

Wie sollen Überschriften markiert werden?

Wolfgang Schmidle

18. März 2010

In diesem Text geht es zum einen um die Frage, ob Überschriften als Absätze mit speziellem Typ (`<p type="head">`) oder mit einem eigenen Element (`<head>`) markiert werden sollen. These 1: Überschriften sind keine Absätze und sollten daher mit `<head>` markiert werden. Zum anderen geht es um die Frage, ob Überschriften `<s>` enthalten sollen oder nicht. These 2: Überschriften sollten eigentlich kein `<s>` enthalten; aus pragmatischen Gründen könnte man beides erlauben, um verschiedenen Workflows gerecht zu werden.

Inhaltsverzeichnis

1 Überschriften und Absätze	2
1.1 Merkmale von Überschriften und Absätzen	2
1.2 Historisch	2
1.3 Wie machen es andere?	3
1.3.1 HTML	3
1.3.2 TEI	3
1.3.3 Open XML	3
1.3.4 OpenDocument	3
1.4 Text-Container in Arboreal	3
1.4.1 MPDL-Schema	4
1.4.2 Archimedes	4
2 <code><s></code> in Überschriften	5
2.1 Überschriften im Archimedes-Korpus	5
2.1.1 Texte mit den Kategorien C, T, I, F	6
2.1.2 Nicht-triviale Fehler	6
2.1.3 Interessante Fälle	6
2.1.4 Verbesserungen im Archimedes-Korpus	9
2.2 Überschriften im MPDL-Korpus	9
2.2.1 Benedetti	10
2.3 Technische Argumente	11
2.4 Pragmatische Gründe	12

1 Überschriften und Absätze

These: Überschriften sind keine Absätze und sollten daher nicht als Absätze markiert werden.

1.1 Merkmale von Überschriften und Absätzen

Überschriften und Absätze unterscheiden sich funktional, inhaltlich und optisch. Überschriften sind daher keine "Absätze mit besonderer Funktion".

Ein Absatz besteht aus einem oder mehreren ganzen Sätzen. Ein Überschrift besteht normalerweise aus einem unvollständigen Satz ("Capitulum primum de proportione et eius diuisione.") bzw. einer Passiv-Konstruktion, oder aus reinen Struktur-Angaben ("THEOREMA CXLI."). Im Gegensatz zu Absätzen gehören Überschriften zu verschiedenen Hierarchie-Ebenen.

Überschriften sind optisch meist leicht erkennbar, zum Beispiel sind sie oft zentriert. Die Chinesen können Überschriften und Absätze rein optisch mit einer hohen Trefferquote unterscheiden und deshalb aufgrund der Optik semantisch markieren.

1.2 Historisch

Ob Überschriften historisch mit Absätzen vergleichbar sind oder nicht, wäre noch zu prüfen.

Michael Giesecke, *Der Buchdruck in der frühen Neuzeit*, schreibt: "Die Ersetzung der 'Incipits' und der 'wundersamen' Benennungen von Manuskripten durch Titelblätter und Überschriften wird später noch genauer nachzugehen sein" (p.324), bezieht sich dann aber im Abschnitt *Titelblatt und Paginierung: Die Informationen bekommen eine Adresse* (p.420ff) nicht mehr direkt auf Überschriften. Die Aussage "Soll [über das Titelblatt] hinaus auf einzelne Argumente eines Autors Bezug genommen werden, so muss auch der Text selbst sequenziert werden und die einzelnen Sequenzen müssen Adressen erhalten" (p.421, 424) bezieht sich offenbar nur auf Seitenzahlen.

Natürlich sollte man nicht die Eigenschaften des früheren Mediums kritiklos in das neue Medium zu übertragen. Andersherum sollte man aber auch typographisches Wissen nicht ignorieren, wenn man dadurch nicht etwas erreicht, was man sonst nicht erreichen könnte. Ein Beispiel ist die Neuordnung der Textflows im Eipo-Text oder in den Conimbricenses, durch die man im Medium XML einen echten Mehrwert hat. Die Markierung von Überschriften als Absätze hat keinen vergleichbaren Mehrwert.

1.3 Wie machen es andere?

1.3.1 HTML

HTML hat `<h1>` bis `<h6>`, und `<p>`.

