

Neue Möglichkeiten zur inhaltlichen Textauszeichnung beim Satz mit TUSTEP

Im Bericht über den TEI-Workshop in BI 96/3+4, S. 10, wurde eine neue Möglichkeit der inhaltlichen Textauszeichnung für das Satzprogramm kurz angedeutet, die hier etwas detaillierter vorgestellt werden soll. Ihre Verwirklichung ist ein weiterer kleiner Schritt zu einer besseren Unterstützung von SGML-konformer Textauszeichnung durch TUSTEP, die nach dem Bericht über »Künftige Software zur Textanalyse« in der letzten BI (96/5+6, S. 24–26) erforderlich ist.

Zum besseren Verständnis der neuen Möglichkeiten sollen vorab einige wichtige Punkte in der Arbeitsweise des Satzprogramms knapp erläutert werden.

Ergebnis-Dateien des Satzprogramms

Neben der Ausgabe-Datei, die das eigentliche Ergebnis des Satzlaufes darstellt (nämlich die Steuerinformation für den Belichter oder den Laserdrucker in noch weitgehend geräteunabhängiger Form), wird beim Satz mit TUSTEP im Unterschied zu vielen anderen Satzprogrammen eine weitere Datei erzeugt, die sog. Ziel-Datei. Diese Datei enthält (analog zur entsprechenden Datei bei FORMATIERE) noch alle Steueranweisungen für den Satz, ist also z. B. nach einer Korrektur wieder als Eingabe-Datei für einen weiteren Satzprogramm-Lauf verwendbar. Andererseits enthält sie schon die durch den Satz erzeugte Zeilen- und Seiteneinteilung des gesetzten Textes und damit die Information, die für alle Zwecke der Weiterverarbeitung wichtig ist, bei denen (z.B. in Registern) Verweise auf die Seiten- und Zeileneinteilung des mit Hilfe des Satzprogramms publizierten Werkes eine Rolle spielen.

Steuerung des Satzablaufs

Beim Satz von Texten mit TUSTEP muß wie bei jedem anderen Satzprogramm die typographische Gestaltung im Detail festgelegt werden. Dazu sind zwei Arten von Angaben notwendig:

1. Auftragsbezogene Angaben (vergleichbar den »style sheets« vieler Formatierprogramme), mit deren Hilfe die für einen ganzen Text gültigen typographischen Eigenschaften festgelegt werden: Satzspiegel, verwendete Schriften, Schriftgröße und Durchschuß für die einzelnen Textteile (Grundtext, Einschaltungen, Fußnoten, Überschriften usw.), Art und Stel-

lung von Kolummentitel und Seitennumerierung usw. Diese Angaben werden dem Programm mit Hilfe von Parametern mitgeteilt.

2. Textbezogene Angaben, mit deren Hilfe der Text untergliedert (z. B. in Abschnitte oder Kapitel eingeteilt) und an bestimmten Stellen typographisch modifiziert (z. B. durch Kursivschrift hervorgehoben) wird. Diese Angaben sind auf unterschiedliche Weise möglich.

Textauszeichnung durch elementare und übergeordnete Steueranweisungen

Für jedes typographische Detail, das ausgeführt werden muß, gibt es deshalb Steueranweisungen (z. B.: Seitenwechsel, wenn nicht noch mindestens 8 Zeilen auf der Seite frei sind; zwei Leerzeilen; Schriftgrößenwechsel auf 10 Punkt). Für häufig vorkommende »Ereignisse« im Text (z. B. Zwischenüberschriften) stehen übergeordneten Steueranweisungen zur Verfügung, in denen solcher Einzelangaben zusammengefaßt sind (z. B. `&& &&{` für »Überschrift der Stufe 2«, vor der, wenn über Parameter nichts anderes verlangt wird, die soeben als Beispiel genannten Einzelanweisungen ausgeführt werden und hinter der eine weitere Leerzeile eingefügt und der Schriftgrößenwechsel rückgängig gemacht wird).

Diese übergeordneten Anweisungen stellen einen ersten Schritt in Richtung »sachliche« bzw. »inhaltliche« Auszeichnung des Textes dar: Statt der typographischen Einzelbefehle stehen im Text die übergeordneten Steueranweisungen, für deren Ausführung es Voreinstellungen gibt, die aber über Parameter abgeändert werden können.

