensätze sorgen für Chaos
- Konkurrierende Codierungen und Zeichensätze
- Mehrere Sprachen in einem Text sind problematisch.

Vereinheitlichung

- Unicode hat für jedes Zeichen eine eindeutige Nummer
- Unicode ist ein allgemein akzeptierter Standard
- Unicode bietet Unterstützung für alle Schriften.

Repräsentation

Die Unicode-Zeichen werden für gewöhnlich im UTF-8 abgespeichert. 1 byte (0xxxxxx) für Codes 0-127 2 bytes (110xxxxx 10xxxxxx) für Codes 128-2047 3 bytes für Codes bis 65535 4 bytes für alle Übrigen

Anwendung in L^AT_EX

Unicode-Text können wir auch direkt eingeben.

```
\usepackage{ucs}
```

```
\usepackage[utf8]{inputenc}
```

- 1 Das Babel-Package
 - Babel einsetzen
 - Was Babel bietet
- 2 Typographische Eigenarten
 - Absätze
 - French Spacing
 - Anführungszeichen
- 3 Unicode
 - Zeichensatzprobleme bei Mehrsprachigkeit
 - Repräsentation von Unicode-Zeichen
- 4 Ostasiatische Sprachen
 - Allgemeines
 - Das CJK Package

Allgemeines

- Die drei wichtigsten Sprachen: Chinesisch, Japanisch, Koreanisch
- haben alle ihre eigenen Textkodierungen
- und eigene \LaTeX -Distributionen
- bei mehrsprachigen Texten: Unicode
- Alternative: Sprachen separat bearbeiten

Das CJK Package

Vorteile:

- alle Sprachen in einem Paket
- Unicode kompatibel
- Extras: Rubies/Furigana, Vertikale Schrift

Das CJK Package

Nachteile:

- keine weitere Lokalisation
- bei Unicode auf Unicode-Fonts beschränkt
- funktioniert scheinbar nicht richtig mit beamer package

Verwendung

Deklaration

```
\usepackage[encapsulated]{CJK}
```

Im Dokument

```
\begin{CJK}{UTF8}{cyberbit}  
\end{CJK}
```

Kurze Info zur koreanischen Sprache

- verwendet eigene Buchstabenschrift: Hangeul
- heutzutage kaum chinesische Zeichen (in NK überhaupt keine)
- Schreibrichtung wie bei uns

Koreanisches Beispiel

Als erstes Beispiel, hier ein kurzer koreanischer Text in der Silbenschrift Hangul.

```
\begin{CJK}{UTF8}{cyberbit}
hier steht der koreanische Text
\end{CJK}
```

한글 배우기가 어렵지않아요.

Auf Deutsch: „Hangul lernen ist nicht schwer“

Kurze Info zur japanischen Sprache

- verwendet 2 japanische Silbenschriften (Hiragana, Katakana) in Kombination mit chinesischen Zeichen
- Schreibrichtung entweder wie bei uns oder traditionell von oben nach unten, rechts nach links
- Lesung kann über dem Zeichen angegeben werden (Furigana/Rubies)

Japanisch mit Furigana

Deklarierung

```
\usepackage[overlap, CJK]{ruby}
```

Im Dokument

```
\ruby{zu lesendes Zeichen}{Lesung}
```

Japanisches Furigana Beispiel

z.B. ergibt `\ruby{日本語}{にほんご}` folgendes: 日本語^{にほんご}

Das selbe funktioniert auch mit lateinischen Lettern:

これは^{Japanisch}日本語ですか？ -え？！^{ドイツ語}Deutschでしょう！

Japanisch in vertikaler Schriftweise

Deklaration

```
\usepackage{CJKvert}
```

- dreht alle Zeichen um 90 Grad
- Problem: passt nicht in ein horizontales Dokument
- Lösung: In Textbox setzen und diese wiederum um 90 Grad drehen

Japanisch in vertikaler Schriftweise

dafür benutzt man z.B. das Rotating Package:

```
\usepackage{rotating}
```

Als Beispiel ein berühmtes Haiku von Basho:

```
\begin{turn}{-90}  
\parbox{3cm}{  
\begin{CJK}{UTF8}{cyberbit}  
hier steht das Haiku  
\end{CJK}  
\end{turn}
```


Japanisch in vertikaler Schriftweise

古池や
蛙とびこむ
水の音

In der deutschen Übersetzung:

Der alte Teich.

Ein Frosch springt hinein -

das Geräusch des Wassers