1.3.2 TEI

TEI hat `<head>` und `<p>`:

"`<head>` (heading) contains any type of heading, for example the title of a section, or the heading of a list, glossary, manuscript description, etc." (TEI P5, p. 971)

"`<p>` (paragraph) marks paragraphs in prose." (TEI P5, p.1081)

1.3.3 Open XML

Open XML ist das XML-Datenformat von MS Word. Dort wird offenbar rein optisch markiert: Überschriften sind Absätze mit einem bestimmten "style". Beispiel:

```
<w: p>  
  <w: r>  
    <w: rPr>  
      <w: rStyle w: val="heading 1"/>  
    </w: rPr>  
    <w: t>Überschrift</w: t>  
  </w: r>  
</w: p>
```

`<w: r>` steht dabei für "run", dem eine Formatierung zugeordnet werden kann.

1.3.4 OpenDocument

OpenDocument ist das XML-Datenformat von OpenOffice. Dort wird zwischen Überschriften und Absätzen semantisch unterschieden: `<text: h>` für Überschriften und `<text: p>`.

1.4 Text-Container in Arboreal

Arboreal kann mit allen Möglichkeiten, Überschriften zu markieren, umgehen. In den docspecs kann man Container definieren, die Text enthalten.

1.4.1 MPDL-Schema

Für das MPDL-Schema gibt es folgende Container:

- s
- head
- caption, description, variables (Unterelemente von figure)
- note, handwritten, xhtml (i.e. Listen und Tabellen) mit inlinemodel

In den docspecs sieht das so aus:

```
<containers>
  <element name="s"/>
  <element name="head"/>
  ...
</containers>
```

Die Elemente des letzten Punktes sind definiert als Pseudoelemente, z.B.

```
<pseudoelement name="note.inlinemodel">
  <element name="note"/>
  <attr-value name="xml:space" value="preserve"/>
</pseudoelement>
```

Außerdem markiert Arboreal Überschriften auf der linken Seite mit §, wenn die XML-Anzeige ausgeschaltet ist. In den docspecs:

```
<self>
  <element name="head"/>
  <render-before>&#x00a7; </render-before>
  <!-- section sign -->
</self>
```

Beispielsweise sind im lateinischen und chinesischen Euklid (Clavius und Ricci) Überschriften als `<head>` markiert und enthalten keine `<s>`.

1.4.2 Archimedes

Beim Archimedes-Format sind die Text-Container `<s>` und `<cell>`. Der Container `<cell>` markiert eine Tabellen-Zelle als Unterstruktur von `<table>` und `<row>` und kommt in `alber_reaed_004_la_1485.xml`, `agric_remet_001_la_1556.xml` und vielen weiteren Archimedes-Texten vor. Es ist also nicht so, dass in Archimedes nur `<s>` Text enthalten kann.

Für Überschriften gibt es das Pseudo-Element "heading":

```
<pseudoelement name="heading">
  <element name="p"/>
  <attr-value name="type" value="head"/>
</pseudoelement>
```

2 <s> in Überschriften

These: Überschriften sollten kein <s> enthalten. Aus rein pragmatischen Gründen, insbesondere um die Korrektur von falschen <p> und <head> zu vereinfachen, kann man eventuell analog zu <note> beides erlauben.

2.1 Überschriften im Archimedes-Korpus

Ich habe alle 110 Archimedes-Texte, die im MPDL-Prototypen enthalten sind, auf Überschriften mit mehr als einem <s> durchgeschaut. Das Ergebnis ist, dass es praktisch keine solchen Fälle gibt.

Es folgt eine meines Wissens vollständige Liste der Fälle. Ich verwende dabei vier Kategorien:

- C für Fälle wie "Di quali cofe è composta l'Architettura. Cap. II.", also

Di quali cofe è composta l'Architettura. Cap. II.

oder

Di quali cofe è composta l'Architettura.
Cap. II.

Hier muss man klären, wie man damit umgehen möchte. Ich denke nicht, dass Cap. II. ein getrenntes <s> ist, und auch in Archimedes wurde offenbar keine bewusste Entscheidung getroffen, sondern es ist das Artefakt eines Skriptes. Beachte, dass es auch Situationen gibt, wo zwei Überschriften direkt hintereinander kommen, aber sowohl optisch als auch semantisch klar getrennt sind und daher nicht zu einer einzigen Überschrift gemacht werden können. Ich diskutiere Kategorie C weiter unten.

- T für Titelseiten: Die Markierungen auf den Titelseiten sind nicht standardisiert, und um Titelseiten geht es hier auch nicht.
- I falsche Verwendung in Index, Inhaltsverzeichnis, etc.: Offensichtliche Fehler in diesen Strukturen.
- F für andere offensichtliche Fehler. Die häufigsten Fehler sind: Punkte als </s> interpretiert, paragraph als heading markiert, heading nicht korrekt beendet.