Textauszeichnung mit Hilfe von Makros

a) Bisherige Form von Makros: `&.ab`

Einen weiteren Schritt in der Abstraktion von der typographischen Textauszeichnung und in Richtung auf inhaltliche bzw. sachlich orientierte Textauszeichnung stellt die Verwendung von Makros dar, die für das Satzprogramm bisher durch `&.ab` kodiert werden, wobei für `ab` zwei beliebige Zeichen außer Blank stehen dürfen. Diesen Makros werden erst bei der Nutzung des Satzprogramms über Parameter typographische Codes (die eigentlichen Satzprogramm-Steueranweisungen) zugeordnet. Es

ist es also möglich, die inhaltliche Markierung (z. B. zur Kennzeichnung der verschiedenen Kategorien wie Autor, Titel, Erscheinungsjahr in den einzelnen Einträgen einer Bibliographie) bis zum Aufruf des Satzprogramms beizubehalten.

Ein Nachteil dieser Form von Makros war und ist, daß sie – von in diesem Zusammenhang uninteressanten Ausnahmen abgesehen – in der Ziel-Datei des Satzprogramms nicht mehr als solche erkennbar sind. An ihre Stelle ist dort die typographische Kodierung durch die Satzprogramm-Steueranweisungen getreten. Dabei sind möglicherweise unterschiedliche Makros durch identische typographische Kodierungen ersetzt, da inhaltlich unterschiedliche Kategorien typographisch gleich dargestellt werden sollen. Die in den Makros enthaltene Information bezüglich der Struktur des Textes ist damit gerade in der Datei nicht mehr greifbar, die für alle Zwecke der Weiterverarbeitung wichtig ist, bei der Verweise auf die Seiten- und Zeileneinteilung des mit Hilfe des Satzprogramms publizierten Werkes eine Rolle spielen.

Dieser Nachteil kann weitgehend vermieden werden, wenn die neue Form von Makros benutzt wird, die in künftigen TUSTEP-Versionen zusätzlich zur bisherigen enthalten sein wird (und im Vorgriff darauf bereits auf den zentralen Servern des ZDV zur Verfügung steht).

b) Neue Form von Makros: <name>

In seiner neuen Form bestehen ein Satzprogramm-Makros aus einer bis zu 40 Zeichen langen, in spitze Klammern eingeschlossenen Zeichenfolge, die selbst aus allen Zeichen außer Blank und spitzen Klammern bestehen darf.

Diese neue Form von Makros stellt in verschiedener Hinsicht eine deutliche Erweiterung der bisherigen Möglichkeiten dar:

– Sie hebt sich vom Text allein durch ihre

graphische Form deutlich besser ab als die mit &.ab kodierte Makros.

- Sie erlaubt mit bis zu 40 Zeichen langen Makronamen eine mnemonisch bessere Kodierung.
- Anfang und Ende von Einheiten lassen sich leichter sichtbar kennzeichnen, z. B. durch die Verwendung des für die Anfangsmarkierung benutzten Namens, dem ein Schrägstrich vorangestellt wird, auch für das Ende (Beispiel: <zitat>Dies ist ein Zitat</zitat>).
- Sie entspricht weitgehend der in SGML vorgesehenen Kodierung.
- Diese Makros bleiben (sofern ihre Auflösung durch Steueranweisungen keine Blanks enthält) auch in der Ziel-Datei des Satzprogramms erhalten.

Der letzte Punkt ist in unserem Zusammenhang der wichtigste. Denn auf diese Weise ist die in diesen Makros enthaltene Information über die inhaltliche Struktur des Textes auch in der Datei noch voll erhalten, die neben den Steuerzeichen auch die Information über die endgültige Anordnung des Textes enthält. Diese Strukturinformation steht also auch für die Weiterverarbeitung eines schon gesetzten Textes viel bequemer als bisher zur Verfügung.

Die neue Regelung hat darüber hinaus einen weiteren durchaus nützlichen Nebeneffekt: Der gesetzte Text bleibt in der Zieldatei sehr viel übersichtlicher als bisher in all den Fällen, in denen einem Makros eine längere Sequenz von Steueranweisungen zugeordnet werden muß. Daß die Makros in der neuen Form in der Zieldatei erhalten bleiben, erleichtert die Weiterverwendung solcher Dateien auch dann erheblich, wenn die entsprechende Folge von Steueranweisungen eindeutig einer einzigen inhaltlichen Markierung zugeordnet werden kann.

Wilhelm Ott
ott@zdv.uni-tuebingen.de