2.1.1 Texte mit den Kategorien C, T, I, F

cered_disco_021_it_1567.xml	F
comma_centra_023_la_1565.xml	F
jorda_ratio_049_la_1565.xml	T
monte_mecha_036_la_1577.xml	T
carda_propo_015_la_1570.xml	T F
borro_demot_011_la_1575.xml	T F C I
monte_mecha_037_it_1581.xml	T (": " als idiosynkratische Strukturierung)
varro_demot_044_la_1584.xml	T
monan_mecha_035_la_1599.xml	T
stell_mecha_041_it_1597.xml	T F (u.a. Appendice. I.)
gheta_promo_029_la_1603.xml	T F
gueva_mecha_005_la_1627.xml	I
lorin_forti_031_it_1609.xml	C (merkwürdig formatiertes XML)
valer_centra_043_la_1604.xml	T
caste_mensu_068_en_1661.xml	T (auch "kleine Titelblätter" im Text) F
fosca_epist_067_en_1661.xml	T (auch Kolophon)
galil_doctr_066_en_1661.xml	F
galil_syste_065_en_1661.xml	T
salus_mathe_040_en_1667.xml	T (auch Kolophon und "kleine Titelblätter") F
casat_mecha_017_la_1684.xml	T I F (paragraph as head)
lucre_natur_082_la_1893.xml	F
agric_remet_002_en.xml	F (und <28> im Text)
desca_princ_081_la_1644.xml	F I

2.1.2 Nicht-triviale Fehler

1. bianc_locam_009_la_1615.xml C T F

p.323: Tabellen-Erläuterung falsch als Überschrift markiert:

Singula porrò fecula ex 100. annis constant. Anno huius feculi 76.
vel ante Vrb. cond. 24. Olympiades initium fumpferunt.

2. schre_qiqit_X03_zh_1627.xml: Chinesischer Text, viele leere <s> (als Artefakte des Satzzeichens "。"?), Paragraphs oft falsch als heading markiert.

2.1.3 Interessante Fälle

(Und das sind die einzigen interessanten Fälle in 110 Texten!)

1. agric_remet_001_la_1556.xml

p.513: Überschrift des Index:

REI METALLICAE
NOMINA LATINA GRAECA-
QVE GERMANICE REDDITA, ET EX
ordine, quo quodque primo occurrit, collocata. Re-
liqua fubiuncta funt libris, De natura foffi-
lium infcriptis: quorum pauca quæ-
dam repetere nunc habui
neceffe.

**Alternative: Hier geht die Überschrift (unvollständiger Satz) optisch nett ge-
setzt in einen Erklärungsabsatz (vollständiger Satz) über.**

2. tarta_quesi_042_la_1554.xml I

(siehe auch it.wikisource.org/wiki/Quesiti_et_inventioni_diverse)

p.050:

Quesito vigesimoterzo fatto da M. Alberghetto di Alberghetti gettadore de artegliaria
l'anno 1545 adi Aprile. In Venetia.

(Der Text ist ziemlich schlecht markiert, alle anderen markierten Beispiele
sind Kategorie F.)

Wenn man wirklich will, kann man In Venetia. als eigenes <s> markieren.

3. vitru_archi_045_it_1567.xml C F I

p.380

Della ragione de i Gnomoni ritrouati per l'ombra
da i raggi del Sole. Et del mondo. Et de i
pianeti. Cap. IIII.

Eher eine Regel wie "verwende kein Komma in der Überschrift" als echte
<s>

4. heron_pneum_112_la_1575.xml C

p.73:

ALITER. Poteft autem & aliter idem (hier ein falscher <lb/>) effici.

Wenn man wirklich will, kann man ALITER. als eigenes <s> markieren.

5. stevi_weegc_085_nl_1586.xml F

(p.54: Ein Absatz nach Cap II ist größer geschrieben, aber als normaler
<p> markiert, wie die ersten Absätze im Clavius, die wir bewusst nicht als
headings markiert haben.)

6. buona_demot_014_la_1591.xml C F

auf mehr als 1000 Seiten insgesamt drei Fälle:

p.167:

Explicatur habitudo actionis ad agens. & quædam corollaria cum nonnullarum
dubitationum folutionibus apponuntur. Cap. X.

p. 173:

Difputatur an fuceffio pertineat ad omnem fpeciem motus, ac problema proponitur.
Vtrum generatio puncto temporis obeatur nec'ne. Cap. XII.

p. 773:

Exploratur vis eius effati. Quod vere eft nulli accidit. Cap. XVIII.

Dies wären Beispiele für echte, wenn auch kurze, <s>. Aber dann muss man Cap. X. etc. ebenfalls in ein eigenes <s> tun, weil es sich sonst nur auf das jeweils letzte der vorhergehende <s> bezieht. Und dann müsste man es, um einheitlich zu sein, wiederum überall machen.

(außerdem p.289 etc.: "q." bewirkt falsches <s>)

7. baldi_mecha_007_la_1621.xml

p.38

QVÆSTIO II.

Cur, fi surfum libræ fulcimentum fit, appofito ad alteram partem pondere, defcendat libra, & eo amoto, iterum afcendat, & ad æquilibrium reuertatur. Si verò deorfum fulcimentum fuerit, depreffa ad æquilibrium non reuertatur?

Aber p.159/160: Quaestio XXIII, dann normaler Absatz, der mit

Quod autem hîc quærit Ariftoteles, ita fe habet.

endet, dann erst Absatz in kursiv/zentriert.

Paul: Es ist die Propositio der Quaestio. Wiedergeben mit <p type="title">. (Hier passt tatsächlich ein Attribut type="title", denn es ist tatsächlich ein Absatz mit einer besonderen Funktion.)

8. alber_archi_003_en_1755.xml

p.100

BOOK V. CHAP. I.

Of Buildings for particular Perfons. Of the Castles or Habitations of a King or a Tyrant; their different Properties and Parts.

sowie weitere Beispiele, die allerdings im Text nicht so markiert wurden, z.B.

p.81

BOOK IV. CHAP. I.

Of Works of a publick Nature. That all Buildings, whether contrived for Neceffity, Conveniency or Pleafure, were intended for the Service of Mankind. Of the feveral Divifions of humane Conditions, whence arifes the Diverfity of Buildings.

(dort ist nur BOOK IV. CHAP. I. als Überschrift markiert.)

Unterüberschriften oder Minimal-Zusammenfassungen?

Paul: wieder <p type="title">

2.1.4 Verbesserungen im Archimedes-Korpus

Ein Nebenergebnis ist, dass die Aufbereitung der Texte teilweise ziemlich schlecht ist. Die Texte sind sehr uneinheitlich markiert, und es gibt viele Skript-Artefakte wie zum Beispiel leere `<s>`. Insbesondere das Skript, das entscheiden sollte, ob ein Punkt das Ende eines `<s>` bedeutet oder nicht, hat viele Fehlentscheidungen in beide Richtungen produziert.

Sind diese Texte denn die aktuellen Versionen? Paul: ja.

Paul: Es wäre sinnvoll, die Texte zu überarbeiten. Zum Beispiel leere `<s>` raus. Das ist ein Problem, wenn es bereits ein matching file in Arboreal gibt, aber mindestens für Texte ohne matching file ist es wohl kein Problem. (Eventuell noch Vokabellisten in Arboreal, die IDs verwenden.)

Vielleicht doch wenigstens einige Texte in das neue Schema überführen?

2.2 Überschriften im MPDL-Korpus

Der MPDL-Korpus besteht zurzeit aus Benedetti, Clavius, Ricci und dem Eipo-Text. Ich gehe im Folgenden auf Benedetti ein.

Grundsätzlich habe ich die Markierungen aus China übernommen, mit wenigen Korrekturen. Den Fall C habe ich daher bisher so markiert:

```
<head>Di quali cofe è composta l'Architettura. Cap. II.</head>
```

bzw.

```
<head>Di quali cofe è composta l'Architettura.</head>  
<head>Cap. II.</head>
```

Das ist nicht einheitlich. Wenn man es einheitlich haben möchte, könnte man den zweiten Fall so markieren:

```
<head>Di quali cofe è composta l'Architettura. <lb/>Cap. II.</head>
```

Allerdings gibt es Fälle, in denen tatsächlich Überschriften direkt aufeinanderfolgen, die man weder optisch noch semantisch zu einer einzigen Überschrift machen kann. Man kann also Gruppen von aufeinanderfolgenden Überschriften nicht vermeiden, und sie sind auch kein besonderes Problem.

2.2.1 Benedetti

Im folgenden gehe ich einige Fälle von mehreren `<head>` in Folge in Benedetti durch. Fälle, wo die Überschriften offensichtlich auf verschiedenen Hierarchie-Ebenen sind, habe ich weggelassen. (Meistens sind sie dann durch ein `<div>` getrennt.)

Ich halte keines der Beispiele für ein Argument für `<s>` in `<head>`.

- p.13: Kategorie T

```
<head>IO. BAPTISTAE <lb/>BENEDICTI <lb/> ... <lb/>PHILOSOPHI.</head>
<head style="it">Theoremata Arithmetica.</head>
```

Optisch zwei klar getrennte Teile (Autor und Titel), inhaltlich (Autor im Genitiv) eine einzige Überschrift. Wenn man es als eine Überschrift markieren möchte:

```
<head>IO. BAPTISTAE ... PHILOSOPHI. <lb/><emph style="it">Theoremata ...</emph></head>
```

- p.153 etc.: Kategorie C

```
<head style="it">De differentia fitus brachiorum libra.</head>
<head>CAP. I.</head>
```

müsste dann werden zu:

```
<head><emph style="it">De differentia ... libra.</emph> <lb/>CAP. I.</head>
```

Ich habe es aber erstmal so gelassen, weil es optisch klar getrennt ist.

Im Inhaltsverzeichnis möchte man wohl beides haben. Und man will insbesondere nicht davon abhängig sein, wie und in welcher Reihenfolge es im Text vorkommt (d.h. CAP. I. vor oder nach De differentia ...).

- p.210: Entscheidung: Erklärungsabsatz in kursiv

```
QVamuis omnia libri quinti Euclid. uerißima fint. ...
```

...

ist keine Überschrift. Hier passt wieder `<p type="title">`.

und:

```
<head>TERTIVM.</head>
<head style="it">Qüē est euclidis feptima propofitio.</head>
```

- p.216: Kategorie T (kleines Titelblatt)

```
<head>PHYSICA, <lb/>ET MATHEMATICA <lb/>RESPONSA.</head>
<head style="it">FO. BAPTISTAE BēNēDICTI PATRITII</head>
<head style="it">Veneti, Philofophi Mathematici.</head>
```

geändert zu:

```
<head>PHYSICA, <lb/>ET MATHEMATICA <lb/>RESPONSA.</head>
<head style="it">FO. BAPTISTAE ... <lb/>Veneti, Philofophi Mathematici.</head>
```

- p.217: wieder T, aber im Text aus China nicht sinnvoll aufgeteilt. Neu:

```
<head>DE TEMPORVM ... BENEDICTI <lb/>Patritij Veneti, Philofophi <lb/>Mathematici.</head>
<head style="it">AD SERENISS. CMANVELEM PHILIB. <lb/>Allobrogum ... Inuictißimum.</head>
<head>EPISTOLA.</head>
```

Es sind zumindest optisch drei getrennte Teile. Wenn die Punkte am Ende nicht wären, könnte man es inhaltlich als eine einzige Überschrift deuten.

- Überschriften in Kapitel 6: Dieses Kapitel besteht aus Briefen. Dabei gibt es immer Paare von Titel und Adressat. Beispiel p.297:

```
<head>DE MOTV MOLAE ...</head>
<head style="it">Illust. Ioanni Paulo ...</head>
...
<head style="it">De rèuolutione rota putealis ...</head>
<head>AD EVNDEM.</head>
...
<head style="it">De machina ...</head>
<head>AD EVNDEM.</head>
...
```

Jeder Brief ist ein `<div type="letter">`, und ein oder mehrere aufeinanderfolgende Briefe an denselben Adressaten bilden zusammen ein `<div type="section">`.

2.3 Technische Argumente

Zum Argument des Programmieraufwandes für einen Programmierer, der in fünf Jahren unsere Texte bearbeiten will und die Situation vorfindet, dass nicht jedes Textstück in einem `<s>` ist, weil es kein `<s>` in `<head>` gibt und in `<note>` sogar entweder `<p>` oder `<s>` oder direkt Text enthalten sein kann:

- Es gibt ein ganz einfaches Kriterium, ob ein Element direkt Text enthält oder nicht, nämlich ob es das Attribut `xml:space="preserve"` hat. Ein XSLT-Template, das jedes Textstück und nichts anderes findet, sieht dementsprechend so aus:

```
<xsl:template match="*[@xml:space='preserve']">
...
</xsl:template>
```

- Wir haben in section 1.4.2 gesehen, dass auch in den Archimedes-Texten nicht jedes Textstück in einem `<s>` ist.

Zum Argument der Verschachtelungstiefe, dass es `<s>` in `<head>` geben sollte, damit der Text in `<head>` und `<p>` die gleiche Verschachtelungstiefe hat:

1. Ich halte dieses Argument für "un-XML". Jedes `<div>` verändert die Verschachtelungstiefe. Eine `<note>` in einem `<div float>` hat eine andere Verschachtelungstiefe als eine `<note>`, die nicht in einem `<div float>` ist. Wenn eine `<note>` in `<s>` nicht ganz aus dem `<p>`, sondern nur aus dem `<s>` verschoben wird, ändert sich die Verschachtelungstiefe der `<note>`.
2. Wie wir bereits in section 1.4 gesehen haben, gibt es in Arboreal keine Gründe für oder gegen `<s>` in Überschriften.

3. TEI hat <s> (Kapitel 17.1). Beispiel p.1153:

```
<head>
  <s>A short affair</s>
</head>
<s>When are you leaving?</s>
<s>Tomorrow.</s>
```

Offenbar ist es TEI nicht wichtig, dass alle <s> die gleiche Verschachtelungstiefe haben.

(Laut diesem Beispiel ist <s> in TEI auch in <head> erlaubt. Allerdings glaube ich nicht, dass das für unsere Situation ein Argument ist, denn in TEI sind wiederum Strukturen erlaubt, die wir nicht erlauben wollen. TEI ist quasi eine Tag-Soße, die über den Text geschüttet wird. Das Beispiel wirkt so, als ob die <s> bereits da waren und dann ein <s> zusätzlich zu <head> erklärt wurde; dafür gibt es bei uns kein Gegenstück.)

2.4 Pragmatische Gründe

Ein einfacher Workflow ist folgender: Jeder Absatz wird mit <p> markiert und mit <s> versehen. Die weitere Bearbeitung (IDs vergeben, etc.) erfolgt in Arboreal. Das Schema soll mit diesem Workflow funktionieren.

Das bedeutet aber nicht, dass alles erlaubt sein muss, was durch dieses Vorgehen entsteht. Beispielsweise sind leere <p>, die bei der Konversion aus Word-Dokumenten entstehen können, zu Recht verboten. Arboreal stört sich nicht daran, sodass man problemlos mit diesem Text arbeiten kann. Aber wenn man ein vollständig Schema-konformes Dokument erstellen will, muss man ein (triviales) Korrektur-Skript verwenden. Die Belohnung ist, dass die Textqualität verbessert wird.

1. Einfacher Fall:

Di quali cofe è composta l'Architettura.

in Archimedes:

```
<p type="main" xml:id="N001">
  <s xml:id="N002">Di quali cofe è composta l'Architettura.</s>
</p>
```

wird zu

```
<p type="head" xml:id="N001">
  <s xml:id="N002">Di quali cofe è composta l'Architettura.</s>
</p>
```

In MPDL:

```
<p xml:id="N001">
  <s xml:id="N002">Di quali cofe è composta l'Architettura.</s>
</p>
```

wird zu

```
<head xml:id="N002">Di quali cofe è composta l'Architettura.</head>
```

Es wird also `<p>` entfernt und `<s>` in `<head>` umbenannt. Beachte, dass der Text immer noch die richtige ID hat.

2. Schwierigerer Fall: Wir haben gesehen, dass es praktische keine Fälle gibt, in denen eine Überschrift wirklich aus mehreren `<s>` besteht. Ich gehe daher in diesem Beispiel von einer falsch aufgeteilten Überschrift aus, nämlich:

Appendice. I.

In Archimedes:

```
<p type="main" xml:id="N001">
  <s xml:id="N002">Appendice. </s>
  <s xml:id="N003">I. </s>
</p>
```

wird zu

```
<p type="head" xml:id="N001">
  <s xml:id="N002">Appendice. </s>
  <s xml:id="N003">I. </s>
</p>
```

In MPDL:

```
<p xml:id="N001">
  <s xml:id="N002">Appendice. </s>
  <s xml:id="N003">I. </s>
</p>
```

wird zu

```
<head xml:id="N002">Appendice. I. </head>
```

In beiden Fällen kann man die Änderung wohl nicht in Arboreal durchführen, sondern sinnvollerweise in einem XML-Editor. Aber genauso wie bei leeren `<p>` hat hier die Forderung, dass eine Überschrift nicht aus mehreren `<s>` besteht, die Funktion der Fehlerkontrolle: Wenn man ein `<head>` mit mehr als einem `<s>` hat, ist sehr wahrscheinlich etwas verkehrt. Meistens ist es ein Punkt, der eigentlich kein Satzpunkt ist.